

**ROYAUME DU MAROC**

**UNIVERSITE SIDI MOHAMMED BEN ABDELLAH**

-----

**FACULTE DE MEDECINE ET DE PHARMACIE DE FES**



**Projet de traduction de l'ADI-R (Autism Diagnostic Interview - Revised) en arabe marocain**

**Mémoire de fin de spécialité**

Présenté par :

**Dr. Benbrahim Mohammed**

Né le **28/05/1980**

Sous la direction du

**Pr R. AALOUANE**

Pour l'obtention du diplôme de fin de spécialité en Pédopsychiatrie

Session Novembre 2016

# Remerciements

*MONSIEUR LE PROFESSEUR SMAIL RAMMOUZ*

*PROFESSEUR DE PSYCHIATRIE*

*Recevez ce travail en témoignage de mon respect profond.*

*Je suis reconnaissant pour votre apprentissage,*

*Veillez accepter, cher Maître, l'assurance de mon estime et profond respect.*

*MONSIEUR LE PROFESSEUR RACHID AALOUANE  
PROFESSEUR AGREGÉ DE PSYCHIATRIE*

*Je vous remercie pour la gentillesse et la spontanéité avec lesquelles vous avez bien voulu diriger ce travail.*

*J'ai eu le grand plaisir de travailler sous votre direction, et J'ai trouvé auprès de vous le conseiller et le guide qui m'a reçu en toute circonstance avec sympathie, sourire, bienveillance et patience inépuisable.*

*Votre simplicité, votre compétence, et vos qualités humaines et professionnelles font que vous serez toujours un exemple pour nous.*

*MADAME LE PROFESSEUR CHADYA AARAB  
PROFESSEUR ASSISTANT EN PSYCHIATRIE*

*Votre compétence professionnelle ainsi que vos qualités humaines ont toujours suscité mon admiration.*

*Qu'il me soit permis en ce jour, de vous exprimer, mon profond respect et ma très haute considération.*

## Résumé

**Projet de traduction de l'ADI-R (Autism Diagnostic Interview – Revised) en arabe marocain : Revue de littérature sur la validité transculturelle de l'ADI-R et méthodologie de la traduction, l'adaptation socioculturelle et la validation de la future version marocaine du questionnaire**

L'utilisation d'outils de diagnostic standard dans l'évaluation des Troubles du Spectre de l'Autisme est essentielle à l'avancement de la recherche, la pratique clinique et la prise en charge des personnes atteintes de TSA. Un des instruments standards, celui qui est souvent utilisé en combinaison avec un outil d'observation est l'*Autism Diagnostic Interview-Revised* ADI-R. Cet outil a été décrit comme le "gold standard" de diagnostic des TSA dans la pratique et la recherche internationale.

Le but de ce travail est de :

1. Décrire l'ADI-R, son intérêt et les modalités de sa passation ;
2. Faire une revue de littérature sur sa validité et sa fiabilité en tant que test diagnostique de l'autisme ;
3. Et élaborer une stratégie et une méthodologie pour : sa traduction à l'arabe dialectale, son adaptation à notre contexte socio-culturel et sa validation dans les temps qui suivent.

Ce projet de traduction de l'ADI-R à l'arabe marocain vise l'amélioration de la pratique clinique et la recherche scientifique dans le domaine de l'autisme au Maroc.

La traduction, l'adaptation et la validation de l'ADI-R en arabe dialectale marocaine aura certainement un impact positif direct sur la prise en charge basée sur les preuves et sera à la longue un promoteur essentiel pour pousser davantage la recherche dans le domaine de l'autisme.

# Sommaire

Remerciements .....	2
Résumé .....	6
Sommaire.....	8
Abréviations.....	10
Introduction .....	12
Objectifs .....	15
Chapitre 1 : Qu'est-ce que l'ADI-R ? .....	17
I. Le questionnaire : .....	18
II. Historique .....	21
III. Données psychométriques .....	22
IV. Utilisations cliniques .....	24
Chapitre 2 : Validité de l'ADI-R à l'échelle internationale.....	29
I. Validité de l'ADI-R .....	30
II. Validité de l'ADI-R à travers les cultures.....	32
III. Influence des facteurs socioculturels sur l'évaluation des TSA .....	32
Chapitre 3 : Traduction, adaptation et validation des échelles et questionnaires.....	34
I. Contexte et signification .....	35
II. Approches méthodologiques.....	36
Chapitre 4 : Traduction et Validation de l'ADI-R en arabe dialectale .....	46
I. Traduction de l'ADI-R en arabe dialectale .....	47
II. Adaptation de l'ADI-R au contexte socio-culturel marocain .....	49
III. Projet de validation de la version marocaine de l'ADI-R.....	52
Conclusion.....	59
Références .....	61

## **Abréviations**

ADI-R	Autism Diagnostic Interview-Revised
ADI-R-VM	Autism Diagnostic Interview-Revised-Version Marocaine
ADI-TS	Autism Diagnostic Inventory-Telephone Screening
ADOS	Autism Diagnostic Observation Schedule-Generic
BNV	Domaine B pour les sujets non verbaux
BV	Domaine B pour les sujets verbaux
CARS	Childhood Autism Rating Scale
I-CVI	Item - Content Validity Index
ICC	Intraclass Correlation Coefficient
K <sub>w</sub>	Kappa pondéré
LC	Langue cible
LS	Langue source
QI	Quotient Intellectuel
S-CVI	Scale - Content Validity Index
TED	Trouble Envahissant du Développement
TED-NoS	Trouble Envahissant du Développement Non Spécifié
TSA	Troubles du Spectre de l'Autisme
VPN	Valeur prédictive négative
VPP	Valeur prédictive positive
WISC	Wechsler Intelligence Scale for Children
WNV	WISC Non Verbal
WPS	Western Psychological Services

# Introduction

Les études épidémiologiques européennes et américaines révèlent une augmentation de la prévalence des syndromes autistiques [1]. Malheureusement, On manque d'études épidémiologiques au Maroc, mais les données les plus récentes recueillies dans d'autres pays indiquent une prévalence entre 0.6% et 1% [2-4], dont un tiers environ de formes typiques d'autisme. Si l'on extrapole ces données au Maroc, on peut estimer qu'il y aurait au Maroc de 338.000 à 563.000 [5] personnes atteintes du syndrome autistique au sens le plus large du terme « Troubles du Spectre l'Autisme » (TSA), dont 108.000 à 180.000 enfants. 6.400 à 10.680 nouveau-nés par an développeront ce handicap.

L'utilisation d'outils de diagnostic standard dans l'évaluation des Troubles du Spectre de l'Autisme est essentielle à l'avancement de la recherche, la pratique clinique et la prise en charge des personnes atteintes de TSA.

Un des instruments standards, celui qui est souvent utilisé en combinaison avec un outil d'observation est l'*Autism Diagnostic Interview-Revised* ADI-R [6,7]. Cet outil a été décrit comme le "gold standard" [8-12] de diagnostic des TSA dans la pratique et la recherche internationale.

L'ADI-R est un questionnaire structuré qu'un professionnel qualifié mène avec les parents des personnes faisant l'objet d'une évaluation de l'autisme. Ce questionnaire est utilisé par les chercheurs et les cliniciens depuis plusieurs années. Il peut être utilisé à des fins de diagnostic pour toute personne ayant un âge mental d'au moins 18 mois [6] et il évalue les comportements dans les domaines de la communication et le langage, l'interaction sociale réciproque, et les comportement stéréotypés et restreints.

L'ADI-R est l'un des instruments les plus fréquemment utilisés dans la recherche et les publications dans le domaine de l'autisme [13]. Ses propriétés diagnostiques et sa validité sont très bien documentées [8]. C'est un outil de diagnostic utile pour distinguer entre les enfants atteints d'autisme et les enfants souffrant d'autres troubles du langage [14]. La classification diagnostique ADI-R reste relativement stable au fil du temps dans les

études prospectives [15]. Cependant, l'ADI-R ne classifie pas l'autisme en ses formes légères, modérées ou sévères [16].

L'ADI-R a été traduit à plus de trente langues [11], il a été rigoureusement évalué et validé dans de nombreux pays [17-22]. Il a été traduit en arabe ; Une équipe de professionnels bilingues (anglais/arabe) de l'Arabie Saoudite, l'Egypte et le Qatar a été chargée de développer une version arabe de l'ADI-R. L'instrument a ensuite été examiné par un professionnel qualifié et puis a été retraduit afin de s'assurer que les deux versions sont identiques [23].

Dans ce manuscrit, l'usage du terme TSA fait plutôt référence à la constellation des troubles apparentés à l'autisme et pas nécessairement à la catégorie nosographique nouvellement définie par le DSM-5. Rappelons que le concept TSA est très ancien, il date des années 40 [24] juste quelques temps après la description des premiers cas d'autisme par Leo Kanner en 1943 [25].

# Objectifs

Le but de ce travail est de :

4. Décrire l'ADI-R, son intérêt et les modalités de sa passation ;
5. Faire une revue de littérature sur sa validité et sa fiabilité en tant que test diagnostique de l'autisme ;
6. Elaborer une stratégie et une méthodologie pour sa traduction à l'arabe dialectale, son adaptation à notre contexte socio-culturel et sa validation dans les temps qui suivent.

La seule version arabe disponible actuellement de cet outil a été élaborée au centre de recherche et hôpital spécialisé du Roi Faisal en Arabie saoudite [23]. Son usage est encore limité à la recherche et son adaptation et validation n'ont pas été encore appuyées par des études scientifiques rigoureuses.

Ce projet de traduction de l'ADI-R à l'arabe marocain vise l'amélioration de la pratique clinique et la recherche scientifique dans le domaine de l'autisme au Maroc.

La traduction, l'adaptation et la validation de l'ADI-R en arabe dialectale marocaine aura certainement un impact positif direct sur la prise en charge basée sur les preuves et sera à la longue un promoteur essentiel pour pousser davantage la recherche dans le domaine de l'autisme.

## **Chapitre 1 : Qu'est-ce que l'ADI-R ?**

## I. Le questionnaire :

L'ADI-R est un hétéro-questionnaire standardisé semi-structuré administré par des cliniciens qualifiés aux parents ou aux aidants des personnes adressées pour suspicion de TSA. C'est un outil qui permet de faire le diagnostic de l'autisme et planifier sa prise en charge. Il est très utile pour distinguer l'autisme des autres troubles du développement. L'ADI-R couvre l'histoire du développement du sujet. Il est habituellement réalisé dans un bureau, à la maison ou un autre cadre calme par un psychologue ou autre professionnel qualifié.

L'ADI-R comprend 93 items dans trois domaines de fonctionnement : la communication et le langage, la qualité des interactions sociales réciproques et les comportements dits restreints, répétitifs et stéréotypé, ce dernier domaine est communément appelé « la troisième sphère ». Le questionnaire comprend également d'autres éléments pertinents pour la planification du traitement, tels que l'automutilation et l'hyperactivité.

