

# mtss-k

Entre standards de qualité de la recherche et contraintes pragmatiques, quelle cohabitation dans le cadre de trois essais contrôlés randomisés sur les effets des interventions Multi-Tiered System of Supports (MTSS) ?

Dachet D., Dierendonck, C., Aubry, A. & Baye A.



# Plan de la présentation

1. **Quelle définition de la RCT ?**
2. **Quels standards de qualité pour les designs expérimentaux en SE ?**
3. **Quels biais sont pris en considération par ces standards ?**
4. **Comment une RCT essaye-t-elle d'éviter ces biais ?**
  1. absence de dissimulation de la séquence d'allocation
  2. absence d'aveuglement
  3. séquence d'allocation non aléatoire
  4. facteur confondant (facteur prédictif du résultat prédit également l'assignement à l'intervention)
  5. sélection des participants
  6. méthodes utilisées pour mesurer les variables et les résultats
  7. sélection des résultats rapportés
5. **Bibliographie**

# 1° Quelle définition de la RCT ?

Méthode de recherche permettant d'inférer des liens de causalité entre une variable indépendante (souvent un traitement) et une (ou plusieurs) variable(s) dépendante(s) grâce à la comparaison entre deux groupes de sujets : l'un soumis à la variable indépendante (groupe expérimental) et l'autre pas (groupe contrôle) servant de contrefactuel

(Tanner, 2002 ; Tanner, 2018)



Méthode de constitution des groupes



Organisation de l'administration des instruments de mesure



Mise en œuvre du *blinding*

## 2° Quels standards de qualité pour les designs expérimentaux en SE ?

### *Best Evidence Encyclopedia*

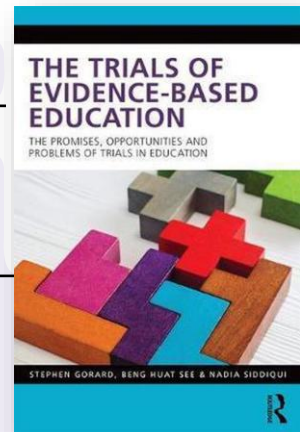
#	Critères d'exclusion
1	Présence d'un groupe expérimental et d'un vrai groupe contrôle
2	Pas de sélection des élèves à posteriori
3	Groupes contrôles et expérimentaux équivalents au départ
4	Dispositif testé par de « vrais » enseignants
5	Tests (standardisés) équitables pour les deux groupes
6	Tests non développés par les promoteurs du dispositif
7	Dispositifs testés sur le moyen terme (1 semestre)
8	2 enseignants et 30 élèves par groupe



## 2° Quels standards de qualité pour les designs expérimentaux en SE ?

### **Gorard Sieve** (Gorard et al., 2017)

<i>Design</i>	Échelle	Attrition	Qualité des données	Autres menaces	Évaluation
<i>Design</i> rigoureux pour la question de recherche	Large nombre de cas (par groupe de comparaison)	Attrition minimale et absence de preuve d'un impact sur les résultats	Préspécifiées, standardisées et indépendantes	Aucun conflit d'intérêt et déclaration de l'absence de conflit d'intérêts	4*
Bon <i>design</i> pour la question de recherche	Nombre moyen de cas (par groupe de comparaison)	Faible attrition (ou déséquilibre initial)	Préspécifiées, non standardisées ou non indépendantes	Aucun conflit d'intérêts sans déclaration de l'absence de conflit d'intérêts	3*
Faible <i>design</i> pour la question de recherche	Faible nombre de cas (par groupe de comparaison)	Attrition modérée (ou déséquilibre initial)	Non préspécifiées, mais valides en contexte	Présence d'un conflit d'intérêts, déclaré par les auteurs et mise en œuvre d'une procédure d'aveuglement	2*
Très faible <i>design</i> pour la question de recherche	Très faible nombre de cas (par groupe de comparaison)	Attrition élevée (ou déséquilibre initial)	Problème de validité ou de pertinence	Présence d'un conflit d'intérêts déclaré par les auteurs	1*
Aucune considération pour le <i>design</i>	Une échelle d'étude triviale, ou N peu clair	Attrition très élevée ou non rapportée	Faible fiabilité, trop de variables d'intérêt, faible mesure	Présence d'un conflit d'intérêts sans qu'il soit déclaré par les auteurs	0*



