

# Interpréter les résultats

*Objectif : s'assurer de l'effet de l'intervention sur  
les progrès observés*

# Ce qu'il faut retenir

L'analyse des résultats repose sur :

- 👉 leur significativité clinique, qui pourra être mise en évidence par des questionnaires qualitatifs auprès du patient et/ou de son entourage (ces aspects sont développés dans la Synthèse d'articles #5 Mesures fonctionnelles),
- 👉 leur significativité numérique, à l'appui des mesures prises lors des lignes de base.

Pour évaluer l'effet de votre intervention grâce aux mesures prises lors des lignes de base :

- 👉 la lecture visuelle des graphiques donne une première idée des progrès du patient

Pour s'assurer que les résultats sont bien la conséquence de l'intervention thérapeutique, il existe des tests statistiques :

- 👉 Test de TAU-U pour comparer deux périodes avec / sans intervention
- 👉 Test de Mac Nemar pour comparer des résultats échoués / réussis entre le pré-test et le post-test
- 👉 Test des rangs de Wilcoxon pour comparer des mesures de temps (durée, vitesse) entre le pré-test et le post-test

# Analyser les résultats

Martinez Perez, T., Geurten, M., & Willems, S. (2021). Collecter des données sur l'évolution du patient pour augmenter son efficacité thérapeutique en orthophonie. ANAE, 33(173), 433-431.

Pour analyser les résultats des mesures prises avec les lignes de base, T. Martinez Perez propose deux axes d'interprétation :

- Quelle **significativité clinique** reflète les mesures ? Pour évaluer le transfert réel des capacités travaillées dans les activités quotidiennes et le ressenti du patient, l'auteur propose d'utiliser l'échelle GAS, qui permet de questionner l'atteinte des objectifs fixés. Cet outil est présenté dans la **Synthèse d'articles #5** (les mesures fonctionnelles).
- Quelle **significativité numérique** entre les mesures pré/post, les mesures contrôle ? En effet, une simple analyse visuelle sur un graphique permet d'observer si les performances s'améliorent, stagnent ou se dégradent. Mais comment savoir si ces tendances sont significatives et reflètent réellement l'effet de l'intervention thérapeutique ?

# Interpréter les résultats visuellement

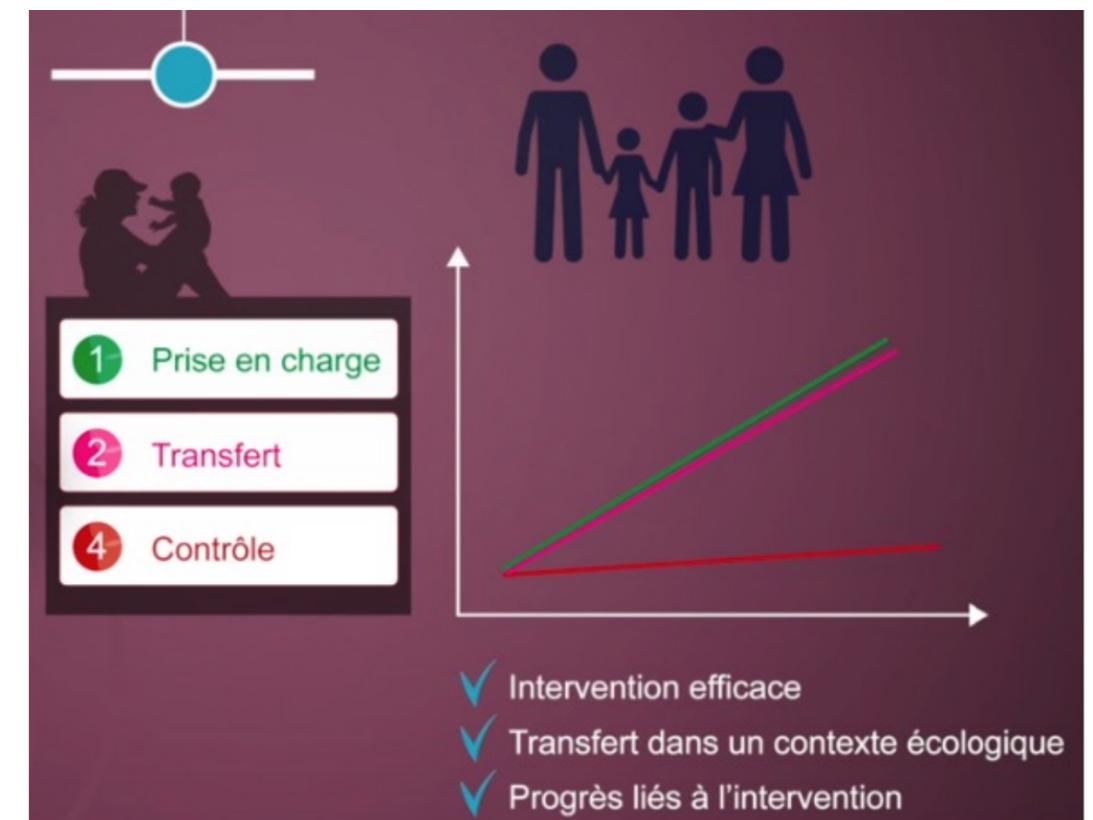
Source : Fun Mooc EBP 2020, Module 6, séquence 4