L'enquêteur enregistre et cote les réponses de l'informateur selon une procédure standardisée. Les questions couvrent huit domaines essentiels :

1. Les antécédents : la famille, l'éducation, les diagnostics précédents et les médicaments pris ;
2. Le comportement global ;
3. Les points de repère essentiels du développement ;
4. L'acquisition du langage et d'autres compétences ;
5. Fonctionnement actuel (langage et communication) ;
6. Interaction social et jeu ;
7. Intérêts et comportements ;
8. Comportements cliniquement significatives : agressivité, automutilation, épilepsie

... Tous les items de l'ADI-R sont codés pour les comportements actuels et passés, 'actuels' faisant référence aux comportements ayant lieu dans les 3 derniers mois. Pour certains items, 'passé' signifie que le comportement en question a eu lieu au moins une fois, alors

que pour d'autres on se demande si le comportement était présent pendant une période spécifiquement définie, entre 4 et 5 ans par exemple ( cette période est souvent appelée 'la plus anormale 4 à 5' parce que même les enfants ayant une déficience intellectuelle qui fonctionnent à moins de la moitié de leur âge chronologique devraient régulièrement présenter ces comportements à l'âge de 4 ans ) [26].

42 items de l'ADI-R sont systématiquement combinés pour produire un algorithme de diagnostic de l'autisme sur la base des critères de la CIM-10 [27] et du DSM-IV [28] comme il est spécifié par les auteurs. Outre les trois sphères du comportement, L'ADI-R comporte un quatrième domaine : Anomalie du développement évidente avant l'âge de 36 mois, pour indiquer si l'enfant répond aux critères de l'âge d'apparition.

L'entretien commence par une question d'introduction suivie par des questions sur le développement du sujet. Puis 41 questions portent sur la communication verbale et non verbale [11]. Les questions 50 à 66 concernent le développement social et le jeu. Les 13 questions suivantes portent sur les intérêts et les comportements. Les 14 dernières questions portent sur '*le comportement global*' y compris des questions sur les compétences mnésiques, les habilités motrices, l'hyperactivité et les crises.

Les évaluations sont basées sur le jugement du clinicien qui note les impressions de l'aidant sur le comportement et le développement du sujet. Pour chaque item, le clinicien donne un score allant de 0 (non autiste) à 3 (autiste).

Chaque domaine a un seuil '*cutoff*'; le sujet doit atteindre ou dépasser les cutoffs dans les quatre domaines pour avoir un diagnostic ADI-R de "l'autisme". Des cutoffs à part sont disponibles pour le domaine de la communication, en fonction de si oui ou non l'enfant est verbal (défini comme ayant un score 0 à l'item 30 du niveau global du langage : Utilisation fonctionnelle d'un langage spontané, échoïque, ou stéréotypé compréhensible par d'autres personnes, impliquant des phrases de trois mots ou plus et incluant au moins un verbe) [29].

D'autres critères utilisant des cutoffs inférieurs avec le même ensemble d'items ont été utilisés pour créer un algorithme de classification plus large des troubles du spectre de l'autisme comme dans plusieurs études de collaboration [12,30,31]. L'algorithme de diagnostic pour les enfants de 4 ans et au-delà est basé sur les codes '*les plus anormaux*' [6,11,26,32-37], des formulaires de l'algorithme du '*comportement actuel*' [6,24,27,30,32,32] sont disponibles pour faciliter un diagnostic clinique pour les enfants de 2 ans et plus.

Les résultats peuvent à présent être marqués et interprétés en utilisant un seul formulaire pratique plutôt que les cinq formulaires précédemment requis [38]. Le formulaire de l'Algorithme complet (W-382E) permet de calculer et d'interpréter quelconque des cinq algorithmes ADI-R spécifiques à chaque âge (deux algorithmes de diagnostic basés sur l'histoire du développement et utilisés pour le diagnostic formel, et trois algorithmes du *comportement actuel* mettant l'accent sur le fonctionnement actuel et utilisés pour le planning de la prise en charge). Les algorithmes eux-mêmes n'ont pas changé : le formulaire révisé ne fait que remplacer les cinq formulaires précédemment nécessaires au calcul des algorithmes.

Une version pour les tout petits '*Toddler ADI-R*' [39] a été mise au point il y a quelques années afin de fournir des données descriptives pour la recherche chez les enfants de moins de 4 ans. La *Toddler ADI-R* comporte 125 items, dont 32 nouvelles questions, et des codes relatifs à l'apparition des symptômes de l'autisme et au développement général [29]. Les autres items sont identiques à l'ADI-R, sauf que la *Toddler ADI-R* n'a pas de codes pour les comportements entre 4 et 5 ans.

Des analyses antérieures ont suggéré que l'algorithme de diagnostic a été utile pour les enfants ayant un âge mental non verbale supérieure à 2 ans [6,7,16]. Ainsi, l'ADI-R avait été approprié pour l'évaluation diagnostique de toute personne au sein de la tranche d'âge s'étendant de la petite enfance à la vie adulte, à condition qu'elle ait un âge mental non verbale supérieure à 2 ans. Cependant, des algorithmes récemment développés pour les tout-petits et les jeunes enfants d'âge préscolaire ont montré une bonne validité prédictive

par rapport aux algorithmes préexistants pour les jeunes enfants de 12 à 47 mois d'âge [35]. Ces algorithmes étendent l'utilisation de l'ADI-R pour les enfants aussi jeunes que 12 mois et un niveau de développement non verbale d'au moins 10 mois. En outre, ces nouveaux algorithmes comprennent des éléments présents dans les deux versions de l'ADI-R (*standard* et *Toddler*), permettant l'utilisation des algorithmes avec n'importe laquelle des deux versions.

La plupart des éléments de l'ADI-R se rapportent à des comportements qui sont rares chez les personnes ne présentant pas de TSA et/ou ne présentant pas de déficience intellectuelle profonde. Ainsi, les estimations numériques des scores des enfants au développement normal de la population générale n'ont pas été obtenues [29]. Cependant, il y a eu plusieurs comparaisons avec les enfants et adolescents atteints d'autres troubles, ce qui a été utilisé dans le développement des algorithmes de diagnostic [6,16,35]. Des chercheurs ont utilisé des scores de domaines individuels ou le score global de trois domaines pour estimer la sévérité des symptômes autistiques [26], bien que la validité de cette approche n'a pas été directement testée. Les évaluations ont été publiées pour de nombreuses populations de recherche, mais n'ont pas été encore systématiquement standardisées [29].

Parce que l'ADI-R est un hétéro-questionnaire, les questions sont ouvertes et l'enquêteur est en mesure d'obtenir toutes les informations nécessaires pour déterminer une cote valide pour chaque comportement. Pour cette raison, les parents et aidants se sentent généralement très à l'aise lorsqu'ils participent à cet entretien parce que ce qu'ils ont à dire à propos de leurs enfants est évalué par intervieweur qualifié [6]. Aussi, prendre part à cet entretien aide les parents à obtenir une meilleure compréhension des TSA et des éléments qui aident au diagnostic.

## II. Historique

L'ADI-R a été développée en 1989 [16] initialement dans le but de la recherche scientifique [40]. Il a été ensuite modifié en 1994 [6]. Cette nouvelle version est un peu

plus courte que l'originale afin de rendre l'entretien plus réalisable dans les milieux cliniques et de la recherche. La version actuelle de l'ADI-R a été publiée en 2003 par les *Western Psychological Services (WPS)* [7].

Le développement de la version *Toddler ADI-R* a été achevé en 2006 [29] à des fins de recherche. Juste après, il y a eu une augmentation de la demande pour des instruments de diagnostic pour les très petits, ce qui a conduit au développement de nouveaux algorithmes de diagnostic pour les tout-petits et les jeunes enfants d'âge préscolaire [35]. Les algorithmes finaux pour cette tranche d'âge contiennent moins d'éléments que les algorithmes originaux et sont appropriés pour une utilisation chez les enfants de 12 à 47 mois [29].

### III. Données psychométriques

Les propriétés psychométriques de la version originale de l'ADI ont été rapportées pour un échantillon de 16 enfants et adultes atteints d'autisme et 16 enfants et adultes ayant une déficience intellectuelle ; chaque groupe comprenait des personnes de différents âges et QI (Quotient Intellectuel) de performance (L'âge moyen est de 12,28 ans avec un écart-type de 3,43 et un QI de performance de 43 à 71) [16].

Les participants ont été sélectionnés, interrogés et codés soigneusement et à l'aveugle. La fiabilité inter-juges a été évaluée, avec des kappas variant de 0,25 à 1 pour chaque élément. Les Coefficients ICC '*Intraclasse Correlations Coefficient*' étaient au-dessus de 0,94 pour tous les scores des domaines et sous-domaines. La majorité des items individuels ont montré une bonne validité discriminante entre le groupe de l'autisme et le groupe des personnes ayant une déficience intellectuelle sans autisme

[16]. Les propriétés psychométriques pour le développement des algorithmes de l'ADI-R ont été recueillies à partir d'échantillons de 25 enfants autistes et 25 enfants avec déficience intellectuelle qui ont été aussi sélectionnés, interviewés et codés à l'aveugle [6,7,29]. Ces enfants avaient des âges chronologiques entre 36 à 59 mois, avec des âges mentaux non

verbaux allant de 21 à 74 mois. En utilisant un échantillon de 10 enfants, la fiabilité inter-juges été évaluée ; les kappas étaient de 0,63 à 0,89 pour chaque item. En utilisant le même échantillon, les ICC étaient au-dessus de 0,92 pour tous les scores des domaines et sous-domaines. En outre, après la normalisation initiale de l'ADI-R, un échantillon distinct de 53 enfants autistes et 41 enfants non autistes ayant une déficience intellectuelle ou des troubles du langage ont été utilisés pour évaluer davantage la validité de l'ADI-R [41]. Les résultats de l'étude ont montré que la fiabilité inter-juges était aussi élevée que celle de l'étude initiale, avec des kappas allant de 0,62 à 0,96 pour les items individuels. La fiabilité test-retest était également très élevé, avec tous les coefficients dans la zone 0,93 au 0,97.

La majorité des items individuels de l'ADI-R dans sa forme actuelle a montré une bonne validité discriminante entre les enfants atteints d'autisme et les enfants ayant une déficience intellectuelle [6]. Les algorithmes actuels différencient les enfants autistes de plus de 36 mois des enfants souffrant d'autres troubles développementaux, avec une grande sensibilité et spécificité (les deux supérieurs à 90%).

D'autres analyses des données auprès des enfants d'âge préscolaire ont révélé que les algorithmes de l'ADI-R différencient les enfants de plus de 2 ans atteints de TSA de ceux avec d'autres troubles du développement. Cependant, pour les enfants de moins de 2 ans, la discrimination entre les enfants TSA non verbaux et les enfants non autistes non verbaux était pauvre, ce qui entraîne une faible spécificité, surtout pour les enfants d'âge mental inférieur à 18 mois [41].

Dans une étude plus récente comprenant un échantillon plus grand [12], l'ADI-R a montré une sensibilité élevée (supérieure à 80%) pour les enfants atteints de TSA de moins de 3 ans, mais plus faible spécificité pour la discrimination TSA/non TSA (environ 70 %). Ventola et al [42] ont rapporté que, pour les enfants entre 16 et 37 mois, les classifications diagnostiques effectuées en fonction de l'algorithme ADI-R ont montré une sensibilité plus faible que celles faites à l'aide de l'*Autism Diagnostic Observation Schedule* (ADOS) [43], du *Childhood Autism Rating Scale* (CARS) [44] et du jugement

clinique en utilisant les critères du DSM-IV. Wiggins et Robins [45] ont également constaté que les algorithmes ADI-R ont une faible sensibilité pour les enfants de la même tranche d'âge lorsque les cutoffs standards de la troisième sphère sont inclus dans les critères de diagnostic. Étant donné les faibles sensibilités et spécificités signalées pour les enfants très jeunes, de nouveaux algorithmes ADI-R ont été développés pour les tout-petits et les enfants d'âge préscolaire entre 12 à 47 mois en utilisant un échantillon de 491 enfants atteints de TSA, 136 avec des troubles du développement et 67 enfants neurotypiques [35]. Les nouveaux algorithmes ADI-R comportent deux cutoffs différents : un pour la recherche (plus restrictif avec une plus grande spécificité et une sensibilité plus faible) et un à des fins cliniques (plus inclusives avec une plus grande sensibilité et une spécificité plus faible). Ils comprennent également des "zones de suspicion" pour un usage clinique. Dans le dernier échantillon, la sensibilité selon les cutoffs cliniques variait de 80% à 94% et la spécificité de 70% à 81% pour la comparaison des TSA/non TSA. Selon les cutoffs de recherche, la comparaison montre à une sensibilité de 80% à 84% et une spécificité de 85% à 90% [29].

