

## Collecter des données sur l'évolution du patient pour augmenter son efficacité thérapeutique en orthophonie

T. Martinez Perez\*, M. Geurten\*\* et S. Willems\*\*\*

\*Unité de Recherche Enfances

\*\*Unité de Recherche PsyNCog

\*\*\*Clinique Psychologique et Logopédique Universitaire

Faculté de Psychologie, Logopédie et des Sciences de l'Education  
Université de Liège, Belgique

Article publié en 2021 dans le périodique « *ANAE : Approche Neuropsychologique des Apprentissages chez l'Enfant* »

### Matériel supplémentaire : tutoriel pour les analyses

Par rapport aux différents cas de figure envisagés dans l'article « Collecter des données sur l'évolution du patient pour augmenter son efficacité thérapeutique en orthophonie », plusieurs analyses (Gage & Lewis, 2013 ; Vannest & Ninci, 2015) peuvent être réalisées pour déterminer si les résultats d'un patient se sont améliorés au cours de l'intervention ou entre deux périodes de temps dites phases (intervention/absence d'intervention, pré/post-intervention, intervention/stagnation des progrès) :

1. Les analyses visuelles
2. Le pourcentage de non-chevauchement
3. La statistique Tau
4. L'erreur type de mesure
5. Le test de McNemar

Nous nous attacherons ici à détailler le principe général de chaque analyse et comment les réaliser. Les avantages et les limites de chacune d'entre elles seront également développés plus longuement.

#### 1. L'ANALYSE VISUELLE

Le principe de l'analyse visuelle est extrêmement simple (Gage & Lewis, 2013). Il s'agit de réaliser une représentation graphique de la performance obtenue par le patient lors des différentes évaluations effectuées avant, durant ou après l'intervention. Il est ensuite possible d'examiner visuellement si :

- a. La performance a tendance à diminuer, augmenter ou rester stable au fur et à mesure de l'intervention.
- b. La performance moyenne obtenue lors d'une phase est inférieure, égale ou supérieure à la performance moyenne obtenue lors d'une seconde phase.

##### 1.1. Comment faire ?

Voyons cela à l'aide d'un exemple. Le pourcentage de productions correctes des phonèmes K/T chez le patient, en l'absence d'intervention puis lors de la reprise de l'intervention (monitoring envisagé sous l'angle du design ABAB, Figure 2c dans l'article de référence), sont repris dans le tableau ci-dessous (Tableau 1).

Absence d'intervention		Reprise de l'intervention	
Semaine 6	57	Semaine 11	55
Semaine 7	61	Semaine 12	65
Semaine 8	58	Semaine 13	83
Semaine 9	48	Semaine 14	91
Semaine 10	57	Semaine 15	88

		Semaine 16	94
Moyenne	56,2	Moyenne	79,3

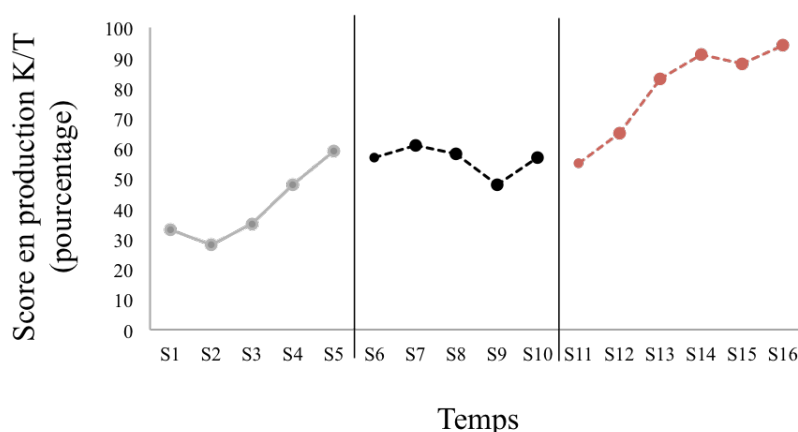
**Tableau 1.** Résultats du patient en l'absence d'intervention sur la cible puis lors de la reprise de l'intervention sur cette cible.

Afin de mener une analyse visuelle, deux étapes doivent être réalisées : le dessin du graphique et l'analyse de celui-ci.