## 2° Quels standards de qualité pour les designs expérimentaux en SE ?

### *What Works Clearinghouse - Procedures & standards handbook V5.0.*

#	Exclusion	
	Type	Critère
1	Design	Le design n'est pas éligible, c'est-à-dire qu'il ne suit pas un design ECR, PQE, RDD/ DSC
2		Le design n'est pas une étude primaire des impacts ou de l'efficacité d'une intervention
3	Échantillon	L'échantillon n'est pas conforme au protocole spécifique : âge, niveau, SES, genre, handicap, ...
4	Variable(s)	L'étude n'inclut pas au moins un résultat qui relève des domaines identifiés dans le protocole
5	Délai de pub.	L'étude est publiée il y a plus de 20 ans (sauf indication contraire au sein du protocole)
6	Disponibilité	L'étude est disponible publiquement et est décrite complètement
7	Localisation	L'étude est implémentée en-dehors des limites géographiques choisies
8	Intervention	L'intervention étudiée n'est pas conforme au protocole spécifique décrivant les interventions qui sont admissibles et toute autre caractéristique d'admissibilité liée aux interventions, comme des informations sur la méthode de présentation, la reproductibilité de l'intervention et les caractéristiques des personnes qui mettent en œuvre l'intervention.

## 2° Quels standards de qualité pour les designs expérimentaux en SE ?

### ***What Works Clearinghouse - Procedures & standards handbook V5.0.***

Se centre avant tout sur la validité des résultats de la recherche en éducation visant à déterminer des inférences causales.

<b>Rencontre les standards de WWC sans réserve</b>	Résultats issus d'une recherche dont le design est rigoureux et bien exécuté. <b>→ Plus haut degré de confiance que l'intervention cause les effets observés</b>
<b>Rencontre les standards de WWC avec réserve</b>	À cause de carences naturelles dans le design de recherche ou à cause de circonstances autour de la mise en œuvre de la recherche, la recherche n'exclut pas suffisamment la possibilité que quelque chose d'autre que l'intervention ait causé l'effet observé.
<b>Ne rencontre pas les standards de WWC</b>	Les résultats ne sont pas accompagnés de preuves suffisantes que l'intervention a causé les effets observés.

## 2° Quels standards de qualité pour les designs expérimentaux en SE ?

### ***What Works Clearinghouse - Procedures & standards handbook V5.0.***

Étape 1 : évaluer les instruments de mesure et vérification des facteurs de confusion

#### **Instrument de mesure**

- Standard 1 = validité de façade de la mesure
- Standard 2 = fiabilité de la mesure (consistance interne d'au moins 0.60 ; stabilité temporelle et fiabilité test-retest d'au moins 0.40 ; fiabilité inter-codeurs d'au moins 0.50 ; accord inter-codeurs d'au moins 80 %).
- Standard 3 = instrument de mesure n'est pas sur-aligné.
- Standard 4 = procédures de récolte de données sont cohérentes
- Standard additionnel = indépendance de l'instrument de mesure

**Facteurs de confusion** (aspect d'une étude qui est toujours présent pour les membres d'un groupe et jamais pour les membres de l'autre groupe)

- N-1 : unité unique
- Caractéristiques qui peuvent affecter les résultats différent entre les conditions sans chevauchement
- Timing en tant que facteur confondant

# 2° Quels standards de qualité pour les designs expérimentaux en SE ?

## ***What Works Clearinghouse - Procedures & standards handbook V5.0.***

### Étape 2 : assignement aux conditions

- Procédure aléatoire = Rencontre les standards de WWC sans réserve
- Procédure non aléatoire = Rencontre les standards de WWC avec réserves