#### **IV. Utilisations cliniques**

L'ADI-R montre le profil d'un enfant, un adolescent ou un adulte en décrivant la qualité des interactions et réciprocités sociales, le langage et la communication, et les intérêts restreints et comportements répétitifs et stéréotypés. Les items sont évalués sur la base des descriptions détaillées des aidants sur l'histoire et les comportements de leur enfant, ce qui permet au clinicien de recueillir à la fois des informations à la fois quantitatives et qualitatives.

Une mise en garde importante pour les cliniciens est de reconnaître que les classifications diagnostiques basées sur des algorithmes et le diagnostic clinique réel ne sont pas pareilles. Le diagnostic clinique est basé sur de multiples sources d'information, y compris des observations directes [12,35,46]. Risi et al. [12] ont trouvé un meilleur équilibre de sensibilité et spécificité lorsque l'ADI-R et l'ADOS ont été utilisés en combinaison par rapport à quand chaque instrument a été utilisé seul. L'utilisation combinée de ces deux

instruments a donné lieu à une sensibilité de 82% et une spécificité de 86% pour les enfants TSA par rapport aux enfants atteints d'autres troubles du développement au-delà de 3 ans. Pour les enfants plus jeunes, la sensibilité et la spécificité utilisant les deux instruments étaient de 81% et 87%, respectivement. En revanche, lorsque chaque instrument a été utilisé seul, les spécificités ont varié de 59% à 72%.

Le Couteur et al. [46] ont également examiné l'utilisation combinée de l'ADOS et l'ADI-R pour les enfants TSA d'âge préscolaire en utilisant des algorithmes révisés de l'ADOS [47]. Conformément à l'étude de Risi de 2006 [12], les auteurs ont aussi constaté que la combinaison des informations à la fois de l'ADOS et de l'ADI-R améliorait la précision du diagnostic par rapport à l'usage de l'un ou l'autre instrument utilisé seul. De même, en utilisant les algorithmes nouvellement développés du *Toddler-ADI-R* pour les tout-petits et les jeunes enfants d'âge préscolaire, l'ADOS révisé et les nouveaux algorithmes *Toddler-ADOS*, Kim & Lord [35] ont également constaté que, pour les très jeunes enfants, l'utilisation combinée de l'ADOS et de l'ADI-R améliorait plus la validité du diagnostic par rapport à quand chaque instrument a été utilisé seul. Ainsi, même si l'ADI-R fournit des informations sur l'histoire et la description de l'individu et de son fonctionnement actuel dans un large éventail de contextes, l'ADI-R ne peut pas être utilisé seul pour poser un diagnostic clinique.

Les cutoffs de l'algorithme de diagnostic permettent une classification des TSA basée sur les patrons du comportement en se référant aux critères du DSM-IV et de la CIM-10 pour les troubles autistiques. En plus des scores de cutoff uniques, les nouveaux algorithmes pour les tout-petits et les jeunes enfants d'âge préscolaire fournissent aux cliniciens et aux chercheurs plusieurs options différentes pour la classification diagnostique des jeunes enfants.

Pour des raisons cliniques, des zones de suspicion (peu à aucune suspicion, suspicion légère à modérée, et suspicion modérée à sévère) qui représentent la sévérité des symptômes autistiques chez les jeunes enfants sont également fournis. Le clinicien ou le chercheur peut utiliser ces zones afin d'éclairer les décisions quant à savoir si un enfant

doit bénéficier d'autres évaluations ou devrait être rapidement adressé aux services de prise en charge indépendamment des cutoffs de diagnostic.

Les scores de la zone 'peu à aucune suspicion' indiquent que l'enfant n'a pas plus de comportements associés aux TSA que les autres enfants sans TSA de la même tranche d'âge. Au contraire, un enfant marqué dans la gamme de 'suspicion légère à modérée' a un certain nombre de comportements compatibles avec un TSA sans qu'ils soient spécifiques à l'autisme pour autant. Pour des raisons cliniques, les enfants de la zone de suspicion 'légère à modérée' devraient recevoir une nouvelle évaluation et un suivi, d'autres évaluations cognitives et linguistiques sont souvent très utiles, ainsi que les recommandations de prise en charge.

En plus des zones de suspicion, les scores de cutoff peuvent être utilisés séparément lorsque des groupements strictement stratifiés sont nécessaires, tels que pour des interventions, pour la neuro-imagerie ou pour la recherche génétique [29]. Ces différentes solutions permettent aux cliniciens et aux chercheurs d'être transparents sur les choix qu'ils font, tout en reconnaissant que les décisions de diagnostic de TSA chez les très jeunes enfants sont moins stables et précises que pour les enfants plus âgés et les adolescents.

Outre les algorithmes de diagnostic, l'ADI-R comprend un formulaire pour l'algorithme du comportement actuel qui peut être utilisé dans des milieux cliniques afin d'évaluer les changements qui se produisent pendant ou après des interventions ou qui peuvent refléter une maturité croissante du développement ou une évolution des conditions de vie [19].

L'ADI-R fournit une structuration utile des éléments de l'anamnèse et permet ainsi de comprendre la perspective de l'aidant des symptômes autistiques de son enfant. Cependant, il nécessite une formation spécifique quant à l'administration et la cotation. L'ADI-R ne doit être utilisé que par des cliniciens expérimentés qui sont familiers avec les TSA et leurs symptômes relatifs. Des ateliers de formation et des bandes vidéo sont disponibles pour aider les cliniciens et les chercheurs à comprendre la cotation et l'administration de l'ADI-R. Pour la recherche, les enquêteurs doivent répondre aux normes de fiabilité. Un clinicien

qualifié peut administrer l'ADI-R au parent d'un enfant soupçonné d'autisme dans environ 2 heures et demi [22].

Des études ont prouvé également la possibilité de passer le questionnaire ADI-R en sa forme originale [48] ou modifiée ADI-TS [49] ( *Autism Diagnostic Inventory-Telephone Screening* ) par téléphone avec des résultats très prometteurs en terme de fiabilité et de validité.

Il faut rappeler que le diagnostic posé sur la base de l'ADI-R n'a pas à se conformer au diagnostic clinique. Ce dernier est basé sur de nombreuses sources d'information, y compris des données à partir des observations directes. Alors que l'ADI-R est un outil standardisé pour l'obtention d'une grande quantité d'informations et son utilité dans la recherche est prouvée. A la lecture des publications scientifiques occidentales, en particulier dans le domaine de la biologie de l'autisme, on se rend compte que très souvent les auteurs ne posent pas le diagnostic de l'autisme en se référant à l'ADI-R et/ou ADOS [19].

Parce que l'ADI-R est un questionnaire plutôt qu'un test [38], et parce qu'il met l'accent sur des comportements qui sont rares chez les individus non TSA, il fournit des résultats catégoriques plutôt que des échelles ou des normes. Les résultats peuvent être utilisés pour appuyer un diagnostic d'autisme ou déterminer les besoins cliniques de différents groupes à risque de TSA [38] (par exemple, les personnes ayant un trouble sévère du langage ou autres conditions médicales, les enfants atteints de cécité congénitale, et jeunes souffrant de privation institutionnelle). L'ADI-R est très efficace pour différencier l'autisme des autres troubles du développement et dans l'évaluation des cas limites du syndrome autistique en identifiant des nouveaux sous-groupes et en quantifiant la symptomatologie autistique. L'utilisation intensive de l'ADI-R dans la communauté de recherche internationale a fourni des preuves solides sur la fiabilité et la validité de ses résultats catégoriques.

Bien que le DSM-5 [42] a été publié avec des critères mis à jour pour les TSA (La communication sociale et les comportement restreints et stéréotypés), A ce jour, la plupart des enfants ont eu des diagnostics basés sur le DSM-IV dont les critères diagnostiques ont été à la base du questionnaire ADI-R [50].

## **Chapitre 2 : Validité de l'ADI-R à l'échelle internationale**

L'utilisation d'outils standards de diagnostic dans l'évaluation des troubles du spectre de l'autisme est essentielle à l'avancement de la recherche, la pratique clinique et la prise en charge des personnes atteintes de ces troubles [50]. *L'Autism Diagnostic Interview- Revised* (ADI-R) est un instrument standard souvent utilisé en combinaison avec un outil d'observation [6]. Cette association est considérée comme le '*golden standard*' dans l'évaluation des enfants et des adultes atteints de TSA [51]. Initialement développé comme un outil de recherche, l'ADI-R est devenu plus largement utilisé dans les milieux cliniques pour aider à l'évaluation diagnostique des individus suspectés de TSA.

Cette introduction de l'ADI-R dans la pratique clinique, ainsi que son utilisation extensive dans les études de recherche, sollicite une forte demande pour une évaluation complète de ses propriétés psychométriques allant au-delà de la normalisation au sein de quelques échantillons. À ce jour, plusieurs études à travers le monde ont évalué la validité de l'ADI-R à identifier les individus atteints de TSA. Cependant, la recherche sur l'utilisation de l'ADI-R et son applicabilité dans l'identification précise des individus atteints de TSA dans les populations non anglophones manquent considérablement [50].

Nous allons essayer dans ce chapitre de faire le point sur la validité de l'ADI-R à l'échelle internationale et plus particulièrement dans les pays non anglophones afin d'évaluer son efficacité dans l'identification des enfants et des adolescents atteints de TSA à travers d'autres cultures.

## **I. Validité de l'ADI-R**

Plusieurs études internationales ont exploré les propriétés psychométriques de l'ADI-R chez des sujets avec et sans TSA chez tous les groupes d'âge et à travers les différents niveaux de capacité intellectuelle [50]. L'étude originale de validation de l'ADI-R 'en langue anglaise' a été menée par Lord et al. [6], elle a comparé les enfants atteints d'un trouble autistique aux enfants souffrant de retard mental ou des troubles du langage. Les résultats de cette étude suggèrent que l'ADI-R a une sensibilité élevée (96%) et une spécificité aussi élevée (92%).

Plusieurs études ont montré également la validité et la fiabilité de l'ADI-R chez les tout-petits et les enfants d'âge préscolaire, même si les chiffres chez cette population sont moins impressionnants que ceux des enfants d'âge supérieur [52,53]. Ces résultats sporadiques sont repris par une revue systématique récente [54] montrant qu'en moyenne la sensibilité de l'ADI-R est beaucoup plus faible chez les enfants de moins de 3 ans (82%) que chez les enfants d'âge supérieur (91%).

D'autres études ont mis en évidence que le niveau intellectuel de l'individu peut avoir une influence directe sur les scores ADI-R, avec des taux de faux positifs plus élevés lorsque les sujets évalués présentent une déficience intellectuelle [55]. Cependant, d'autres études ont rapporté des taux de sensibilité et spécificité adéquats de l'ADI-R dans l'identification de TSA chez les enfants et les adultes avec retard mental [56,57].