### ÉTAPE 1 - DESSIN DU GRAPHIQUE

Pour toute analyse visuelle, l'abscisse représente la variable « temps », c'est-à-dire le moment où les données sont collectées (ici, les données ont été prises en semaine 6 à 10 pour la mesure d'absence d'intervention et en semaine 11 à 16 pour la mesure de reprise de l'intervention). L'ordonnée, quant à elle, représente les scores obtenus par le patient pour le comportement considéré dans l'analyse (ici, les pourcentages de productions correctes varient entre 48 et 61 pour la mesure d'absence d'intervention et entre 55 et 94 pour la mesure de reprise de l'intervention). La Figure 1 représente ces différentes données. Notons qu'il est important de respecter l'échelle des axes (p. ex. un axe de 0 à 100 pour une mesure de pourcentage) car une modification des échelles pourrait entraîner des interprétations opposées (en amplifiant ou en diminuant les effets).

**Figure 1.** Représentation graphique des données pendant l'intervention (en gris), en l'absence d'intervention (en noir) et lors de la reprise de l'intervention (en rouge).



### ÉTAPE 2 - ANALYSE DU GRAPHIQUE

En analysant ce graphique, on constate :

- Que la moyenne des performances en l'absence d'intervention (moyenne = 56,2) est inférieure à la moyenne des performances lors de la reprise de l'intervention (moyenne = 79,3)
- Compte tenu de la courbe, il semble que la performance du patient lors de l'intervention tend à s'améliorer au fur et à mesure des séances.

Face à un tel profil, le clinicien peut conclure à une possible efficacité de l'intervention.

#### 1.2. Avantages des analyses visuelles

Ce type d'analyse est intuitif, rapide et facile à réaliser. Aucun outil statistique particulier n'est nécessaire, il suffit de tracer un graphique pour représenter les données visuellement (ceci peut être effectué à l'aide d'un logiciel de type « Excel » ou simplement à la main). Ce graphique pourra ensuite être présenté au patient pour discuter avec lui de l'évolution de sa performance.

#### 1.3. Inconvénients des analyses visuelles

Comme leur nom l'indique, les analyses visuelles reposent uniquement sur un examen visuel de la performance du patient. Ainsi, dans les cas où une différence de petite taille est observée entre deux phases, il est extrêmement difficile de déterminer, sur base de ce type d'analyse, si le changement observé est significatif d'une évolution réelle dans le comportement du patient ou s'il s'agit d'une simple variation due au hasard. Afin de pouvoir conclure avec certitude à l'efficacité de l'intervention, des statistiques spécifiques doivent être réalisées.

## 2. LE POURCENTAGE DE NON-CHEVAUCHEMENT

Le pourcentage de non-chevauchement (Non-overlap of All Pairs (NAP); Parker & Vannest, 2009) est une technique statistique non paramétrique permettant d'estimer le non-chevauchement entre deux phases (p. ex. en absence versus en présence de l'intervention). Le NAP est approprié pour presque tous les types de données et distributions, y compris les données dichotomiques. En outre, il peut être calculé manuellement à partir de petits ensembles de données.

### 2.1. Comment faire ?

Reprenons l'exemple des résultats obtenus par le patient lorsqu'il n'y avait pas d'intervention sur la production correcte du K/T puis lorsque l'intervention a repris sur cette cible (Tableau 1).

Absence d'intervention		Reprise de l'intervention	
Semaine 6	57	Semaine 11	55
Semaine 7	61	Semaine 12	65
Semaine 8	58	Semaine 13	83
Semaine 9	48	Semaine 14	91
Semaine 10	57	Semaine 15	88
		Semaine 16	94

*Tableau 1. Résultats du patient en l'absence d'intervention sur la cible puis lors de la reprise de l'intervention sur cette cible.*

Pour calculer le pourcentage de non-chevauchement, chaque score d'une première phase doit être comparé à chaque score obtenu lors d'une seconde phase afin de déterminer, pour chaque score obtenu lors d'une première phase, s'il est inférieur, égal ou supérieur à chacun des scores obtenus lors d'une seconde phase. Chaque fois que le score de la phase 1 est inférieur à celui de la phase 2, on parle de non-chevauchement entre les deux phases. Chaque fois que le score de la phase 1 est égal ou supérieur à celui de la phase 2, on parle de chevauchement. Une fois tous les scores comparés deux à deux, il convient de compter le nombre de non-chevauchement entre les deux phases et de diviser le résultat obtenu par le nombre total de comparaisons réalisées.