Psychologue et orthophoniste : L'EBP au service du patient. (s. d.). FUN MOOC, à l'adresse <https://www.fun-mooc.fr/fr/cours/psychologue-et-orthophoniste-lebp-au-service-du-patient/#/> (disponible jusqu'au 01/07/2021)

## Interprétation : 4 configurations de lignes de base

Exemple proposé par T. Martinez-Perez : évaluation d'un programme d'intervention centrée sur le parent

- #1 Dans une **première famille**, le clinicien pourrait uniquement constater une amélioration des résultats dans les mesures "prise en charge" et "transfert". Il s'agit de la configuration idéale qui peut être interprétée comme suit : l'intervention du clinicien a été efficace car elle a permis un changement chez le parent, transférable dans un autre contexte écologique. De plus, les progrès peuvent être interprétés comme étant la conséquence de l'intervention du clinicien car le parent n'a pas progressé pour l'instant sur la mesure contrôle.

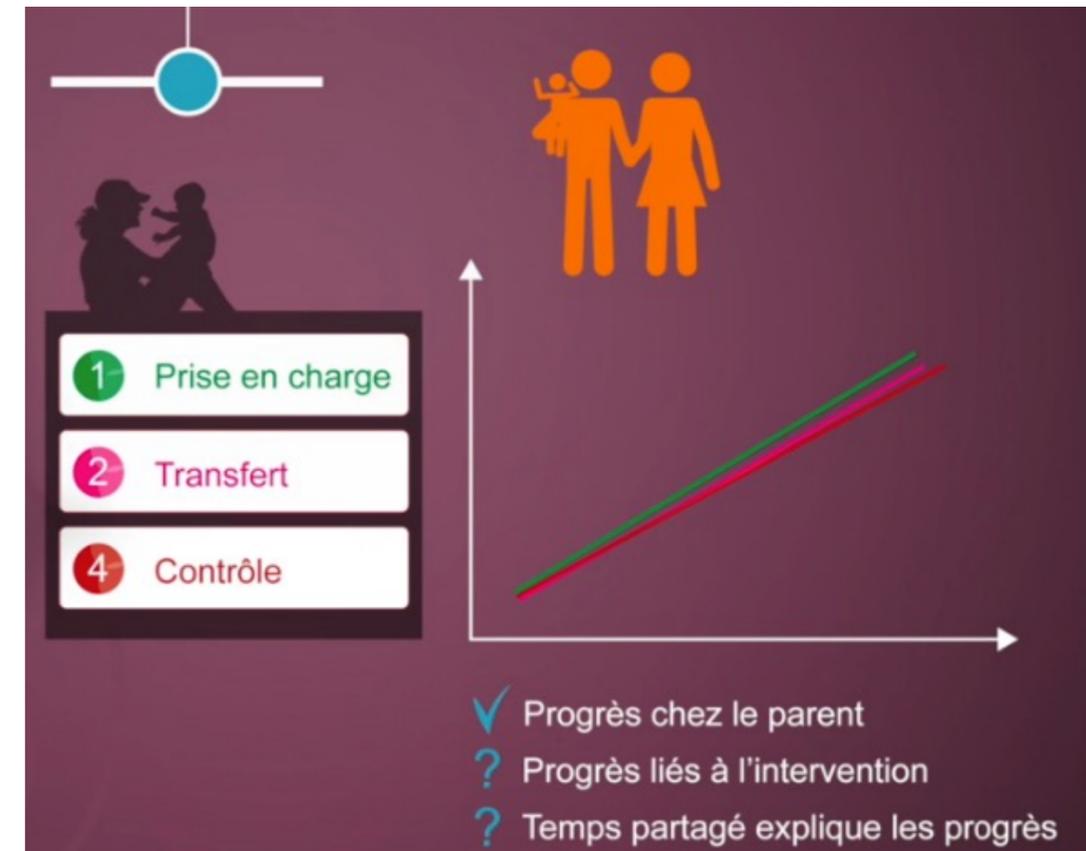


# Interpréter les résultats visuellement

**Interprétation : 4 configurations de lignes de base**

**Exemple proposé par T. Martinez-Perez : évaluation d'un programme d'intervention centrée sur le parent**

**#2** Dans une **deuxième famille**, une amélioration des résultats pourrait s'observer dans les trois mesures. Il est ici impossible d'affirmer que les changements chez le parent sont la conséquence de l'intervention du clinicien. Il est notamment possible que les progrès soient simplement la conséquence d'une augmentation d'un temps partagé entre enfant et parent.