En gros, la validité de l'ADI-R dans sa version originale anglaise a été prouvée dans les situations communes et pour des âges supérieurs à 3 ans. Cependant, beaucoup d'études varient en fonction de la gravité du TSA (Autisme, le Trouble Envahissant du Développement Non Spécifié Autrement ...), des âges d'inclusion (tout-petits, enfants d'âge préscolaire, les jeunes), et des éléments de comparaison (algorithmes, scores de domaine, classification de diagnostic) [50]. Les études menées à ce jour montrent que certains facteurs comme le QI des sujets [54] et les conditions de l'entretien ( par exemple l'ordre des questions concernant les comportements actuels et passés) [32] peuvent influencer les résultats de l'ADI-R. Ces études suggèrent que l'ADI-R même dans sa version originale en langue anglaise peut présenter quelques incohérences sous certaines conditions. Il n'est donc pas tout à fait claire comment les facteurs linguistiques et culturels peuvent influencer l'efficacité de l'ADI-R [50]. Ainsi, il est essentiel de déterminer comment ces questions peuvent être redressées lorsque des facteurs socioculturels supplémentaires entrent en jeu.

## II. Validité de l'ADI-R à travers les cultures

A présent, l'ADI-R est publié et commercialisé en 17 langues [38]; cependant, le processus par lequel l'ADI-R est traduit de l'anglais et adapté aux différentes langues et cultures est très variable [50]. Pour déterminer si l'utilisation de l'ADI-R à travers les cultures est valide, les études devraient évaluer non seulement l'utilisation de l'ADI-R comme outil de diagnostic, mais aussi d'évaluer son équivalence dans la culture cible [58].

Les évaluations des versions traduites de l'ADI-R dans leurs cultures respectives ont donné des informations utiles, mais quelque peu contradictoires [50]. Au Brésil, Becker et al. [18] ont rapporté une sensibilité élevée (100%) et une spécificité parfaite (100%) dans un échantillon clinique d'enfants et d'adolescents avec des diagnostics de TSA et de déficience intellectuelle modérée. Ces taux peuvent être dus à l'homogénéité du groupe de contrôle supérieure à celle des autres études et aussi aux différences dans la façon dont l'ADI-R a été administré et coté [50]. Tsuchiya et al. [17] ont montré que la sensibilité et la spécificité de la version japonaise de l'ADI-R étaient inférieures, mais quand même élevées (92% et 89% respectivement) dans un échantillon clinique et communautaire mixte d'enfants et d'adolescents.

Dans un échantillon clinique grec d'enfants et de jeunes, la sensibilité était élevée (88%); mais la spécificité était beaucoup plus faible (69%) pour identifier l'autisme avec la version grecque de l'ADI-R [59]. Ces derniers résultats sont similaires à ceux rapportés dans un échantillon finlandais de diagnostics basés sur le registre national de l'autisme, avec une sensibilité élevée de 95%, cependant, une spécificité beaucoup plus faible de 74% [21].

Ces résultats mettent en évidence la nécessité d'étudier davantage l'ADI-R avant son utilisation à travers des cultures et des langues différentes.

## III. Influence des facteurs socioculturels sur l'évaluation des TSA

Les facteurs socioculturels peuvent jouer un rôle important dans l'utilisation des outils de diagnostic du TSA [11,50]. Des travaux récents sur l'adaptation culturelle des outils de

diagnostic et d'éducation ont constaté que les perceptions culturelles de tous les handicaps en général et de l'autisme en particulier peuvent varier considérablement et que celles-ci peuvent modifier la façon dont les parents comprennent et rapportent le développement de leur enfant [60,61]. Par exemple, les familles mexicaines peuvent percevoir le retard du développement du langage comme anodin, ce qui peut influencer et retarder la procédure de diagnostic et de prise en charge [62].

En outre, un comportement jugé acceptable et attendu des jeunes enfants peut varier considérablement entre les cultures et parfois même au sein de la même culture. Par exemple, l'un des items clés de l'ADI-R évalue l'absence du contact visuel de l'enfant. Dans de nombreuses cultures asiatiques et latino-américaines, le contact visuel est jugé inapproprié pour un enfant [11,50] et, par conséquent, le manque de contact avec les yeux peut ne pas être indicatif d'une symptomatologie autistique sous-jacente, mais plutôt du respect des normes sociales par l'enfant [50].

Par conséquent, les études antérieures comparant l'ADI-R dans des contextes culturels différents, peuvent avoir été intrinsèquement biaisées du fait que l'équivalence des items individuels, des domaines et des scores totaux n'a pas été confirmée [50]. Pourtant, en tant qu'outil de diagnostic, l'ADI-R doit distinguer entre les enfants atteints de TSA et ceux avec d'autres handicaps ou déficiences au sein d'une culture ou d'une population donnée.

## **Chapitre 3 : Traduction, adaptation et validation des échelles et questionnaires**

## I. Contexte et signification

La mondialisation a permis de mettre en évidence la diversité croissante de la population dans un monde où le contexte des différents pays est extrêmement diversifié quant à la culture, la langue et l'appartenance ethnique [63,64]. Par conséquent, il est nécessaire pour une recherche interculturelle pertinente de discuter et résoudre les problèmes de ces disparités à travers ces populations multinationales et multiculturelles [65].

Les chercheurs en santé et qui mènent des études interculturelles doivent avoir accès à des instruments fiables à validation croisée dans d'autres cultures et/ou dans d'autres langues [66-68]. Les résultats de la recherche interculturelle peuvent avoir de grandes implications cliniques pour les médecins, les infirmiers, les psychologues et autres professionnels de santé qui fournissent des soins aux diverses populations parce que la prestation des soins de qualité dépend d'une évaluation précise et d'une meilleure compréhension du contexte culturel, linguistique et ethnique des individus.

Bien qu'il existe bel et bien des approches méthodologiques pour la traduction, l'adaptation et la validation des instruments pour une utilisation dans la recherche interculturelle dans le domaine de la santé [64,66,68-76], une grande variation dans l'utilisation de ces guidelines continue à dominer la littérature scientifique [65].

Une revue récente de 47 études méthodologiques portant sur la traduction et la validation des instruments pour la recherche interculturelle a rapporté que la qualité et les approches méthodologiques des études examinées variaient considérablement [77]. Il n'y avait pas de consensus clair parmi les chercheurs sur la façon dont les approches doivent être utilisées ou combinées. On note également une grande variation sur les qualifications des traducteurs et un manque d'informations détaillées sur le processus de traduction, la retraduction, la validation, les tests, la révision et les ajustements de l'instrument [65].

Dans le même sens, Sperber et al. [64] ont rapporté que, malheureusement, le processus de la traduction, de l'adaptation et de validation transculturelle d'un

instrument est considéré comme une étape de faible importance dans les protocoles des études. En outre, l'approche méthodologique la plus couramment utilisée et signalée était la traduction uni-sens sans retraduction, et souvent en utilisant un traducteur non qualifié [65]. Les mêmes auteurs [64] ont rapporté que seuls quelques chercheurs ont décrit les stratégies utilisées et les étapes ou les processus d'adaptation et/ou la validation des instruments. Les auteurs ont souligné également qu'il ne suffit pas de traduire un instrument sans évaluation approfondie de son adaptation et sa validation transculturelle.

Ces données suggèrent que beaucoup de chercheurs ne suivent pas les recommandations et les guidelines existants sur la nécessité d'un processus complet à plusieurs étapes pour la traduction, l'adaptation et la validation transculturelle des instruments [65]. Il se peut que les approches méthodologiques ne sont pas clairement présentées dans un format convivial, ce qui pourrait expliquer la difficulté des chercheurs à adopter et suivre les recommandations.

Par conséquent, l'objectif de ce chapitre est d'examiner les recommandations publiées et ainsi présenter les approches méthodologiques hautement recommandées pour la traduction, l'adaptation et la validation des instruments pour la recherche en soins de santé interculturelle dans un guide clair et convivial. Cela conduirait à une meilleure compréhension de l'utilisation de ces approches et donnerait quelques points de repère essentiels dans notre projet de traduction de l'ADI-R.

## **II. Approches méthodologiques**

Werner et Campbell [78] ont défini deux catégories de traduction: symétrique (ou décentrée) et asymétrique (ou uni-centrée). Le but de la traduction symétrique est la loyauté au sens et à l'égalité des familiarités dans chaque langue. Une traduction symétrique comprend des éléments de test communs aux deux (ou toutes) les cultures de l'étude. La traduction symétrique implique la traduction supra-linguistique des symboles culturels et est essentielle lorsque l'objectif de l'étude est comparatif. Si un item est spécifique à l'une des cultures, il doit être remplacé par un autre qui fait référence au même concept dans les

deux cultures. Le jugement sur la pertinence culturelle des items de l'instrument doit être fondé sur la connaissance approfondie des deux cultures et résulte souvent d'un travail qualitatif préliminaire. Un problème potentiel avec la traduction symétrique se pose lorsque l'instrument original a une histoire d'utilisation dans la langue originale et a des propriétés psychométriques liées au son, surtout lorsque le chercheur ne veut pas altérer le contenu des items originaux au moment de la traduction [71]. Le chercheur doit être conscient que l'amélioration de la validité d'une étude comparative par la traduction symétrique peut compromettre la fiabilité parce que certains éléments de l'instrument original peuvent nécessiter des modifications pour atteindre l'équivalence conceptuelle entre les cultures.

La traduction asymétrique souligne la fidélité à une langue, généralement la langue source (LS, d'origine) [78]. Elle est plus appropriée pour les études interculturelles dont les objectifs sont opérationnels. Les traductions asymétriques donnent souvent des versions traduites qui semblent exotiques et non naturelles dans la langue cible (LC), bien que la version de la LC est fidèle au texte original. La traduction asymétrique est la méthode appropriée si le but des tests est de mesurer les réponses d'une culture par rapport aux réponses dans une autre culture dans une situation spécifique à l'une des deux.

Dans notre cas, on opte pour une traduction symétrique car celle-ci préserve le sens tant dans la LS que dans la LC et ne constitue pas seulement une traduction littérale [74], le but de notre traduction étant d'obtenir l'équivalence du sens et des implications entre l'instrument dans la LS et l'instrument dans la LC [64].

La traduction symétrique est la seule méthode qui facilite la comparaison des réponses des individus d'une culture à celles d'une autre [67,74,79]. Elle permet également de déterminer les points les plus pertinents de l'équivalence interculturelle : sémantique, conceptuel, contenu, critères et techniques [63,80]. En outre, elle utilise le processus connu sous le nom de *Décentration* [69], dans lequel à la fois la LS et la LC d'un instrument sont aussi importantes l'une que l'autre.

Le décentrage se réfère à un processus de traduction dans lequel la LS et les versions de la LC ont la même importance au cours du processus traduction [69]. Une langue ne contient pas un contenu qui doit être traduit sans modification dans l'autre. Par exemple, un test de personnalité standardisé peut être modifié après traduction et retraduction dans une autre langue. Une telle modification pourrait être fondée sur la connaissance des termes qu'on ne peut pas bien traduire, des termes familiers (tels que '*take advantage of*' ou '*playing sick*') pourraient être modifiés. Dans le processus de décentrage, la version source et la version cible contribuent toutes les deux et de façon égale à la version finale des items et peuvent donc être sujet de révision.