### Étape 3 : changement de composition

- Attrition (totale et différentielle) dans l'échantillon
- Attrition (totale et différentielle) au niveau des clusters
- Nouveaux arrivants

### Étape 4 : équivalence au prétest

- $0.00 \leq ES \leq 0.05$
- $0.05 \leq ES \leq 0.25$
- $ES > 0.25$

Teams conducting WWC reviews are responsible for determining whether to use optimistic or cautious assumptions when assessing sample attrition, and should document their reasoning. In general, choice of the optimistic boundary indicates the review team's assessment that attrition (or lack thereof) in the sample is unlikely to be related to the intervention. WWC (2022, P. 41)

**Table C.1.** Highest differential attrition rate for a sample to maintain low attrition, by overall attrition rate, under cautious and optimistic assumptions

Overall attrition	Differential attrition		Overall attrition	Differential attrition		Overall attrition	Differential attrition	
	Cautious boundary	Optimistic boundary		Cautious boundary	Optimistic boundary		Cautious boundary	Optimistic boundary
0	5.7	10.0	12	6.2	10.9	24	4.9	9.4
1	5.8	10.1	13	6.1	10.8	25	4.8	9.2
2	5.9	10.2	14	6.0	10.8	26	4.7	9.0
3	5.9	10.3	15	5.9	10.7	27	4.5	8.8
4	6.0	10.4	16	5.9	10.6	28	4.4	8.6
5	6.1	10.5	17	5.8	10.5	29	4.3	8.4
6	6.2	10.7	18	5.7	10.3	30	4.1	8.2
7	6.3	10.8	19	5.5	10.2	31	4.0	8.0
8	6.3	10.9	20	5.4	10.0	32	3.8	7.8
9	6.3	10.9	21	5.3	9.9	33	3.6	7.6
10	6.3	10.9	22	5.2	9.7	34	3.5	7.4
11	6.2	10.9	23	5.1	9.5	35	3.3	7.2

Continued on next page

# 2° Quels standards de qualité pour les designs expérimentaux en SE ?

## ***What Works Clearinghouse - Procedures & standards handbook V5.0.***

### Étape 2 : assignement aux conditions

- Procédure aléatoire = Rencontre les standards de WWC sans réserve
- Procédure non aléatoire = Rencontre les standards de WWC avec réserves

### Étape 3 : changement de composition

- Attrition (totale et différentielle) dans l'échantillon
- Attrition (totale et différentielle) au niveau des clusters
- Nouveaux arrivants

### Étape 4 : équivalence au prétest

- $0.00 \leq ES \leq 0.05$
- $0.05 \leq ES \leq 0.25$
- $ES > 0.25$

Teams conducting WWC reviews are responsible for determining whether to use optimistic or cautious assumptions when assessing sample attrition, and should document their reasoning. In general, choice of the optimistic boundary indicates the review team's assessment that attrition (or lack thereof) in the sample is unlikely to be related to the intervention. WWC (2022, P. 41)

**Table C.1.** Highest differential attrition rate for a sample to maintain low attrition, by overall attrition rate, under cautious and optimistic assumptions (continued)

Overall attrition	Differential attrition		Overall attrition	Differential attrition		Overall attrition	Differential attrition	
	Cautious boundary	Optimistic boundary		Cautious boundary	Optimistic boundary		Cautious boundary	Optimistic boundary
36	3.2	7.0	46	1.6	4.6	56	0.2	2.3
37	3.1	6.7	47	1.5	4.4	57	0.0	2.1
38	2.9	6.5	48	1.3	4.2	58	N/A	1.9
39	2.8	6.3	49	1.2	3.9	59	N/A	1.6
40	2.6	6.0	50	1.0	3.7	60	N/A	1.4
41	2.5	5.8	51	0.9	3.5	61	N/A	1.1
42	2.3	5.6	52	0.7	3.2	62	N/A	0.9
43	2.1	5.3	53	0.6	3.0	63	N/A	0.7
44	2.0	5.1	54	0.4	2.8	64	N/A	0.5
45	1.8	4.9	55	0.3	2.6	65	N/A	0.3

Note: Overall attrition rates are given as percentages. Differential attrition rates are given as percentage points. Not every combination of differential and overall attrition is possible for any given study. N/A is not applicable (if the total attrition rate is 58 percent or higher, no differential attrition rate will yield low attrition under the cautious boundary).