Dans le cas présent, nous allons donc comparer deux par deux les scores de la semaine 6 à 10 (absence d'intervention) aux scores de la semaine 11 à 16 (reprise de l'intervention). Nous compterons ensuite le nombre de non-chevauchements avant de le diviser par le nombre de comparaisons réalisées.

Semaine 6

57 < 55 = chevauchement

57 < 65 = non-chevauchement

57 < 83 = non-chevauchement

57 < 91 = non-chevauchement

57 < 88 = non-chevauchement

57 < 94 = non-chevauchement

Semaine 7

61 < 55 = chevauchement

61 < 65 = non-chevauchement

61 < 83 = non-chevauchement

61 < 91 = non-chevauchement

61 < 88 = non-chevauchement

61 < 94 = non-chevauchement  
 Semaine 8  
 58 < 55 = chevauchement  
 58 < 65 = non-chevauchement  
 58 < 83 = non-chevauchement  
 58 < 91 = non-chevauchement  
 58 < 88 = non-chevauchement  
 58 < 94 = non-chevauchement  
 Semaine 9  
 48 < 55 = non-chevauchement  
 48 < 65 = non-chevauchement  
 48 < 83 = non-chevauchement  
 48 < 91 = non-chevauchement  
 48 < 88 = non-chevauchement  
 48 < 94 = non-chevauchement  
 Semaine 10  
 57 < 55 = chevauchement  
 57 < 65 = non-chevauchement  
 57 < 83 = non-chevauchement  
 57 < 91 = non-chevauchement  
 57 < 88 = non-chevauchement  
 57 < 94 = non-chevauchement

On compte donc un total de 26 non-chevauchements sur trente comparaisons réalisées :  $26/30 = 0.86$ .  
 Nous avons donc ici un pourcentage de non-chevauchement de 86%.

Afin de ne pas avoir à calculer ce score manuellement (ce qui peut être laborieux pour les designs impliquant un grand nombre de mesures) et afin de déterminer si le pourcentage de non-chevauchement entre une première phase et une seconde phase est bien statistiquement significatif, un calculateur est disponible en ligne à l'adresse : <http://www.singlecaseresearch.org/calculators/nap>  
 À l'aide de ce calculateur, réaliser l'analyse est alors très simple. Cela peut être décomposé en trois étapes.

### ÉTAPE 1 - ENCODAGE

Il s'agit d'encoder les résultats obtenus par le patient lors de la première phase (ici, absence d'intervention) dans une première colonne, puis d'encoder les résultats obtenus par le patient lors de la seconde phase (ici, reprise de l'intervention) dans une deuxième colonne (Figure 2).

*Figure 2. Étape d'encodage pour calculer un pourcentage de non-chevauchement.*

### ÉTAPE 2 - CALCUL DE LA STATISTIQUE

Une fois les résultats encodés, le calcul du NAP et de la valeur p associée est très rapide. Pour le réaliser, il suffit de sélectionner les deux colonnes que vous souhaitez comparer (ici l'absence

d'intervention et la reprise de l'intervention) en cochant la case en haut à gauche de chaque colonne concernée, puis d'appuyer sur le bouton « Contrast » (Figure 3).

Figure 3. Étape de calcul du pourcentage de non-chevauchement.

### ÉTAPE 3 - INTERPRÉTER

Une fois la statistique calculée, reste encore à l'interpréter. Si le pourcentage de non-chevauchement est élevé et que la valeur p obtenue est inférieure à .05, vous pouvez considérer que la différence observée entre les deux phases est significative et que votre prise en charge a été suffisante pour induire un changement significatif dans le comportement du patient. Dans le cas de notre exemple, le pourcentage de non-chevauchement (NAP) est de 86% et la valeur p est de .04 (Figure 4). Nous pouvons donc supposer que l'intervention réalisée sur le patient s'est montrée efficace, en tenant compte des recommandations de Byiers et al. (2012) qui impliquent d'observer des différences significatives entre 4 phases (et pas uniquement entre deux phases comme dans cette illustration).

#### Results

id	Label	S	PAIRS	NAP	VARs	SD	SDnap	Z	P Value	CI 85%	CI 90%
0	Absence vs Reprise	22	30	0.8667	120	10.9545	0.3651	2.0083	0.0446	0.208<>1	0.133<>1

Figure 4. Étape d'interprétation du pourcentage de non-chevauchement.