Le processus de traduction, de l'adaptation et de la validation transculturelle d'un instrument destiné à être utilisé dans d'autres cultures, langues et pays exige la planification et l'adoption d'approches méthodologiques complètes, rigoureuses et plus établies [64,66,68-76]. Parce qu'il y a des variations entre ces approches, nous intégrons ici les plus recommandées dans un guideline convivial et cohérent avec les étapes essentielles.

### **1. Étape 1 : traduction de la version originale dans la langue cible**

L'instrument est initialement traduit à la LC (langue cible) par au moins deux traducteurs indépendants, de préférence certifiés [65,79], dont la langue maternelle est la LC souhaitée de l'instrument. Les traducteurs doivent être parfaitement bilingues et de préférence biculturels. Ils doivent provenir d'environnements différents. Le premier traducteur doit être familier avec la terminologie des soins de santé et du domaine du contenu de l'instrument dans la LC. Le second traducteur doit connaître les phrases familières, le jargon médical, les expressions idiomatiques et les termes émotionnels d'usage courant dans la LC souhaitée sans connaissance du domaine du contenu de l'instrument en question. Cette approche doit générer deux versions traduites LC1 et LC2 qui contiennent des mots et des phrases qui couvrent à la fois le domaine médical et la langue parlée habituelle avec ses nuances culturelles. Par conséquent, le choix des traducteurs qualifiés est la clé des traductions de haute qualité [65].

## **2. Étape 2 : comparaison des deux versions traduites (LC1 et LC2)**

Les instructions, les items et le format de réponse des deux versions traduites de l'instrument (LC1 et LC2) avec la version originale de l'instrument dans la LS sont d'abord comparés par un tiers bilingue et biculturel indépendant concernant les ambiguïtés et les incohérences des mots, des phrases et des significations. Ces ambiguïtés et incohérences doivent être discutées et résolues en utilisant une approche de comité [63,64,75,77,80]. Un consensus devrait être atteint avec la participation du troisième traducteur [80], les deux traducteurs de l'étape 1, et l'enquêteur et/ou d'autres membres de l'équipe de recherche. Ce processus doit générer une version préliminaire initiale traduite de l'instrument dans la LC (P-LC).

## **3. Étape 3 : Retraduction à l'aveugle de la version préliminaire initiale traduite**

Cette version préliminaire (P-LC) est retraduite de nouveau dans le LS par deux autres traducteurs indépendants ayant les mêmes qualifications, les caractéristiques et les conditions décrites et détaillées à l'étape 1 [65,69]. Pour cette étape, la langue maternelle des traducteurs mère devrait être la LS de l'instrument original, et ils devraient être complètement aveugles à la version originale de l'instrument (ils n'avaient jamais vu la version originale de l'instrument) [65]. Ils doivent produire deux versions retraduites de l'instrument dans la LS (R-LS1 et R-LS2). Cette étape permet de clarifier les mots et les phrases utilisés dans les traductions.

## **4. Étape 4 : Comparaison des deux versions retraduites**

Dans un premier temps, les instructions, les items et les formats de réponse des deux versions retraduites (R-LS1 et R-LS2) sont comparés par un comité multidisciplinaire avec leurs équivalents de l'instrument original dans la LS en ce qui concerne le format, le texte, et la structure grammaticale des phrases, la similitude du sens et sa pertinence. Il est fortement recommandé que le comité doit comprendre au moins un méthodologiste (qui

peut être le chercheur et/ou un membre de l'équipe de recherche), un professionnel de santé qui est familier avec le domaine du contenu et le concept de l'instrument, et tous les quatre traducteurs impliqués dans les étapes 1 et 3. Il est également recommandé qu'un des auteurs de l'instrument original de la LS participe en donnant ses idées sur la construction de la version retraduite [65].

Avoir au comité au moins un membre monolingue dont la langue maternelle est la LC améliorerait la qualité de la version pré-finale de l'instrument traduit. Toutes les ambiguïtés et les divergences concernant la signification culturelle, les expressions familières ou idiomatiques dans les mots et les phrases des instructions, les items, le format de réponse entre les deux versions retraduites (R-LS1 et R-LS2) ou vis-à-vis de l'instrument original dans la LS doivent être discutées et résolues par un consensus entre les membres du comité pour obtenir une version pré-finale de l'instrument dans la LC

Si les discordances ne peuvent être résolues, il peut être nécessaire de répéter les étapes 1 à 4 ; deux autres traducteurs bilingues et biculturels indépendants doivent être sollicités pour traduire l'instrument d'origine à nouveau pour générer deux traductions, et deux autres traducteurs bilingues et biculturels indépendants doivent être recrutés pour retraduire les versions traduites de l'instrument suivant la même procédure décrite ci-dessus. Sinon, seuls les éléments qui ne conservent pas leur sens original sont retraduits. L'évaluation des versions traduites et retraduites suit le même processus de validation déjà décrit. Ce processus est répété jusqu'à ce qu'aucune ambiguïté ou divergence persiste.

Cette approche méthodologique établira l'équivalence initiale conceptuelle, sémantique et du contenu de la version pré-finale :

- **L'équivalence conceptuelle** : degré auquel les concepts des items de l'instrument existent à la fois dans les deux cultures source et cible.
- **L'équivalence sémantique** : structure des phrases, des expressions familières ou idiomatiques qui assurent que le sens du texte ou l'idée des items de l'instrument dans la LS est présent dans la LC.

- **L'équivalence du contenu** : pertinence du texte ou l'idée des items de l'instrument dans chaque culture.

Le rôle du comité est d'évaluer, réviser et consolider les instructions, les items et le format de réponse des instruments ainsi traduits et de développer la version pré-finale pour des tests pilotes et psychométriques.

#### **5. Étape 5 : Essai pilote de la version pré-finale dans la langue cible avec un échantillon monolingue : débriefing cognitif**

La version pré-finale (P-FLC) est mise à l'essai chez des participants dont la langue est la LC de l'instrument afin d'évaluer les instructions, le format de réponse et les items pour plus de clarté. Les participants devraient être recrutés dans la population ciblée par l'instrument. Une taille d'échantillon de 10-40 individus est recommandée [66,81]. Chaque participant est invité à noter les instructions et les items de l'échelle en utilisant une échelle dichotomique (claire ou pas claire). Les participants qui évaluent des éléments de l'instrument comme peu claire sont invités à fournir des suggestions quant à la façon de réécrire les déclarations pour rendre le contenu plus clair. Les éléments de l'instrument que l'on trouve peu claire dans au moins 20% de l'échantillon doivent être réévalués [82]. Par conséquent, l'accord minimum inter-évaluateurs de l'échantillon est de 80%. Cette étape est utilisée pour soutenir davantage l'équivalence conceptuelle, sémantique et du contenu de l'instrument traduit et pour améliorer encore la structure des phrases utilisées dans les instructions et les items de la P-FLC pour être facilement comprises par la population cible avant de faire des tests psychométriques.

Pour déterminer davantage l'équivalence conceptuelle et sémantique des articles de la P-FLC, il est fortement recommandé de solliciter un groupe d'experts. Les éléments de l'instrument sont évalués pour l'équivalence conceptuelle (clarté) par six à dix membres d'un groupe d'experts [83,84], qui sont bien informés sur le domaine du contenu et concept de l'instrument et la population cible chez qui l'instrument sera utilisé. Les membres du groupe d'experts évalueront les éléments de l'entretien avec les mêmes

considérations et modalités de réévaluation que celles citées pour les participants au test pilote susmentionné.

On demande aux membres du groupe d'experts d'évaluer chaque élément de l'instrument pour l'équivalence du contenu (pour la validité et la pertinence du contenu) en utilisant l'échelle suivante :

- 1 = pas pertinent ;
- 2 = pertinence impossible à évaluer ;
- 3 = pertinent, mais a besoin de modification mineure ;
- 4 = très pertinent.

Les items classés comme 1 (non pertinent) ou 2 (incapables d'en évaluer la pertinence) devraient être révisés [83–85]. Les indices de validité du contenu au niveau de l'item (I-CVI) et au niveau de l'échelle (S-CVI) [85] doivent être calculés. Il y a trois méthodes pour calculer le S-CVI [83–85], mais la méthode de calcul de la moyenne (S-CVI/Ave) est la préférée [65,85]. En utilisant 10 experts, un I-CVI de 0,78 ou plus [83] et un S-CVI/Ave de 0,90 ou plus [84] sont les valeurs minimales acceptables. Les items dont ces indices n'atteignent pas les minimums acceptables sont révisés et réévalués. Puis des nouveaux indices de validité sont recalculés. Ce processus se poursuit jusqu'à ce que tous les minimums acceptables soient atteints. Il est également recommandé que le coefficient kappa pour l'accord entre évaluateurs soit déterminé afin d'accroître la confiance dans la validité du contenu de l'instrument [86]. Un kappa de 0,60 est généralement le coefficient minimum acceptable pour déterminer un bon accord [87].

Le but de cette étape est de continuer à développer la P-FLC pour le test pré-terrain, pour des tests psychométriques préliminaires et/ou complètes.

## **6. Étape 6 : Test psychométrique préliminaire de la version pré-finale de l'instrument traduit avec un échantillon**

Cette étape est rarement utilisée et, de toute façon, n'est pas indispensable [65]; cependant, quand une population bilingue est accessible, il est recommandé que l'instrument soit pré-testé sur le terrain entre des individus bilingues. En règle générale, la recommandation est d'utiliser au moins cinq sujets par article de l'instrument pour effectuer le test psychométrique préliminaire [88]. Idéalement, l'échantillon bilingue devrait être de la population cible dans laquelle l'instrument sera utilisé. Cependant, dans de nombreux cas, cela peut être difficile et irréaliste ; donc d'autres solutions peuvent être utilisées telles que l'échantillonnage sur des étudiants bilingues et des professeurs ou des travailleurs dans les agences de voyage, les agences de change, les sociétés commerciales internationales, les ambassades et consulats, et écoles de langues.

Initialement, les participants reçoivent la P-FLC et sont invités à répondre aux items sans voir l'instrument original. Puis, on leur donne l'instrument original et on les invite à répondre de nouveaux aux items. Ils peuvent remplir un questionnaire démographique et/ou d'autres instruments d'intérêt. L'ordre des items de l'instrument original doit être mélangé pour être dans un ordre différent de celui des éléments de la P-FLC [65]. Les réponses des deux versions de l'instrument sont alors comparées pour établir les critères d'équivalence. Les analyses statistiques utilisées à des fins de comparaison peuvent comporter des statistiques descriptives, les coefficients de corrélation, le test t apparié ou l'analyse de variance unidirectionnelle (ANOVA).

Lorsque l'objectif de l'instrument est de servir de test de diagnostic ou de dépistage, le calcul préliminaire de la sensibilité et de la spécificité est recommandé.

## 7. Étape 7 : Test psychométrique complet de la version pré-finale de l'instrument traduit dans un échantillon de la population

Cette dernière étape vise à établir les propriétés psychométriques complètes initiales de l'instrument nouvellement traduit, adapté et validé sur un échantillon de la population cible d'intérêt. La taille de l'échantillon pour cette étape dépend des types d'approches psychométriques qui seront utilisés. Plus les approches psychométriques pour l'évaluation de l'instrument traduit sont complètes, plus la confiance dans ses propriétés de fiabilité et de validité sera plus grande. En général, il est fortement recommandé d'utiliser au moins 10 sujets par item de l'échelle de l'instrument, de l'analyse des items et de l'analyse factorielle exploratoire [88-91]. S'il existe un projet d'utiliser l'analyse factorielle confirmatoire pour tester la structure factorielle de l'instrument, la recommandation générale est d'environ 300-500 sujets par item [91,92]. L'analyse de puissance en fonction du nombre de degrés de liberté, un niveau alpha (*p-value*) (0,05 ou 0,01) et une puissance désirée (80% ou plus) peuvent aussi être calculées [93,94].