### 3° Quels biais sont pris en considération par ces standards ?

Risque de biais issus ...	WWC	Weight of Evidence	Quality Indicators	BEE	Tamis de Gorard
de l'absence de dissimulation de la séquence d'allocation	X	X	X	X	X
de l'absence d'aveuglement	X	X	X	X	X
d'une séquence d'allocation non aléatoire	V (sans critère sur la génération)	X	(V) Vraie randomisation	(V) Vraie randomisation	(V) Vraie randomisation
d'un facteur confondant (facteur prédictif du résultat prédit également l'assignement à l'intervention)	V N=1 Caractéristiques ≠ Alignement temporel Equivalence initiale	X	V Contamination Équivalence initiale	V N=1 Équivalence initiale	X
de la sélection des participants	(V) Attrition et primo-arrivants	X	(V) Attrition	X	(V) Attrition
des méthodes utilisées pour mesurer les variables et les résultats	V Validité Fiabilité Suralignement Procédures cohérentes Indépendance	X	V Validité Fiabilité	V Standardisation Indépendance Inhérence	V Standardisation Indépendance Pré-spécification
de la sélection des résultats rapportés	X	X	V Effets iatrogènes	X	(V) Pré-spécification
<b>Autres</b>	Procédures d'ajustement statistiques	Transparence Cohérence Adéquation avec QR Pertinence du sujet	Qualité du <i>reporting</i> Fidélité d'imp. Qualité des analyses	Taille des échantillons Utilisabilité de l'intervention	Taille des échantillons Conflit d'intérêt

# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.1. Absence de dissimulation de la séquence d'allocation

### Allocation des écoles aux conditions expérimentales :

- Par un chercheur extérieur au recrutement et non impliqué dans la mise en œuvre de l'intervention
- Au sein de chacun des pays en utilisant
  - Un critère identique (SES)
  - Un/plusieurs critères spécifiques (région, organisation des classes...)



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.2. Absence d'aveuglement

### Trois types d'aveuglement :

- Aveuglement des acteurs/praticiens en charge de l'intervention
  - Très complexe en sciences de l'éducation et surtout dans la recherche interventionnelle
- Aveuglement des chercheurs/administrateurs/... en charge de l'évaluation
  - Utilisation d'administrateur externe de tests pour quelques des variables d'intérêts
  - Mais certaines (comportement en classe, par exemple) nécessitent une évaluation par l'enseignant
  - ... par les parents
- Aveuglement des chercheurs/statisticiens/... en charge des analyses
  - Aveuglement partiel (jusqu'à un certain point) d'un des chercheurs
  - Confrontation systématiques des analyses avec celles réalisés par d'autres issus d'équipes extérieures



## 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

### 4.3. Séquence d'allocation non aléatoire

Dans l'idéal : séquence aléatoire simple sans contrainte (Higgins et al., 2019) ...

Mais financement suisse et européen qui implique d'avoir des écoles expérimentales dans chaque pays  
→ **STRATIFIED RANDOMIZATION**

Mais pragmatiquement plus « simple » de former la moitié des écoles au sein de chaque pays  
→ **BLOCKED RANDOMISATION**

→ Technique de la minimisation (Higgins et al., 2019) avec un assignement par paire