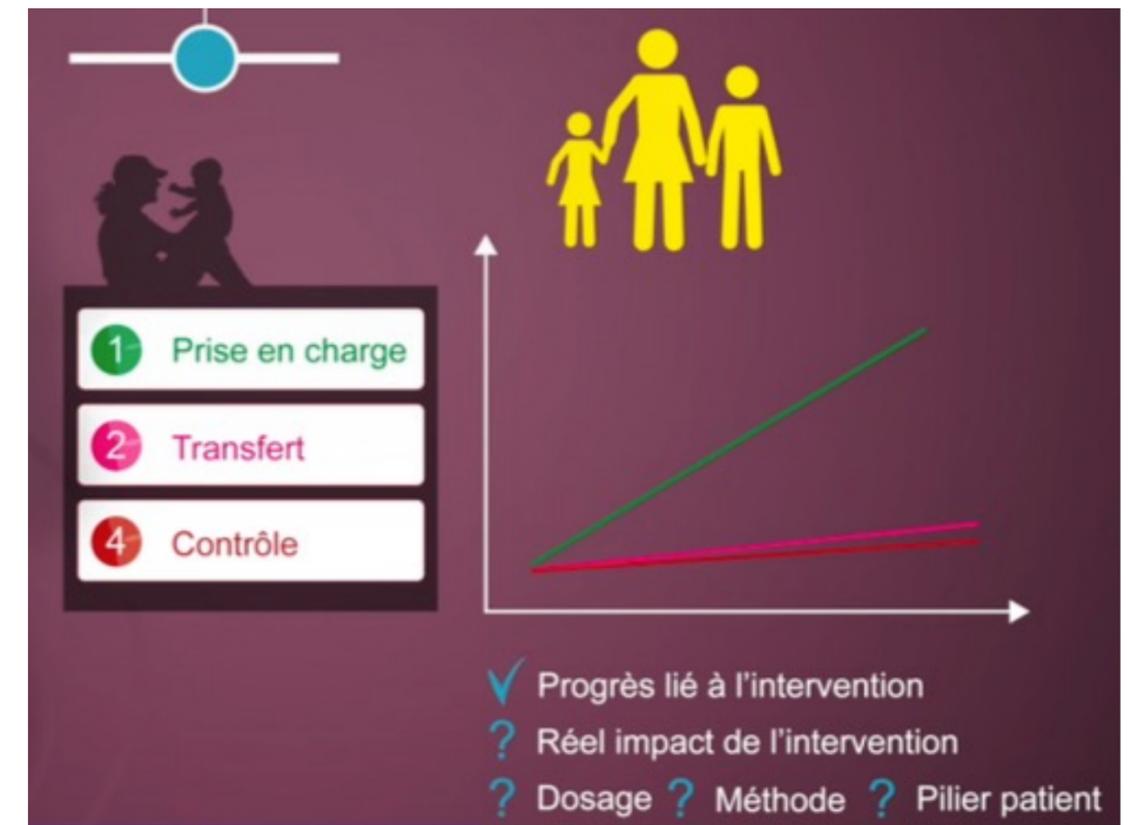


# Interpréter les résultats visuellement

## Interprétation : 4 configurations de lignes de base

Exemple proposé par T. Martinez-Perez : évaluation d'un programme d'intervention centrée sur le parent

- #3 Dans une **troisième famille**, seule la mesure "prise en charge" pourrait s'être améliorée. Le changement du parent peut être attribué à l'intervention du clinicien car la mesure "contrôle" est restée stable. Mais l'absence de résultat dans la mesure "transfert" suggère que l'intervention parentale risque d'être peu généralisable et donc d'entraîner peu de diminution des troubles chez l'enfant. Plusieurs hypothèses peuvent alors être énoncées : est-ce le dosage de l'intervention, notamment le nombre de séances avec le parent, qui était insuffisant pour permettre un transfert ? Est-ce la méthode d'intervention du clinicien qui était inadéquate ? Est-ce que certains éléments du contexte familial, comme la présence d'un nourrisson, n'ont pas été pris en compte et complexifient la gestion des comportements de l'enfant par le parent ?