Les approches psychométriques les plus recommandées et couramment utilisées dans cette étape sont l'estimation de :

1. La fiabilité de la cohérence interne (ou sensibilité et spécificité) ;
2. La fiabilité test-retest (fiabilité de stabilité) ;
3. L'homogénéité ;
4. La validité liée au concept, comme la validité convergente et/ou divergente (discriminante) ;
5. La validité liée aux critères tels que la validité concurrente et/ou prédictive ;
6. La structure du facteur de l'instrument (dimensionnalité) ;
7. La pertinence du modèle.

Le but de cette étape finale est de réviser et d'affiner les éléments de la P-FLC afin d'obtenir une version finale pertinente valide, fiable et homogène.

Les étapes des recommandations proposées ici fournissent des lignes directives claires à la traduction, l'adaptation et la validation transculturelle des échelles et des instruments de recherche.

Choisir les bons traducteurs et les membres du comité est un aspect clé qui doit être considéré pour améliorer la qualité de la traduction.

L'essai pilote parmi des participants dont la langue maternelle est la LC pour évaluer les instructions, les items et le format pour la clarté de réponse de l'instrument améliore également la qualité de la version finale.

Enfin, valider la traduction proposée nécessite l'utilisation d'approches bien établies pour tester les propriétés psychométriques préliminaires parmi des individus bilingues et des tests psychométriques complets à un échantillon de la population cible.

## **Chapitre 4 : Traduction et Validation de l'ADI-R en arabe dialectale**

## I. Traduction de l'ADI-R en arabe dialectale

Traduire un questionnaire en arabe dialectale est une tâche difficile à plusieurs niveaux :

1. D'abord le dialecte marocain n'est pas une langue formelle et chaque région dispose de ses propres nuances lexicales et sémantiques. Dans le projet actuel de traduction, on essaiera de proposer comme expressions ce qui est généralement compréhensible par la population marocaine sans entrer dans les spécificités propres à chaque région. L'arabe dialectale n'est pas une langue écrite, il serait difficile d'écrire les expressions en texte arabe vu que cette pratique est très informelle.

2. Le dialectal n'étant pas aussi riche du point de vue lexical et sémantique que la langue formelle, des empreints de la langue arabe classique peuvent donc s'imposer.

Le processus de traduction de l'ADI-R doit passer par plusieurs étapes :

D'abord, le projet doit être évalué et autorisé par le *Western Psychological Services (WPS)*, l'éditeur qui détient les droits d'auteur de l'ADI-R. Ensuite, l'achat du kit pour la traduction de l'instrument, le paiement de la redevance et la formation aux intervieweurs doivent être arrangés. Il est important de noter que l'ADI-R est une échelle de diagnostic appartenant au WPS et que son utilisation nécessite une formation préalable dans un centre reconnu par le titulaire du droit d'auteur.

Une fois les droits à la traduction acquis, Un traducteur qualifié doit assurer la traduction de la version originale anglaise au dialecte marocain. Si les moyens le permettent, la traduction initiale doit se faire séparément par deux traducteurs et la version finale de la version traduite du questionnaire doit être décidée en consultant un juge compétent indépendant.

Bien qu'il est aussi important de traduire les explications complémentaires, les notes et les commentaires du questionnaire, dans le cas de l'ADI-R, ce choix peut demander beaucoup de temps et de ressources vu la quantité du texte à traduire, l'ADI-R

comportant approximativement 80 pages de texte au format standard [19]. Une approche minimale pourrait donc consister à traduire seulement l'essentiel du questionnaire.

À cause du contenu très spécialisé du questionnaire, la version initiale traduite de l'ADI-R va nécessiter sûrement des interventions et des modifications par les auteurs du projet. Après, le texte doit être soumis pour correction du contenu.

L'étape prochaine est la retraduction à l'anglais « retraduction aveugle » : Un service spécialisé de traduction et qualifié se chargera de cette tâche.

Un comité incluant les personnes qui se sont chargé de la traduction, de la retraduction, les auteurs du projet et si possible un des auteurs de l'ADI-R sera chargé d'évaluer et de comparer la version originale du questionnaire à la version retraduite.

Cette procédure permet d'obtenir un meilleur agrément entre les deux versions originale et marocaine du questionnaire.

Nous gardons à l'esprit qu'il existe des expressions spécifiques à chaque langue et qui sont difficilement traduisibles. Ainsi par exemple l'expression « *fingers walking towards someone* » trouvera difficilement son équivalent marocain [19]!

Les réalisateurs du projet de traduction doivent aussi respecter les recommandations et guidelines des auteurs de la version originale de l'ADI-R. Ainsi, l'apparence, le style et la structure du contenu doivent être respectés. Ceci concerne les graphiques, les tableaux, le nombre de questions et catégories et leur ordre, les instructions et les exemples et les résultats des algorithmes de calcul. Chaque graphique de la version marocaine de l'ADI-R doit correspondre à un graphique de la version originale. Une page standard comprend le nom du département, la fonction de l'intervieweur ou clinicien, les questions, les catégories, les scores et les champs de l'évaluation. Les couleurs de tous les champs, la taille et la police (serif ou sans-serif), et les formats de la police (gras, italique) doivent être préservés.

La même règle d'équivalence s'applique au contenu des phrases : le syntaxe, grammaire, conjugaison, typographie, ponctuation, Les position et catégories des réponses dans leur ordre respectif ainsi que la méthode de calcul des scores doivent être préservés.

Notant que l'arabe marocain étant une langue moins technique et moins formelle que l'anglais, les phrases seront certainement plus langue que dans la version originale [19].

## II. Adaptation de l'ADI-R au contexte socio-culturel marocain

La traduction d'un questionnaire d'une langue à une autre ne se limite pas seulement au contenant de l'information, aux mots et phrase. Le but d'un questionnaire est d'établir un score qui reflète un certain nombre de caractères cognitifs ou émotionnels de l'individu en question. Ces paramètres sont intimement liés au contexte socio-culturel de la personne. Ainsi la réponse à la question '*c'est quoi la capitale du suède*' n'a pas la même valeur psychométrique pour un européen que pour un sud-américain par exemple !

En ce qui concerne l'autisme, L'accent mis sur l'identification des facteurs culturels influençant la compréhension et la vision des gens quant à ce trouble reste limité. Même si plusieurs études insistent sur les différences culturelles majeurs quant aux attitudes des sujets vis-à-vis de l'autisme voire même vis-à-vis de la déclaration des symptômes [11].

La nécessité de prendre en compte l'impact des différences individuelles et des valeurs culturelles attachées à ces différences peut être illustrée par les taux de prévalence des TSA variables à travers le monde (Tableau 1) [95].

Les variables culturelles et linguistiques peuvent contribuer à des difficultés dans l'identification des enfants atteints de TSA et de ce fait renforcent une certaine disparité dans le diagnostic de l'autisme chez certains groupes raciaux/ethniques [96,97]. Par exemple, Begeer et al. [97] ont constaté en 2009 que les pédiatres néerlandais seraient enclins à attribuer les problèmes sociaux et de la communication des groupes minoritaires non européens à leur origine ethnique, tout en attribuant ces mêmes caractéristiques à des troubles autistiques chez les enfants des groupes majoritaires.

Bien que les caractéristiques essentielles des TSA sont communes à toutes les cultures, les réponses des parents aux symptômes ne le sont pas; ces caractéristiques peuvent être vues à travers une lentille culturelle menant à un diagnostic erroné [98]. Les signes et les symptômes qui sont clairement démonstratifs dans les soins de santé des Etats-Unis ou d'un système d'éducation ne peuvent pas être considérés de la même manière pour quelqu'un d'une culture qui ne peut pas définir le trouble.

Un des facteurs contribuant à l'erreur de la classification et du diagnostic des sujets atteints d'autisme est l'interprétation culturelle et linguistique des familles ainsi que leur réaction au diagnostic [99]. Certaines cultures considèrent le handicap dans une lumière négative et estiment que c'est quelque chose qui doit être cachée des autres, ce qui évidemment a une influence sur le type de soins que la famille cherche.

Tous ces éléments doivent être pris en compte au cours et après la traduction du questionnaire de l'ADI-R pour éviter toute source de confusion ou biais socio-culturel potentiel.

Après ces deux étapes majeures de traduction et d'adaptation de l'ADI-R au contexte socio-culturel marocain. Le questionnaire peut être utilisé à des fins de recherche. Il ne peut être publié et utilisé dans un but clinique qu'après avoir fait l'objet d'un processus de validation [19] .

**Tableau 1** : Taux de prévalence des TSA à travers le monde.

<b>Pays</b>	<b>Prévalence (par 10,000)</b>	<b>Référence</b>
<b>Brésil</b>	27.2	Paula, Ribeiro, Fombonne, and Mercadante (2011)
<b>Canada</b>	60	Fombonne, Zakarian, Bennett, Meng, and McLean-Heywood
<b>Chine</b>	0.32-17.9	Sun and Allison (2010)
<b>Danemark</b>	7.2	Madsen et al. (2002)
<b>Danemark</b>	68.5	Parner et al. (2011)
<b>France</b>	16.3	Fombonne, Du Mazaubrun, Cans, and Grandjean (1997)
<b>Allemagne</b>	1.9	Steinhausen, Gobel, Breinlinger, and Wohlleben (1986)
<b>Island</b>	13.2	Magnu´sson and Saemundsen (2001)
<b>Indonésie</b>	11.7	Sun and Allison (2010)
<b>Israël</b>	10	Davidovich, Holtzman, and Tirosh (2001)
<b>Japon</b>	1.1-21.1	Sun and Allison (2010)
<b>japon</b>	181	Kawamura, Takahashi, and Ishii (2008)
<b>Oman</b>	1.4	Al-Farsi et al. (2010)
<b>Corée du sud</b>	260	Kim et al. (2011)
<b>Suède</b>	32.9	Gillberg, Cederlund, Lamberg, and Zeijlon (2006)
<b>Angleterre</b>	116	Baird et al. (2006)
<b>États-Unis</b>	42-121	Autism and Developmental Disabilities Monitoring Network
<b>Australie</b>	51	Parner et al. (2011)

### III. Projet de validation de la version marocaine de l'ADI-R

L'ADI-R est beaucoup utilisé à travers le monde, sa fiabilité et validité ont été examinées et prouvées dans sa version originale et dans ses versions traduites [6,14,21,100,101]. La tâche de validation du questionnaire ADI-R peut parfois prendre plusieurs années. Citons ici l'exemple de la validation de la version japonaise qui a nécessité 7 ans (de 2005 à 2012) [17]. Souvent, les auteurs des versions traduites des tests et questionnaires n'envisagent même pas la validation du questionnaire vu que cette procédure est très exigeante en matière d'effort, de temps et de savoir-faire.

Dans cette préparation du projet de validation de la future traduction marocaine du questionnaire ADI-R ; Les auteurs sont menés à tester la fiabilité inter-évaluateurs et discriminante et la validité du diagnostic de l'ADI-R, version Marocaine (ADI-R-VM).