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.3. Séquence d'allocation non aléatoire

```
library(readxl)
library(readr)
CHY <- read.csv2("CHY.csv")
View(CHY)

install.packages("randomizr")
library(randomizr)

set.seed(123)

CHY$composite <- interaction(CHY$SES, CHY$City, drop = TRUE)

treat2 <- randomizr::block_and_cluster_rpa(
  blocks = CHY$composite,
  clusters = 1:nrow(CHY),
  num_arms = 2,
  conditions = c("EXP", "CTRL"),
  block_m = NULL,
  block_m_each = NULL
)
```

SES	treatment	Quantités	% du Total
L	CTRL	16	34.0 %
	EXP	16	34.0 %
O	CTRL	3	6.4 %
	EXP	5	10.6 %
V	CTRL	4	8.5 %
	EXP	3	6.4 %

City	treatment	Quantités	% du Total
Larna	CTRL	4	8.5 %
	EXP	3	6.4 %
Limassol	CTRL	6	12.8 %
	EXP	8	17.0 %
Nicosia	CTRL	13	27.7 %
	EXP	13	27.7 %

# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.3. Séquence d'allocation non aléatoire

**Minimization (Blocked + Stratified Randomization)**

**Année 1, cohorte 1 (2025-2026)**

**Groupe expérimental**

- 17 + 12 classes belges
- 8 classes allemandes
- 17 classes françaises
- 10 classes suisses
- 24 classes chypriotes

**Prétest**

SES + Type de classe	
SES + Type de classe	
SES + Type de classe + département	
SES + Type de classe	
SES + ville + lieu de vie	

**Posttest**

**Groupe contrôle**

- 17 + 12 classes belges
- 8 classes allemandes
- 17 classes françaises
- 10 classes suisses
- 23 classes chypriotes



Ecoles au sein de la population qui respectent les critères d'inclusion dans l'étude

Assignment aléatoire



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.4. Facteur confondant

- Nombre de clusters  $\neq 1$
- Stratifications  $\rightarrow$  pays et le SES ne seront pas des facteurs confondants
- Equivalence vérifiée sur des variables socio-démographiques parfois récoltées
- Équivalence à vérifier sur les variables d'intérêt ...



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.5. Sélection des participants

- Documentation de l'attrition et des primo-arrivants, mais non évaluable à ce stade du projet



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Méthodes utilisées pour mesurer les variables et les résultats

Country	N students
Belgium	152
Cyprus	173
Germany	138
France	108
Switzerland	144
Italy	141
<b>Total</b>	<b>856</b>

- **Trois types** de mesure

- Mesure directe sur les enfants, les enseignants et les familles
- Mesure indirecte des enfants par les enseignants
- Mesure indirecte des enfants par les familles

- **Trois fonctions**

- Screening
- Progress monitoring
- Effets de l'intervention

- **Trois variables** d'intérêt

- Littératie précoce
- Numératie précoce
- Compétences SEB

#	Phase	#	Step
1	Item development	1	Identification of the domain and items generation
		2	Assessment of content validity by experts
2	Scale development	3	Qualitative pre-testing
		4	Quantitative pre-testing
		5	Item reduction
		6	Exploratory factor analysis
3	Scale evaluation	7	Dimensionality analysis and creation of scales scores
		8	Tests of reliability
		9	Tests of criterion validity

## 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

### 4.6. Méthodes utilisées pour mesurer les variables et les résultats

- Validité : instruments sont créés/choisis par des experts internationaux des domaines ciblés
- Fiabilité : AFE, AFC, IRT et indices de fidélité
- Standardisation : oui
- Suralignement/inhérence : pour les screening et le monitoring difficile de faire sans ...
- Indépendance : tests de littératie et de numéracie créés par les chercheurs
- Procédure d'administration cohérente : administrateur externe de tests et/ou enseignants
- Pré-spécification : lors d'une phase pilote

# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Méthodes utilisées pour mesurer les variables et les résultats