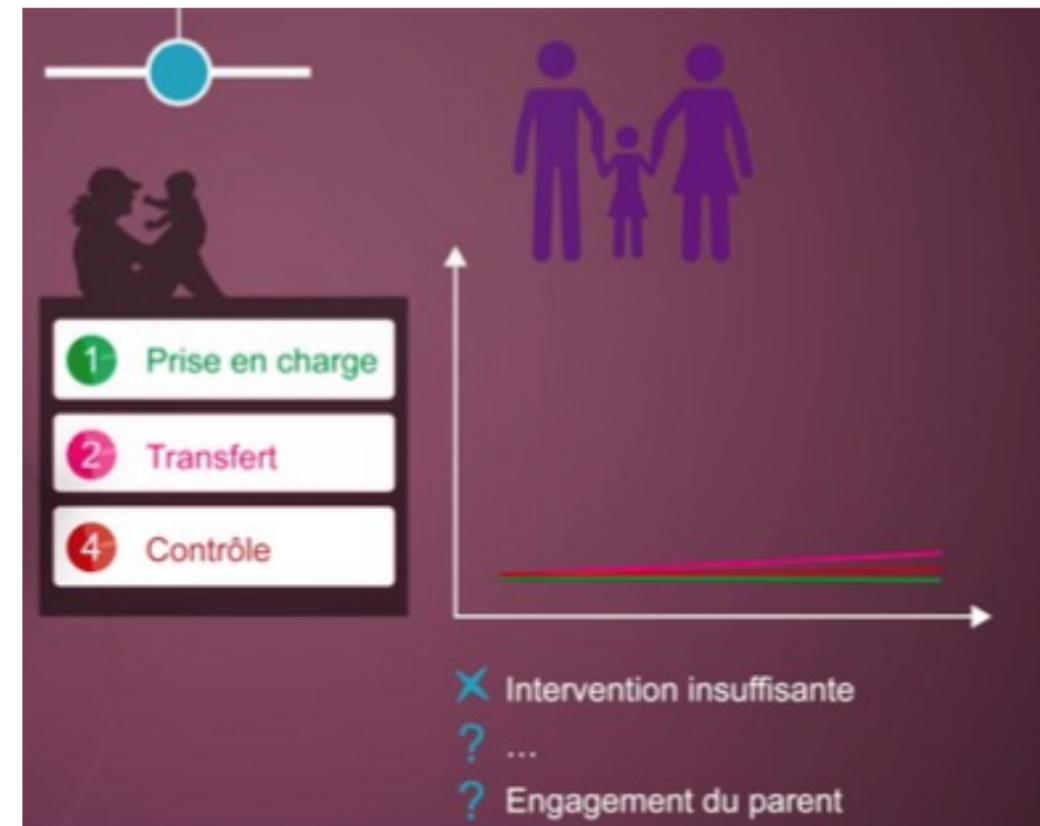


# Interpréter les résultats visuellement

## Interprétation : 4 configurations de lignes de base

Exemple proposé par T. Martinez-Perez : évaluation d'un programme d'intervention centrée sur le parent

- #4 Finalement, dans une **quatrième famille**, aucune mesure n'aurait bougé.  
Cela suggère simplement que le fait d'enclencher une intervention ne permet pas au parent de modifier ses interactions avec son enfant. Face à l'absence de résultat dans la mesure "prise en charge", des nouvelles hypothèses peuvent s'ajouter à celles précédemment citées. Notamment, est-ce un problème d'engagement du parent dans l'intervention ?