Nous proposons dans ce projet de tester la fiabilité inter-évaluateurs en utilisant deux types de mesures : le Kappa pondéré ( $K_w$ ) et le coefficient de corrélation intra-classe (ICC) des scores des items des algorithmes diagnostiques de deux enquêteurs indépendants. La validité discriminante peut être évaluée en comparant les scores moyens des algorithmes diagnostiques (items/sous-domaines/domaines) entre les individus avec et sans diagnostic clinique consensuel.

Nous choisissons l'ICC au lieu du coefficient Cronbach's alpha, parce que c'est l'ICC qui a été utilisé dans les études initiales de validation de l'ADI-R [6,16,41], C'est la mesure utilisée dans les études qui ont développé le plus leur méthodologie de validation de la version traduite de l'ADI-R [17], et de toute façon, généralement les deux tests alpha et ICC donnent des résultats statistiquement équivalents dans la plupart des cas [86]. Parfois et pour être exhaustif, des auteurs préfèrent utiliser les deux mesures à la fois [6,41].

Nous proposons de définir la validité diagnostique dans cette étude comme l'accord entre le diagnostic basé sur l'ADI-R-VM et le diagnostic clinique consensuel. La sensibilité, la spécificité, la valeur prédictive positive et la valeur prédictive négative seront calculés pour évaluer cet accord.

Pour l'évaluation, on fera les hypothèses suivantes :

1. Une bonne à excellente fiabilité inter-évaluateurs en termes de  $K_w$  et ICC de l'ADI-R-VM serait observée, conformément à la littérature publiée [6,100,101] ;
2. La validité discriminante de l'ADI-R-VM serait suffisante, avec des algorithmes diagnostiques aux scores plus élevés parmi les sujets TSA par rapport aux non TSA [6,21] et des scores d'Autisme supérieurs aux scores des TED-NoS ;
3. La validité du diagnostic de l'ADI-R-VM serait satisfaisante quoique à degré moindre chez les jeunes individus et les personnes ayant une déficience intellectuelle [7,53,102].

## 1. Participants et procédure de diagnostic

### A. Étude de fiabilité

On propose les critères d'inclusion suivants :

- Les participants seront sélectionnés sur la base du nombre total de participants inscrits (minimum ciblé :  $N = 30$ ),
- L'âge (enfants de la maternelle et enfants d'âge scolaire/adolescents de moins de 20 ans),
- Le diagnostic clinique (diagnostic de trouble autistique confirmé ou suspecté selon les critères du DSM-IV),
- Consentement à participer à l'étude volontairement, y compris à l'enregistrement vidéo.

Ainsi on optera ici pour un échantillonnage par choix raisonné comme méthode d'échantillonnage dans cette étude.

En plus de ces participants TSA, d'autres sujets non autistes aux caractères démographiques similaires seront invités à participer à l'étude en tant que sujets témoins.

Juste après être recrutés par l'étude, tous les participants feront l'objet d'une évaluation clinique basée sur le DSM-IV-TR et menée par l'un des auteurs. Après ces évaluations cliniques détaillées, des entretiens supplémentaires seront menés afin de recueillir l'histoire du développement des participants. L'équipe de recherche posera alors des diagnostics cliniques consensuels basés sur le DSM-IV-TR [17]. Des évaluations de type cognitif seront également réalisées chez ces participants. Il n'existe pas encore une batterie de tests traduite et validée en arabe marocain pour mesurer le Quotient Intellectuel (QI). On propose alors de faire usage des tests de QI non verbaux, le test qui nous semble le plus intéressant dans cette optique est le Leiter vu qu'il est international, cela dit on pourrait aussi faire usage du WISC Non Verbal (WNV).

## **B. Étude de validité**

Pour cette étude, nous aurons besoin de plus de participants TSA (par exemple 200 sujets), le recrutement des participants peut donc s'étaler sur une longue période. On propose le même modèle d'échantillonnage raisonné [17].

Afin de capturer des différences potentielles entre les deux populations TSA des 2 études, nous proposons de comparer les caractéristiques démographiques de ces 2 échantillons par rapport au sexe, l'âge et les scores QI. Idéalement, il ne doit pas y avoir de différences significatives. Les deux échantillons peuvent alors être rassemblés tous les deux dans un groupe homogène de cas TSA pour l'étude de validité.

Un autre groupe d'enfants non autistes sera invité à participer. Ce groupe devrait avoir également les mêmes caractéristiques démographiques du groupe témoin de l'étude de fiabilité. Ainsi on regroupera ces 2 groupes une fois homogènes dans un échantillon témoin pour l'étude de validité.

Comme il a été fait pour l'étude de fiabilité, les participants de l'étude de validité feront également l'objet d'évaluation clinique basée sur les critères du DSM-IV-TR conduit par l'un des auteurs. Les diagnostics potentiels seront consensuels et selon les mêmes critères.

## 2. Entretiens avec l'ADI-R-VM

Tous les aidants des participants à cette étude seront interrogés à l'aide de l'ADI-R-VM après avoir pris part à l'étude. Les entrevues seront réalisées soit par l'un des auteurs du travail de traduction ou par d'autres professionnels qualifiés et entraînés pour la passation du questionnaire traduit.

## 3. Analyses

### A. Rappel des algorithmes diagnostiques de l'ADI-R-VM

Les algorithmes diagnostiques de l'ADI-R-VM comprennent 4 domaines suivants :

- Anomalies qualitatives des interactions sociales réciproques ;
- Anomalies qualitatives en communication ;
- Comportements répétitifs, stéréotypés et restreints ;
- Anomalies du développement évidents avant l'âge de 36 mois.

Les domaines A, B et C correspondent aux trois catégories de symptômes décrites dans le DSM-IV-TR. Le domaine A se compose de 4 sous-domaines couvrant 16 éléments de l'algorithme ; le domaine B est composé de 4 sous-domaines couvrant 13 éléments ; le domaine C est constitué de 4 sous-domaines couvrant 8 éléments ; et le domaine D n'a pas de sous-domaine et couvre 5 éléments de l'algorithme. Le score de chacun des éléments de l'algorithme prend seulement une des trois valeurs (0, 1 ou 2).

L'évaluation du domaine B est divisée en deux types d'évaluations selon les aptitudes verbales des personnes examinées :

- Les sous-domaines B1, B4, B2 (V) et B3 (V) couvrant 13 articles de l'algorithme vont être utilisés pour les personnes verbales ;
- Alors que seulement B1 et B4 vont être utilisés pour les personnes non-verbales.

Un diagnostic d'Autisme est décidé si tous les scores des quatre domaines (A, B, C et D) égalent ou dépassent les cutoffs suivants :

- 10 points pour le domaine A ;
- 8 points pour le domaine BV (domaine B pour les sujets verbaux) ou 7 points pour le domaine BNV (domaine B pour les sujets non verbaux) ;
- 3 points pour le domaine C ;
- 1 point pour le domaine D.

## **B. L'étude de fiabilité**

La valeur kappa pondéré ( $K_w$ ) doit être calculée pour chacun des 42 éléments de l'algorithme. Nous proposons le système de pondération quadratique de  $w_{ij} = 1 - (i - j)^2 / (k - 1)^2$ . Cela permet aux  $K_w$  et au Coefficient de Corrélation Intra-classe (ICC) d'être considérés équivalents l'un à l'autre [103]. Nous proposons également le calcul de l'ICC pour chacun des 4 domaines et 12 sous-domaines; Les scores additionnés des ICC des domaines et sous-domaines pourraient prendre un certain nombre de valeurs, et donc l'ICC est préférable au  $K_w$  dans ce sens [17].

En ce qui concerne le niveau de significativité, nous proposons les critères des études antérieures [100,104] ; à savoir :

- Les éléments montrant un  $K_w \geq 0,75$  et les domaines/sous-domaines montrant un ICC  $\geq 0,75$  sont considérés comme excellents ;
- $0,60 \leq K_w \leq 0,75$  et  $0,60 \leq ICC \leq 0,75$  sont considérés comme bons ;
- $0,40 \leq K_w \leq 0,60$  et  $0,40 \leq ICC \leq 0,60$  sont considérés comme limites ;
- Tandis que  $K_w \leq 0,40$  et  $ICC \leq 0,40$  manquent de fiabilité inter-évaluateurs.

Compte tenu de la différence dans la répartition par âge des trois groupes de diagnostic des participants (Autisme, TED-NoS et non TSA selon le consensus clinique), les analyses peuvent d'abord être effectuées pour tous les participants, puis une analyse ultérieure pourra être effectuée séparément pour trois tranches d'âge : moins de 5 ans; de 5 à 9 ans et plus de 10 ans [17].

### **C. L'étude de validité – validité discriminante**

Les scores moyens pour les 42 éléments de l'algorithme, les 12 sous-domaines, et les 3 domaines (A, B et C) chez les trois groupes de diagnostic des participants seront comparés à l'aide de la méthode d'analyse ANOVA après la correction de Bonferroni [17]. Les différences dans les scores moyens des items, sous-domaines et domaines devraient également être examinées chez les enfants de moins de 5 ans et/ou ceux avec un QI inférieur à 70 pour voir si à cet âge elles sont moins significatives.

### **D. Étude de validité – Validité diagnostique**

Pour voir si le diagnostic fourni sur la base de l'ADI-R-VM est valide, nous proposons d'estimer la sensibilité, la spécificité, la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN) de l'ADI-R-VM. Dans ce sens, la sensibilité fait référence à la proportion d'individus ayant un diagnostic de TSA sur base d'algorithme ADI-R-VM parmi ceux avec un diagnostic TSA clinique et consensuel. La spécificité est la proportion de ceux qui sont jugés non TSA sur la base de l'ADI-R-VM parmi ceux avec un consensus de diagnostic clinique de non TSA. La VPP est la proportion de sujets ayant un diagnostic clinique de consensus de TSA chez les personnes ayant un diagnostic basé sur un algorithme ADI-R-VM de l'autisme, et la VPN est la proportion de sujets ayant un diagnostic clinique consensuel de non-TSA chez les personnes ayant un diagnostic non-TSA basé sur un algorithme ADI-R-VM.

Selon des critères précédemment rapportés [105], nous proposons de juger la signification clinique de la sensibilité, la spécificité et les valeurs VPP et VPN d'être " limite " si les valeurs de ces mesures sont égales ou dépassent 70%, " bonne " si elles dépassent 80% , et " excellente " si elles sont supérieures à 90%. Nous proposons également d'examiner si les résultats de ceux-ci seraient inférieurs si l'analyse seraient inférieurs chez les individus très jeunes (moins de 5 ans) ou les personnes ayant une déficience intellectuelle ( $QI < 70$ ).

#### 4. Considérations éthiques

Le protocole d'étude doit répondre aux normes des guidelines éthiques de la Déclaration de Helsinki. Tous les participants, ainsi que leurs aidants naturels, doivent recevoir une description complète de l'étude, et les aidants sont invités à donner leur consentement éclairé par écrit de participer à l'étude. En ce qui concerne les individus référés pour suspicion d'autisme, ils doivent être évalués et examinés en combinant les observations cliniques aux évaluations. Puis, au moment de l'entretien ADI-R-VM, un diagnostic de consensus clinique sera formulé par les experts de l'équipe de recherche. Après que l'entretien ADI-R-VM avec les aidants naturels sera mené, l'équipe devra formuler la meilleure estimation diagnostique sur la base à la fois du diagnostic clinique de consensus et du diagnostic de l'algorithme ADI-R et devra communiquer celle-ci aux aidants du sujet.