		Never	Rarely	Sometimes	Often	Very often	Almost always
TOCAC 1	<b>1. Concentrates</b>	<b>Concentration problems</b>					
TOCAC 2							
TOCAC 3							
TOCAC 4							
TOCAC 5							
TOCAC 6	<b>6. Breaks rules</b>	<b>Disruptive behavior</b>					
TOCAC 7							
TOCAC 8							
TOCAC 9							
TOCAC 10	<b>9. Gets angry when provoked by other children</b>	<b>Disruptive behavior</b>					
TOCAC 11							
TOCAC 12							
TOCAC 13	<b>13. Shows empathy and compassion for others feelings</b>	<b>Prosocial behavior</b>					
TOCAC 14							
TOCAC 15							
TOCAC 16	<b>16. Stops and calms down when angry or upset</b>	<b>Emotion Regulation Problems</b>					
TOCAC 17							
TOCAC 18							
TOCAC 19							
TOCAC 20	<b>20. Is easily upset</b>	<b>Internalizing problems</b>					
TOCAC 21							
TOCAC 22							
TOCAC 23							
TOCAC 24							
TOCAC 25	<b>25. Is worries</b>						

# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Méthodes utilisées pour mesurer les variables et les résultats

		Never	Rarely	Sometimes	Often	Very often	Almost always
TOCAC 26	26. Identifies when he/she need help with a task	1	2	3	4	5	6
TOCAC 27	27. Demonstrates confidence in activities he/she has mastered	1	2	3	4	5	6
TOCAC 28	28. Identifies emotions	1	2	3	4	5	6
TOCAC 29	29. Recognizes strengths and limitations	1	2	3	4	5	6
TOCAC 30	30. Links his/her emotions to specific events.	1	2	3	4	5	6
TOCAC 31	31. Respects what others think.	1	2	3	4	5	6
TOCAC 32	32. Notifies others when they are upset, worried or need help.	1	2	3	4	5	6
TOCAC 33	33. Respects others' feelings and ideas.	1	2	3	4	5	6
TOCAC 34	34. Consoles a sad friend.	1	2	3	4	5	6
TOCAC 35	35. Thanks those who help.	1	2	3	4	5	6
TOCAC 36	36. Understands the consequences of his/her actions on the feelings of others.	1	2	3	4	5	6



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Méthodes utilisées pour mesurer les variables et les résultats

```
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = r_poly)
Overall MSA = 0.93
MSA for each item =
  T_TOCA_1  T_TOCA_2  T_TOCA_3  T_TOCA_4  T_TOCA_5  T_TOCA_6  T_TOCA_7  T_TOCA_8
    0.88    0.88    0.96    0.94    0.92    0.95    0.94    0.90
  T_TOCA_9  T_TOCA_10  T_TOCA_11  T_TOCA_12  T_TOCA_13  T_TOCA_14  T_TOCA_15  T_TOCA_16
    0.95    0.89    0.94    0.92    0.95    0.94    0.94    0.96
T_TOCA_17  T_TOCA_18  T_TOCA_19  T_TOCA_20  T_TOCA_21  T_TOCA_22  T_TOCA_23  T_TOCA_24
    0.94    0.93    0.89    0.91    0.96    0.88    0.86    0.91
T_TOCA_25  T_TOCA_26  T_TOCA_27  T_TOCA_28  T_TOCA_29  T_TOCA_30  T_TOCA_31  T_TOCA_32
    0.88    0.96    0.96    0.94    0.95    0.93    0.94    0.96
T_TOCA_33  T_TOCA_34  T_TOCA_35  T_TOCA_36
    0.95    0.95    0.95    0.96
```

```
> corstest.bartlett(r_poly, n = 964)
$chisq
[1] 37766.06

$p.value
[1] 0

$df
[1] 630
```

## 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

### 4.6. Méthodes utilisées pour mesurer les variables et les résultats

```
> fa.parallel(x=r_poly, n.obs=964, fm="pa", plot=FALSE, n.iter=2e3, sim=TRUE, quant=.95)
Parallel analysis suggests that the number of factors = 6 and the number of components = 4
```

```
The Velicer MAP achieves a minimum of 0.02 with 6 factors
BIC achieves a minimum of 1752.03 with 6 factors
Sample Size adjusted BIC achieves a minimum of 3114.53 with 6 factors
```





# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Méthodes utilisées pour mesurer les variables et les résultats

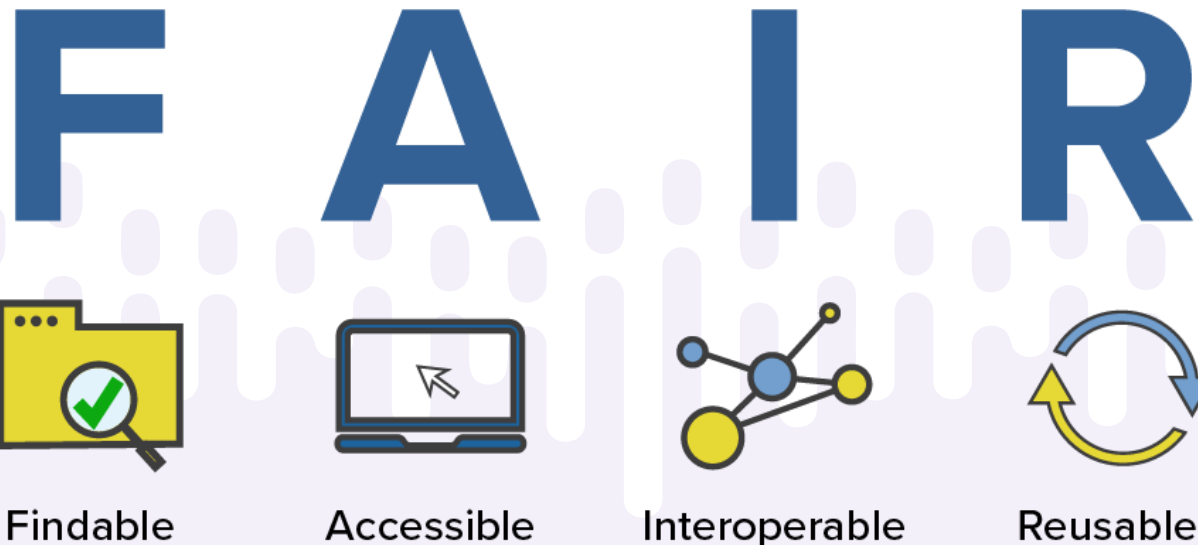
<b>Country</b>	<b>Dimension</b>	<b>Items included</b>	<b>N</b>	<b>Alpha</b>	<b>Mc Donald Omega</b>
All	All	TOCA_1 to TOCA_36	624	0.95	0.97
All	Concentration problems	TOCA_1 to TOCA_5	807	0.91	0.94
All	Disruptive behavior	TOCA_6 to TOCA_11	812	0.90	0.94
All	Prosocial behavior	TOCA_12 to TOCA_15	810	0.78	0.83
All	Emotion Regulation Problems	TOCA_16 to TOCA_20	789	0.85	0.89
All	Internalizing problems	TOCA_21 to TOCA_25	808	0.83	0.86
ALL	SEL	TOCA_26 to TOCA_36	725	0.95	0.97
ALL	SEL - Self Awareness	TOCA_26 to TOCA_30	756	0.91	0.93
ALL	SEL - Social Awareness	TOCA_31 to TOCA-36	742	0.95	0.96

# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Sélection des résultats rapportés

### FAIR (GoFair, n.d.) & Open Data :

- Findable = faciles à retrouver par tous (identificateur pérenne et unique, métadonnées riches)
- Accessible = authentification et autorisation
- Interoperable = intégration avec d'autres données et utilisation avec des applications d'analyse, de stockage et de traitement
- Reusable = description précise et complète



# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Sélection des résultats rapportés

Some rules were defined when creating the full database :

1. the first letter of the variable name indicates the domain
  - L\_ for literacy,
  - N\_ for numeracy,
  - C\_ for the cognitive domain,
  - T\_ for teacher assessment, and
  - P\_ for parent assessment
2. All items have been renamed so that their number always consists of two digits (1 -> 01)
3. Missing data :
  1. Empty cells = missing data that has not been "coded"
  2. NA = 777, 666 and "not applicable"
  3. Complete row of NAs = missing a complete instrument
  4. 888 and 999 have been recoded as 0
4. SEL = DASEC for Belgium, Germany and Cyprus ; SEL = TEC for France, Switzerland and Italy
5. No recoding of values was performed (except for the incorrectly coded VC variable).