Remarque importante : lorsque l'amélioration de la performance se marque par une diminution des scores du patient (p. ex. l'intervention vise à une diminution du nombre d'erreurs) plutôt que par une augmentation des scores (p. ex. l'intervention vise à une augmentation du pointage), le pourcentage de non-chevauchement doit être inversé à l'aide de la formule suivante : 1-NAP. La valeur p reste, cependant, identique.

### 2.2. Avantages des analyses NAP

Les points forts du NAP sont sa simplicité, sa réflexion sur l'analyse visuelle et sa puissance statistique.

### 2.3. Inconvénients des analyses NAP

Par contre, le NAP ne peut être utilisé que lorsque la phase pré ou post-intervention est stable (c'est-à-dire, quand les scores obtenus par le patient lors d'une phase où il n'y a pas d'intervention ne varient pas de manière disproportionnée). Il est toutefois possible de s'assurer que ces phases sont stables à l'aide du calculateur en ligne. Pour ce faire, il convient simplement de comparer la phase pré ou post-intervention avec elle-même, puis de s'assurer que le pourcentage de non-chevauchement obtenu à la suite de cette comparaison est faible (Figure 5).

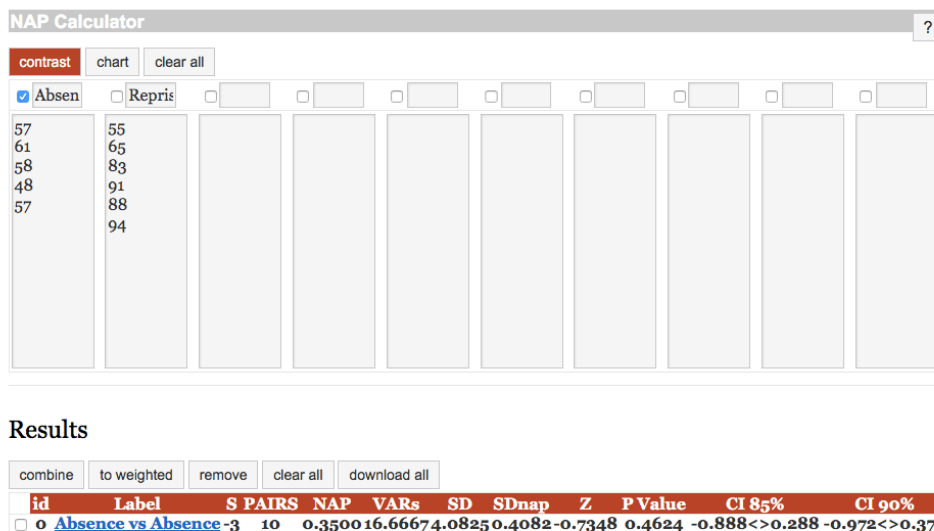


Figure 5. Comparaison de la phase pré-intervention (nommée ici « Absence d'intervention ») avec elle-même.

Si la phase n'est pas stable (valeur  $p < .05$ ), le NAP ne peut pas être employé pour déterminer l'efficacité de l'intervention. Une solution alternative est alors d'avoir recours à la statistique Tau.

### 3. STATISTIQUE TAU

Tau-U (Parker et al., 2010) est également une méthode pour mesurer le non-chevauchement des données entre deux phases (ici l'absence d'intervention et la reprise de l'intervention). Il s'agit d'une technique non paramétrique bien adaptée aux petits ensembles de données. Elle permet, en outre, de contrôler l'influence des variations détectées lors d'une phase pré ou post-intervention.

#### 3.1. Comment faire ?

Tout comme pour le NAP, un calculateur est disponible en ligne pour la statistique Tau à l'adresse : <http://www.singlecaseresearch.org/calculators/tau-u>

#### LES TROIS ÉTAPES : ENCODAGE – CALCUL - INTERPRÉTATION

L'étape d'encodage est strictement identique à l'étape d'encodage pour le NAP, de même que l'étape d'interprétation. La seule différence se situe au niveau de l'étape de calcul. Après avoir sélectionné les deux colonnes qui doivent être comparées et avant d'appuyer sur le bouton « Contrast », la case « correct baseline » doit être cochée afin de contrôler la variabilité présente pendant d'absence d'intervention (l'équivalent de la phase pré-intervention) dans notre illustration (Figure 6).