# Choisir le test statistique

## Test de TAU-U

Comparer deux périodes avec et sans intervention

Pour s'assurer que l'intervention est bien responsable des progrès observés, il est possible de **comparer les performances du patient en post-test 1** (immédiatement après l'intervention), **avec des données en post-test 2** (à l'issue d'un autre type d'intervention).

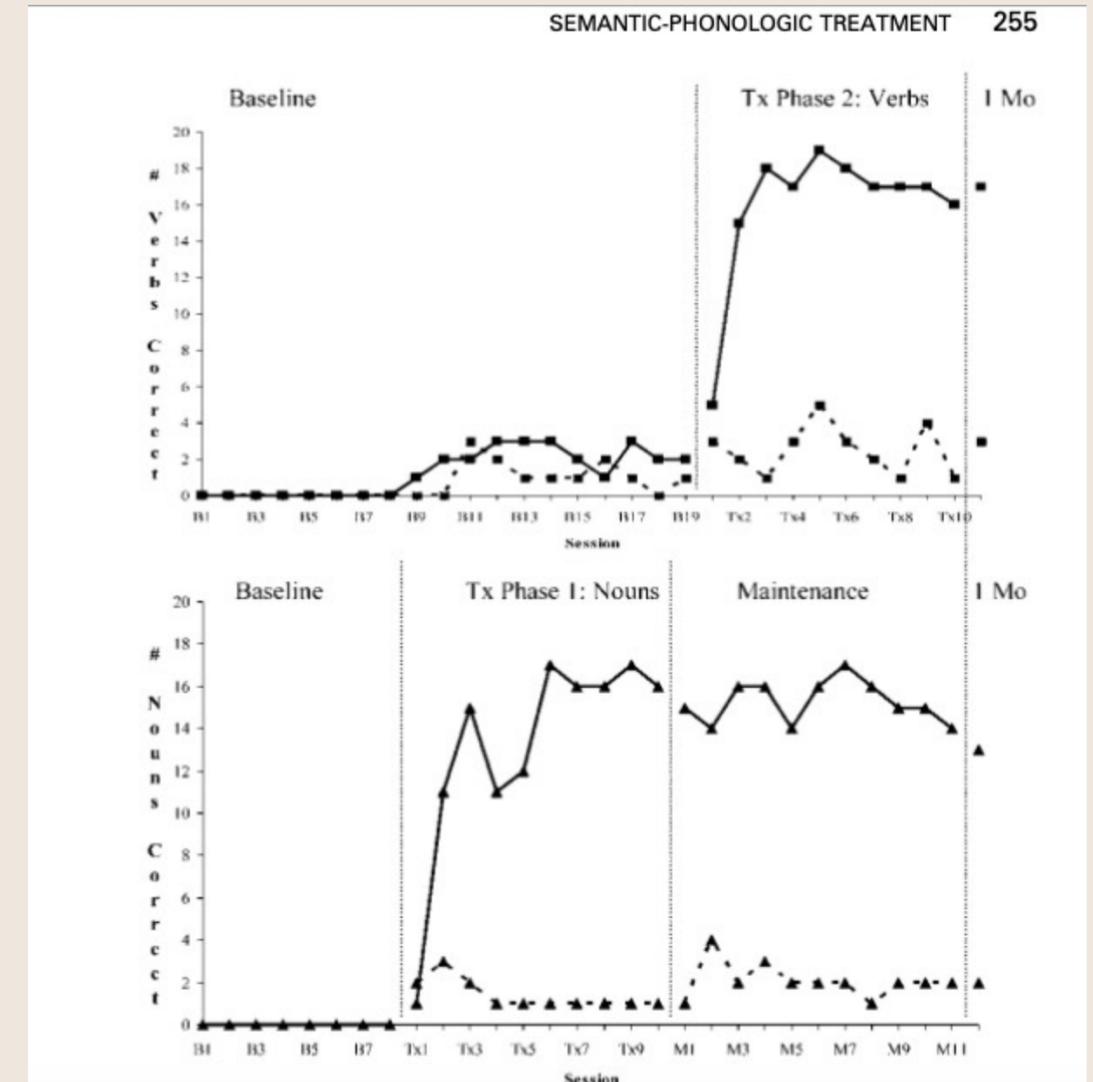
Le test statistique qui s'applique dans ce cas est le TAU-U, qui permet de comparer deux à deux des données numériques entre deux périodes. <http://www.singlecaseresearch.org/#/>