# 4° Comment une RCT essaye-t-elle d'éviter ces biais ?

## 4.6. Sélection des résultats rapportés

```
1 # Installing R Packages
2 install.packages("readxl")
3 install.packages("haven")
4 install.packages("psych")
5 install.packages("ggplot2")
6 install.packages("dplyr")
7 install.packages("tidyr")
8 install.packages("lavaan")
9
10 # Loading R Packages
11 library(readxl)
12 library(psych)
13 library(haven)
14 library(ggplot2)
15 library(dplyr)
16 library(tidyr)
17 library(lavaan)
18 library(lavaanPlot)
19
20 # TOCA-C
21 # Reading the SAV File
22 DATAALL <- read_sav("DATAALL.SAV")
23
24 # Selecting variables to keep in the file for factorial analysis
25 names(DATAALL) # Names function lists the name of all variables in the file.
26 KeepCOL <- c("T_TOCA_1", "T_TOCA_2", "T_TOCA_3", "T_TOCA_4", "T_TOCA_5",
27             "T_TOCA_6", "T_TOCA_7", "T_TOCA_8", "T_TOCA_9", "T_TOCA_10",
28             "T_TOCA_11", "T_TOCA_12", "T_TOCA_13", "T_TOCA_14", "T_TOCA_15",
29             "T_TOCA_16", "T_TOCA_17", "T_TOCA_18", "T_TOCA_19", "T_TOCA_20",
30             "T_TOCA_21", "T_TOCA_22", "T_TOCA_23", "T_TOCA_24", "T_TOCA_25",
31             "T_TOCA_26", "T_TOCA_27", "T_TOCA_28", "T_TOCA_29", "T_TOCA_30",
32             "T_TOCA_31", "T_TOCA_32", "T_TOCA_33", "T_TOCA_34", "T_TOCA_35",
33             "T_TOCA_36") # Name of the variables to keep in the working data file
34
```

# 5° Bibliographie

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in public health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>

Bradshaw, C. P., Debnam, K. J., & Leaf, P. J. (2010). *Teacher Observation of Classroom Adaptation-Checklist (TOCA-C, Version 2)* [Unpublished measure]. Baltimore: Johns Hopkins University, Baltimore. Available for free on request by contacting Catherine Bradshaw at [bradsha@jhsph.edu](mailto:bradsha@jhsph.edu)

Cochrane Methods Bias. (2020). *RoB 2: A revised Cochrane risk-of-bias tool for randomized trials*. Retrieved from: <https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials>

GoFair. (n.d.). *FAIR Principles*. Retrieved from: <https://www.go-fair.org/fair-principles/>

Hedges, L. V., & Rhoads, C. (2010). *Statistical power analysis in education research*. Institute of Education Sciences.

Gorard, S., Haut See, B., & Siddiqui, N. (2017). *The trials of Evidence-Based Education: the promises, opportunities and problems of trials in education*. A David Fulton Book.

Higgins, J. P. T., Savovic, J., Page, M. J., Sterne, J. A. C. (2019). *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2)*. Cochrane.

Slavin, R.E., Lake, C., Inns, A., Baye, A., Dacht, D., & Haslam, J. (2019). *A quantitative synthesis of research on writing approaches in grades 2 to 12*. Education Endowment Foundation.

What works clearinghouse. (2020). *Procedures Handbook, Version 4.1*. Washington, DC : IES

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at <https://ies.ed.gov/ncee/wwc/Handbooks>