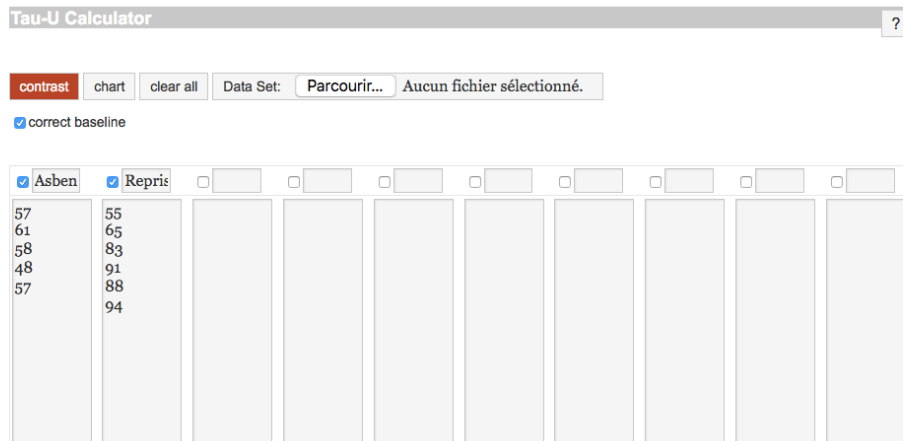


Figure 6. Calcul du Tau en contrôlant la variabilité de phase « absence d'intervention ».

### 3.2. Avantages de la statistique Tau

Le principal avantage de la statistique Tau est sa capacité à contrôler la variabilité des mesures pré-post intervention. Elle possède, en outre, tous les autres avantages du pourcentage de non-chevauchement, si ce n'est qu'il n'est pas possible de la calculer manuellement avec facilité.

### 3.3. Inconvénient de la statistique Tau

Comme le NAP, la statistique Tau n'est pas très facile à adapter au design ABAB ou ABAC : l'ensemble des quatre phases ne sera jamais comparé dans la même analyse ; il est seulement possible de comparer les phases deux à deux (comme dans l'exemple plus haut). De plus, la statistique Tau ne convient pas si une seule donnée est prise pour chaque phase : par exemple une donnée prise au début de l'intervention et une donnée prise à la fin de l'intervention.

## 4. L'ERREUR STANDARD DE MESURE

L'erreur de mesure est l'écart entre la valeur mesurée par un instrument (p. ex. questionnaire, test, etc.) et la valeur vraie (c.-à-d. la performance réelle du patient). Quand l'erreur de mesure peut être calculée pour un outil, sa valeur peut être utilisée pour déterminer si la différence de performance observée entre deux évaluations (ici, entre les phases pré et post-intervention) est liée à une absence de stabilité de la mesure (variations intrinsèques liées à l'outil) ou si elle résulte d'une vraie évolution du comportement du patient.

### 4.1. Comment faire ?

Pour calculer l'erreur de mesure et ensuite établir un intervalle de confiance afin de déterminer si la performance du patient s'est améliorée à la suite de la prise en charge, plusieurs étapes doivent être remplies :

#### ÉTAPE 1 - CALCULER L'ERREUR DE MESURE

L'erreur de mesure se calcule à l'aide de la formule suivante :

$$ETM = \text{ÉT} \sqrt{1-r}$$

ETM = erreur-type de mesure ; ÉT : écart-type de la distribution des scores ; r : coefficient de fidélité (p. ex. coefficient test-retest).

Pour pouvoir calculer l'erreur de mesure, la fidélité de l'outil que vous utilisez et l'écart-type de la distribution doivent être disponibles. Ces informations peuvent être trouvées dans le manuel d'utilisation ou dans l'article de validation de l'outil. Ainsi, par exemple, pour un outil pour lequel le

patient a obtenu un score de 10 en pré-intervention et un score de 13 en post-intervention, dont la fidélité test-retest serait de .85 et l'écart-type de la distribution de 5, l'erreur de mesure serait :

$$ETM = 5\sqrt{1-0.85} = 2.24$$

## ÉTAPE 2 - CALCULER L'INTERVALLE DE CONFIANCE

L'intervalle de confiance (ou IC) est la probabilité que le score vrai du patient se situe à l'intérieur d'une zone spécifique autour du score observé. L'IC varie en fonction du seuil de probabilité utilisé. En calculant l'intervalle de confiance, il est donc possible de déterminer si le score observé après une intervention se situe à l'intérieur de la zone de variation naturelle de l'outil (zone liée à l'erreur de mesure) ou si ce score se situe en dehors de la zone de variation naturelle et donc, peut être expliqué par l'effet de l'intervention. L'IC se calcule à l'aide de la formule suivante :

$$IC = [(x_i - (ETM \times \text{seuil}); (x_i + (ETM \times \text{seuil}))]$$

$x_i$  = le score obtenu par le patient; ETM = erreur de mesure ; seuil = probabilité que le score vrai du patient se situe dans l'intervalle.