Dans cet exemple, des mesures contrôles sur les verbes ont été prises jusqu'au point n°19 (*baseline* du graphique du haut).

A + 1 semaine, on commence l'entraînement sur les noms (graphique du bas).

Au 19<sup>e</sup> jour, on débute l'entraînement sur les verbes et on passe en maintenance sans entraînement des noms (listes variées piochées dans des centaines d'images).

On note que pendant l'entraînement des noms, la *baseline* des verbes n'est plus à zéro.



Exemple tiré de l'article Raymer, M. A. Ciampitti<sup>3</sup>, M., Holliday, B. et al. Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia, *Neuropsychological Rehabilitation* 2007, 17 (2), 244–270



# Choisir le test statistique

## Comparer deux périodes avec et sans intervention (test de TAU-U)

Formuler des **hypothèses**. Par exemple :

Hypothèse 1 : Quel effet de l'entraînement des noms (vs récupération spontanée) ?

Hypothèse 2 : Quel effet de l'entraînement des verbes (vs récupération spontanée) ?

Hypothèse 3 : Quel transfert de la liste entraînée des noms sur la liste non entraînée des verbes ?

Méthodo complète :

*disponible*

*en ressources*

Tau-U Calculator

contrast chart clear all Data Set: Choisir un fichier Aucun fichier

correct baseline

baseli  entrVl

0	5				
0	15				
0	17				
0	16				
0	18				
0	17				
0	16				
0	16				
0	15				

Results

combine to weighted remove clear all download all

id	Label	S	PAIRS	TAU	TAU <sub>b</sub>	VARs	SD	SDtau	Z	P Value	CI 85%	CI 90%
trend:												
<input type="checkbox"/>	-	-	-	-	-	-	-	-	-	-	-	-
phase:												
<input type="checkbox"/>	o baseline vs entrVb	80	80	1	1	506.6667	22.5093	0.2814	3.5541	0.0004	0.595<>1	0.537<>1
corrected baseline:												
<input type="checkbox"/>	-	-	-	-	-	-	-	-	-	-	-	-
combined:												
<input type="checkbox"/>	-	-	-	-	-	-	-	-	-	-	-	-

### Résultats :

La p-value < 0,05, donc l'effet de l'entraînement des verbes est significatif.