Plusieurs seuils peuvent être utilisés en fonction du taux d'erreurs qui est toléré. Si vous souhaitez être certain que l'intervalle de confiance inclut réellement le score vrai du patient, un seuil de 95% ou même de 99% est recommandé. Cela diminuera les chances de mettre en évidence un effet de l'intervention, mais diminuera également les risques de conclure à la présence d'un effet qui n'existe pas. Si, à l'inverse, vous souhaitez augmenter les chances de mettre en évidence un effet de l'intervention, tout en sachant que cela augmentera également les risques de conclure à un effet qui est, en réalité, lié à la variation naturelle de l'outil, un score seuil de 90% peut être accepté.

Ainsi, si l'outil dont nous avons calculé l'erreur de mesure ci-dessus avait été utilisé pour mesurer l'efficacité d'une prise en charge où le patient avait obtenu un score de 10 en pré-intervention et un score de 13 en post-intervention, l'intervalle de confiance à 95% serait de :

$$IC = [(10 - (2.24 \times 0.95); (10 + (2.24 \times 0.95))] = [7.9; 12.1]$$

Si le seuil utilisé était de 99%, l'intervalle de confiance serait de :

$$IC = [(10 - (2.24 \times 0.99); (10 + (2.24 \times 0.99))] = [7.8; 12.2]$$

Si le seuil utilisé était de 90%, l'intervalle de confiance serait de :

$$IC = [(10 - (2.24 \times 0.90); (10 + (2.24 \times 0.90))] = [8; 12]$$

## ÉTAPE 3 - INTERPRÉTER L'INTERVALLE DE CONFIANCE

Si le score obtenu par le patient en post-intervention n'est pas inclus dans l'intervalle de confiance calculé sur base du score en pré-intervention (ici, 10), cela signifie que le score ne se situe pas dans l'intervalle de variation naturelle de l'outil. Dans ce cas, il est possible de conclure – avec une probabilité de 99%, 95% ou 90%, selon le seuil choisi – que le changement observé entre le début et la fin de la prise en charge ne découle pas de l'erreur de mesure. À l'inverse, si le score en post-intervention est inclus dans l'intervalle de confiance, cela signifie qu'un tel score pourrait avoir été obtenu même si aucune intervention n'avait été menée car ce score se situe à l'intérieur de l'intervalle de variation qu'il est possible d'observer pour cet outil.

Dans notre exemple, l'intervalle de confiance à 99% varie entre 7.9 et 12.1, le patient ayant obtenu un score de 13, il se situe en dehors de l'intervalle de confiance et donc nous pouvons conclure, avec une certitude de 99%, que la performance observée n'est pas uniquement due à la variation naturelle attachée à l'outil, mais qu'un autre facteur est intervenu (ici, probablement, la prise en charge).

### 4.1. Avantages de l'erreur de mesure

Le principal avantage de l'erreur de mesure et des intervalles de confiance est qu'ils peuvent être utilisés pour mesurer l'efficacité d'une intervention menée avec un simple design pré-post (c.-à-d., quand une seule mesure est prise au début et une seule mesure est prise en fin d'intervention). La méthode est simple et ne nécessite le recours à aucun logiciel statistique particulier.

### 4.2. Inconvénients de l'erreur de mesure



Afin de pouvoir calculer l'erreur de mesure, l'outil qui est utilisé pour évaluer la performance du patient doit obligatoirement être validé et posséder un indice de fidélité. Sans indice de fidélité, l'erreur de mesure ne peut pas être obtenue. Cette méthode ne peut donc pas être utilisée si une mesure est construite spécifiquement pour la prise en charge du patient. Par ailleurs, en orthophonie, aucune batterie standardisée d'évaluation du langage oral chez l'enfant ne fournit actuellement des informations sur la fidélité de chaque épreuve de cette batterie (Leclercq & Veys, 2014).

Par ailleurs, le fait que l'erreur de mesure puisse être utilisée avec un simple design pré-post ne vient en aucun cas atténuer les inconvénients liés à ce type de design. En effet, si l'erreur de mesure tient compte de la variabilité liée à l'outil, elle ne tient pas compte de la variabilité liée au patient lui-même. Par exemple, elle ne permet pas d'écarter l'hypothèse selon laquelle les changements observés dans la performance du patient puissent être liés à une amélioration naturelle de la performance de celui-ci ou que cette amélioration puisse être due à des facteurs extérieurs à la prise en charge. Pour tenir compte de cette variation propre au patient, il convient de réaliser de multiples mesures à différents moments de la prise en charge et de calculer l'erreur de mesure sur la moyenne de ces différents scores.

## 5. TEST DE McNEMAR

Le test de McNemar permet de comparer deux échantillons de données *appariés*, autrement dit de comparer des données récoltées à deux reprises, sur les *mêmes* items, auprès du patient. Pour les participants familiers aux tests statistiques, le test de McNemar constitue une alternative non-paramétrique au test T pour deux échantillons appariés.

### 5.1. *Dans quels cas ?*

Pour que le test de McNemar puisse être utilisé, les données récoltées doivent toutefois être binaires, c'est-à-dire cotées en 0 ou 1 (0 échoué-1 réussi ; 0 absent-1 présent) (Tableau 3).

Exemples de données éligibles	Exemples de données <u>non</u> éligibles
<ul style="list-style-type: none"> <li>- Liste de mots à orthographier, à prononcer : 0 = erreur, 1 = correct</li> <li>- Liste de calculs/problèmes à résoudre : 0 = erreur, 1 = correct</li> <li>- Présence d'un pointage pour demander un objet : 0 = absent, 1 = présent</li> <li>- Réponses à des questions de compréhension écrite où la réponse est soit juste, soit fausse</li> </ul>	<ul style="list-style-type: none"> <li>- Temps de lecture d'un texte de 100 mots</li> <li>- Nombre de verbes différents utilisés lors d'un récit d'histoire</li> <li>- Pourcentage de dysfluences typiques du bégaiement</li> <li>- Analyse qualitative de la lecture : « la lecture est hachée » : 1 = un peu, 2 = souvent, 3 = tout le temps</li> </ul>

**Tableau 3.** Liste non exhaustive des situations cliniques qui sont éligibles ou non pour l'utilisation du test de McNemar.

### 5.2. *Comment le test fonctionne-t-il ?*

Afin de comprendre le fonctionnement du test de McNemar (et d'appliquer le test), il est nécessaire de réaliser un tableau de contingence 2x2. Ce tableau répartit les résultats du patient en pré-intervention et en post-intervention en 4 catégories :

- A (les items réussis par le patient avant l'intervention mais échoués après l'intervention)
- B (les items réussis par le patient avant l'intervention et toujours réussis après l'intervention)
- C (les items échoués par le patient avant l'intervention et toujours échoués après l'intervention)
- D (les items échoués par le patient avant l'intervention mais réussis après l'intervention).

Le test de McNemar va regarder la distance entre les *couples discordants*, c'est-à-dire qu'il se focalise sur les items qui ont une valeur différente entre les deux prises de mesures (les nombres A et D).

Remarque : Selon certains auteurs, le test ne serait valide que si le nombre total de couples discordants est suffisamment important (un effectif de 10 est souvent retenu) (Dagnelie P. Inférence statistique à

une et à deux dimensions. Statistique théorique et appliquée, Tome 2. Bruxelles : De Boeck et Larcier, 2006).

La statistique de McNemar est la suivante :

$$K = \frac{(A - D)^2}{A + D}$$

Le résultat de cette équation est ensuite comparé à la valeur seuil dans la table de la loi du Chi<sup>2</sup> avec un degré de liberté de 1. Si le résultat de l'équation est strictement supérieur à la valeur seuil, alors on rejette l'hypothèse nulle qui supposait que les différences observées entre les valeurs du patient pré/post-intervention étaient juste dues au hasard. Autrement dit, cela signifie dans ce cas que le comportement du patient est significativement différent entre les 2 moments d'évaluation.

### 5.3. Concrètement, comment faire ?

Via le même url, nous mettons à votre disposition un fichier excell qui applique de manière automatique la formule ci-dessus sur les résultats que vous encodez. Concrètement, dans ce fichier excell, vous devez encoder 4 chiffres (A, B, C, D). Lorsque c'est fait, le fichier fournit la valeur de Chi<sup>2</sup> et indique si la différence est significative ou non. S'il est indiqué « significatif », cela veut dire que l'hypothèse nulle est rejetée (le comportement du patient est significativement différent entre les deux moments d'évaluation). S'il est indiqué « non significatif », cela veut dire que l'hypothèse nulle est acceptée (le comportement du patient n'est pas significativement différent entre les deux moments d'évaluation ; les éventuelles fluctuations observées sont probablement le fruit du hasard).