Hypothèse 2 : Importer les données de la baseline + celles des verbes entraînés

Cocher les colonnes et cliquer sur 'contrast'

# Choisir le test statistique

## Test de Mac Nemar

### Comparer des données catégorielles (juste/faux)

Lorsque les données de pré-test et de post-test sont mesurées selon **un seul critère : échec / réussite**, elles peuvent être codées sous la forme 0 ou 1.

Dans un tableur de type Excel, chaque liste donne lieu à un pourcentage de réponses correctes.

Le tableur permet aussi de créer un graphique montrant l'évolution des réponses correctes entre le pré-test et le post-test.

Pour valider leur significativité, les résultats sont répartis dans un **tableau de contingence** :

Test de McNemar - Calcul du $\chi^2$			
		Après prise en charge	
		Echoué	Réussi
Avant prise en charge	Réussi	5	7
	Echoué	27	15
		$\chi^2 =$	4,05
			<b>Significatif</b>

*Remplacer A, B, C et D dans le tableau par leur valeur correspondante*

A =	Nombres d'items réussis avant prise en charge et échoués après la prise en charge
B =	Nombres d'items réussis avant prise en charge et réussis après la prise en charge
C =	Nombres d'items échoués avant prise en charge et échoués après la prise en charge
D =	Nombres d'items échoués avant prise en charge et réussis après la prise en charge

Méthodo complète :  
*disponible*  
*en ressources*

# Choisir le test statistique

## Test des rangs de Wilcoxon

Comparer des variables continues (mesures de temps : vitesse de lecture, temps de réponse...)

Lorsque les données de pré-test et de post-test sont **mesurées en temps**, elles ne peuvent pas être comparées de la même manière que des variables catégorielles.

Dans un tableur de type Excel, chaque liste donne lieu à un temps moyen de réponse.

Test des rangs de **Wilcoxon** :

<http://www.socscistatistics.com/tests/signedranks/Default2.aspx>

Si la liste = ou > 20 items, choisir la Z-value

Si la liste < ou 20 items, choisir la W-value

Le test est significatif si  $p < 0,05$

Significance Level: <input type="radio"/> 0.01 <input checked="" type="radio"/> 0.05	<u>Result Details</u>  <i>W-value: 53</i> <i>Mean Difference: -161.1</i> <i>Sum of pos. ranks: 157</i> <i>Sum of neg. ranks: 53</i>  <i>Z-value: -1.9413</i> <i>Mean (W): 105</i> <i>Standard Deviation (W): 26.79</i>  <i>Sample Size (N): 20</i>
1 or 2-tailed hypothesis?:  <input checked="" type="radio"/> One-tailed <input type="radio"/> Two-tailed	

*Result 1 - Z-value* nombre d'items = 20 ou plus

The Z-value is -1.9413. The p-value is 0.02619. The result is significant at  $p \leq 0.05$ .

*Result 2 - W-value* nombre d'items <20

The W-value is 53. The critical value of W for  $N = 20$  at  $p \leq 0.05$  is 60. Therefore, the result is significant at  $p \leq 0.05$ .

# Pour aller plus loin

## ✓ **Fiches méthodo déposées sur le site**

- ✓ Martinez Perez, T., Geurten, M., & Willems, S. (2021). Collecter des données sur l'évolution du patient pour augmenter son efficacité thérapeutique en orthophonie. ANAE, 33(173) : matériel supplémentaire : tutoriel pour les analyses
- ✓ Buerki et Cohen, 2018. EBP, Représentation des données et validation statistique. Tutoriel du test de Mac Nemar et du test des rangs de Wilcoxon dans Excel
- ✓ Fichier Excel vierge du test de Mac Nemar

## ✓ **Liens internet**

- ✓ Test de TAU-U <http://www.singlecaseresearch.org/#/>
- ✓ Test des rangs de Wilcoxon <http://www.socscistatistics.com/tests/signedranks/Default2.aspx>