À titre d'illustration, voici les résultats d'un patient à une mesure de pré-intervention et la même mesure en post-intervention (Tableau 4) : votre mesure était composée de 12 mots à produire ; pour chaque item, vous avez coté la performance de votre patient : 0 si le phonème K ou T est omis ou remplacé, 1 si le phonème K ou T est correctement produit.

	Mesure pré-intervention	Mesure post-intervention
Cadeau	0	1
Toupie	1	1
Clou	0	0
Train	0	1
Ecureuil	0	0
Fantôme	0	1
Coeur	1	0
Tigre	1	1
Cinq	0	1
Patte	0	1
Camion	0	1
Tambour	0	1

*Tableau 4. Illustration des données obtenues par un patient à une mesure en pré et post-intervention.*

#### ÉTAPE 1 - ENCODER DANS LE FICHIER EXCELL LES VALEURS A, B, C, D

Il convient tout d'abord de comptabiliser les items qui correspondent à chaque cas de figure: le nombre d'items réussis avant la prise en charge mais échoués après la prise en charge (en vert dans la Figure 7), le nombre d'items réussis avant et après la prise en charge (en gris dans la Figure 7), le

nombre d'items échoués avant et après la prise en charge (en bleu dans la Figure 7) et le nombre d'items échoués avant mais réussis après la prise en charge (en jaune dans la Figure 7).

	Mesure 1 (avant)	Mesure 1 (après)
Item 1	0	1
Item 2	1	1
Item 3	0	0
Item 4	0	1
Item 5	0	0
Item 6	0	1
Item 7	1	0
Item 8	1	1
Item 9	0	1
Item 10	0	1
Item 11	0	1
Item 12	0	1

**Figure 7.** Illustration : identification des 4 configurations.

Ensuite, il est nécessaire d'encoder les 4 valeurs dans les cases associées du fichier excell (Tableau 5).

		Après prise en charge	
		Echoué	Réussi
Avant prise en charge	Réussi	A	B
	Echoué	C	D

		Après prise en charge	
		Echoué	Réussi
Avant prise en charge	Réussi	1	2
	Echoué	2	7

**Tableau 5.** Illustration : fichier excell complété avec les 4 valeurs obtenues.

## ÉTAPE 2 - INTERPRÉTER LE RÉSULTAT

Le fichier excell est configuré pour fournir automatiquement le résultat du test statistique lorsque les 4 valeurs ont été encodées. Deux informations sont disponibles : la valeur du Chi<sup>2</sup> (3.125 dans l'illustration – Tableau 6) et l'interprétation à retenir. L'indication « significatif » précise qu'il y a une différence statistiquement significative entre les performances du patient avant et après l'intervention. Au contraire, l'indication « non significatif » précise qu'il n'y a pas de différence statistiquement significative entre les performances obtenues aux deux moments d'évaluation.

$$X^2 = \begin{array}{|c|c|} \hline 3,125 & \text{Non significatif} \\ \hline \end{array}$$

**Tableau 6.** Illustration : résultat au test de McNemar

## 6. BIBLIOGRAPHIE

Byiers, B. J., Reichle, J., & Symons, F. J. (2012). Single-subject experimental design for evidence-based practice. *American Journal of Speech-Language Pathology*, 21(4), 397-414.

Gage, N. A., & Lewis, T. J. (2013). Analysis of effect for single-case design research. *Journal of Applied Sport Psychology*, 25(1), 46-60.

Leclercq, A. L., & Veys, E. (2014). Réflexions sur le choix de tests standardisés lors du diagnostic de dysphasie. *ANAE: Approche Neuropsychologique des Apprentissages chez l'Enfant*.

Parker, R.I., & Vannest, K.J. (2009). An improved effect size for single case research: Non-overlap of all pairs (NAP). *Behavior Therapy*, 40(4), 357-367.

Parker, R.I., Vannest, K.J., Davis, J.L., Sauber, S.B. (2010). Combining non-overlap and trend for single case research: Tau-U. *Behavior Therapy*, 42, 284–299.

Vannest, K. J., & Ninci, J. (2015). Evaluating treatment effects in single-case research designs. *Journal of Counseling & Development*, 93, 403-411.