## Conclusion

Traduire des outils d'évaluation psychométrique est un vrai travail en soit. Parfois même dépassant les autres travaux de recherche qui se basent sur ces outils en terme de méthodologie, de rigueur scientifique, de temps et d'efforts fournis. Mais les fruits de ce genre de travaux sont indéniables. La traduction de ces outils avec leur adaptation et leur validation dans un contexte socioculturel donné ouvre la voie de la recherche et la pratique clinique sur des nouveaux horizons.

Dans ce travail :

- Nous avons essayé de présenter l'ADI-R dans ces moindres aspects et considérations aussi bien techniques que pratiques ;
- Nous avons fait le tour de la question de la validité internationale et transculturelle de l'ADI-R en tant qu'outil diagnostique de l'autisme ;
- Nous avons essayé de présenter de façon méthodique les différentes étapes du processus de la traduction, l'adaptation et la validation d'un instrument psychométrique quelconque ;
- Et à la fin, nous avons essayé de présenter un projet complet et hiérarchisé des différentes étapes de la traduction en langue marocaine de l'outil ADI-R, son adaptation et sa validation, tout en se limitant à l'essentiel et au réalisable.

Nous souhaitant que les étapes de traduction de l'ADI-R seront concrétisées le plus tôt possible afin d'offrir à la communauté scientifique marocaine un outil aussi indispensable et essentiel dans la pratique clinique de l'autisme.

## Références

1. Ameisen J-C, Deschamps C, Kordon C, Haïm K, Lebatard C, Rouvillois P. Sur la situation en France des personnes, enfants et adultes, atteintes d'autisme [Internet]. Comité Consultatif National d'Ethique pour les Sciences de la Vie et de la Santé; 2007 p. 30. Report No.: 102. Available from: <http://www.ccne-ethique.fr/fr/publications/sur-la-situation-en-france-des-personnes-enfants-et-adultes-atteintes-dautisme#.V7n2uDUTz6g>
2. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D, et al. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet Lond Engl*. 2006 Jul 15;368(9531):210-5.
3. Fombonne E. Epidemiology of pervasive developmental disorders. *Pediatr Res*. 2009 Jun;65(6):591-8.
4. Gillberg C, Cederlund M, Lamberg K, Zeijlon L. Brief report: "the autism epidemic". The registered prevalence of autism in a Swedish urban area. *J Autism Dev Disord*. 2006 Apr;36(3):429-35.
5. Aalouane R. Troubles Du Spectre Autistiques. *Revue de médecine pratique*. 2016 Février;53.
6. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 1994 Oct;24(5):659-85.
7. Le Couteur A, Lord C, Rutter M. Autism diagnostic interview-revised. *Western Psychological Services*; 2003.
8. Gray KM, Tonge BJ, Sweeney DJ. Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: evaluating diagnostic validity. *J Autism Dev Disord*. 2008 Apr;38(4):657-67.
9. Moss J, Magiati I, Charman T, Howlin P. Stability of the Autism Diagnostic Interview-Revised from Pre-School to Elementary School Age in Children with Autism Spectrum Disorders. *J Autism Dev Disord*. 2007 Dec 6;38(6):1081-91.
10. Ozonoff S, Williams BJ, Landa R. Parental report of the early development of children with regressive autism: the delays-plus-regression phenotype. *Autism Int J Res Pract*. 2005 Dec;9(5):461-86.
11. Daley L, Keys C, Henry D, Kim YS, Leventhal B. The CROSS Cultural Examination of a Brief Autism Diagnostic Interview (ADI-R) in KOREA and the United States. In: *ResearchGate* [Internet]. 2014 [cited 2016 Aug 17]. Available from:

[https://www.researchgate.net/publication/268131200\\_The\\_CROSS\\_Cultural\\_Examination\\_of\\_a\\_Brief\\_Autism\\_Diagnostic\\_Interview\\_ADI-R\\_in\\_KOREA\\_and\\_the\\_United\\_States](https://www.researchgate.net/publication/268131200_The_CROSS_Cultural_Examination_of_a_Brief_Autism_Diagnostic_Interview_ADI-R_in_KOREA_and_the_United_States)

12. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. Combining Information From Multiple Sources in the Diagnosis of Autism Spectrum Disorders. *J Am Acad Child Adolesc Psychiatry*. 2006 Sep 1;45(9):1094–103.
13. Spence SJ, Sharifi P, Wiznitzer M. Autism spectrum disorder: screening, diagnosis, and medical evaluation. *Semin Pediatr Neurol*. 2004 Sep;11(3):186–95.
14. Mildenberger K, Sitter S, Noterdaeme M, Amorosa H. The use of the ADI–R as a diagnostic tool in the differential diagnosis of children with infantile autism and children with a receptive language disorder. *Eur Child Adolesc Psychiatry*. 2001 Dec;10(4):248–55.
15. Soke GN, Philofsky A, Diguseppi C, Lezotte D, Rogers S, Hepburn S. Longitudinal changes in Scores on the Autism Diagnostic Interview–Revised (ADI–R) in pre–school children with autism: Implications for diagnostic classification and symptom stability. *Autism Int J Res Pract*. 2011 Sep;15(5):545.
16. Le Couteur A, Rutter M, Lord C, Rios P, Robertson S, Holdgrafer M, et al. Autism diagnostic interview: a standardized investigator–based instrument. *J Autism Dev Disord*. 1989 Sep;19(3):363–87.
17. Tsuchiya KJ, Matsumoto K, Yagi A, Inada N, Kuroda M, Inokuchi E, et al. Reliability and validity of autism diagnostic interview–revised, Japanese version. *J Autism Dev Disord*. 2013 Mar;43(3):643–62.
18. Becker MM, Wagner MB, Bosa CA, Schmidt C, Longo D, Papaleo C, et al. Translation and validation of Autism Diagnostic Interview–Revised (ADI–R) for autism diagnosis in Brazil. *Arq Neuropsiquiatr*. 2012 Mar;70(3):185–90.
19. Chojnicka I, Płoski R. [Polish version of the ADI–R (Autism Diagnostic Interview–Revised)]. *Psychiatr Pol*. 2012 Apr;46(2):249–59.
20. Sun X, Allison C, Auyeung B, Zhang Z, Matthews FE, Baron–Cohen S, et al. Validation of existing diagnosis of autism in mainland China using standardised diagnostic instruments. *Autism Int J Res Pract*. 2015 Nov;19(8):1010.
21. Lampi KM, Sourander A, Gissler M, Niemelä S, Rehnström K, Pulkkinen E, et al. Brief report: validity of Finnish registry–based diagnoses of autism with the ADI–R. *Acta Paediatr Oslo Nor 1992*. 2010 Sep;99(9):1425–8.

22. Bölte S. The Swedish Versions of ADOS and ADI-R: Psychometric Validation [Internet]. 2015 [cited 2016 Aug 14]. Available from: <http://ki.se/en/kind/reliability-and-validity-of-the-swedish-versions-of-the-ados-and-the-adi-r>
23. Aldosari M. Autism Program – Progress as on 31 December 2010 [Internet]. King Salman Center for Disability Research; 2011. Available from: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwj2167lwsHOAhWPzRoKHZW2D08QFggeMAA&url=http%3A%2F%2Fwww.kscdr.org.sa%2Fen%2Fresearch%2Fprojects%2FDocuments%2FAutism%2520Program%2520Progress%2520update.pdf&usg=AFQjCNHeKWK8iruYUgELFfwQT5bA1yRD\\_Q&sig2=3lZc5mpclVvsO7v6J6U4zw](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwj2167lwsHOAhWPzRoKHZW2D08QFggeMAA&url=http%3A%2F%2Fwww.kscdr.org.sa%2Fen%2Fresearch%2Fprojects%2FDocuments%2FAutism%2520Program%2520Progress%2520update.pdf&usg=AFQjCNHeKWK8iruYUgELFfwQT5bA1yRD_Q&sig2=3lZc5mpclVvsO7v6J6U4zw)
24. Minnesota. Minnesota Statutes Annotated. West Publishing Company; 1946. 686 p.
25. Evans B. How autism became autism. *Hist Hum Sci.* 2013 Jul;26(3):3–31.
26. Hus V, Lord C. Effects of child characteristics on the Autism Diagnostic Interview–Revised: Implications for use of scores as a measure of ASD severity. *J Autism Dev Disord.* 2013 Feb;43(2):371.
27. World Health Organization. The ICD–10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines [Internet]. 1992 [cited 2016 Aug 7]. Available from: [http://www.who.int/substance\\_abuse/terminology/icd\\_10/en/](http://www.who.int/substance_abuse/terminology/icd_10/en/)
28. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (4th ed.) [Internet]. 1994 [cited 2016 Aug 7]. Available from: <https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/DSM4/>
29. Kim SH (Sophy), Hus V, Lord DC. Autism Diagnostic Interview–Revised. In: Volkmar FR, editor. *Encyclopedia of Autism Spectrum Disorders* [Internet]. Springer New York; 2013 [cited 2016 Sep 18]. p. 345–9. Available from: [http://link.springer.com/referenceworkentry/10.1007/978-1-4419-1698-3\\_894](http://link.springer.com/referenceworkentry/10.1007/978-1-4419-1698-3_894)
30. Dawson G, Webb SJ, Carver L, Panagiotides H, McPartland J. Young children with autism show atypical brain responses to fearful versus neutral facial expressions of emotion. *Dev Sci.* 2004 Jun;7(3):340–59.
31. Lainhart JE, Bigler ED, Bocian M, Coon H, Dinh E, Dawson G, et al. Head circumference and height in autism: a study by the Collaborative Program of Excellence in Autism. *Am J Med Genet A.* 2006 Nov 1;140(21):2257–74.
32. Jones RM, Risi S, Wexler D, Anderson D, Corsello C, Pickles A, et al. How interview questions are placed in time influences caregiver description of social communication symptoms on the ADI–R. *J Child Psychol Psychiatry.* 2015 May;56(5):577–85.

33. Charman T, Gotham K. Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders – lessons from research and practice. *Child Adolesc Ment Health*. 2013 Feb 1;18(1):52.
34. Kim SH, Thurm A, Shumway S, Lord C. Multisite Study of New Autism Diagnostic Interview–Revised (ADI–R) Algorithms for Toddlers and Young Preschoolers. *J Autism Dev Disord*. 2013 Jul;43(7):1527.
35. Kim SH, Lord C. New autism diagnostic interview–revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *J Autism Dev Disord*. 2012 Jan;42(1):82–93.
36. Zander E. Standardizing diagnostic assessment of autism spectrum disorder: assets and challenges. [Stockholm]: Karolinska Institutet; 2015.
37. Snow AV, Lecavalier L, Houts C. The structure of the Autism Diagnostic Interview–Revised: diagnostic and phenotypic implications. *J Child Psychol Psychiatry*. 2009 Jun;50(6):734–42.
38. Western Psychological Services. WPS titles in commercial translation [Internet]. 2015 [cited 2016 Aug 7]. Available from: <http://www.wpspublish.com/app/OtherServices/PublishedTranslations.aspx>
39. Bildt A de, Sytema S, Zander E, Bölte S, Sturm H, Yirmiya N, et al. Autism Diagnostic Interview–Revised (ADI–R) Algorithms for Toddlers and Young Preschoolers: Application in a Non–US Sample of 1,104 Children. *J Autism Dev Disord*. 2015;45(7):2076.
40. Volkmar FR, Paul R, Klin A, Cohen DJ, editors. *Handbook of Autism and Pervasive Developmental Disorders, Two Volume Set*. 3 edition. Hoboken, NJ: Wiley; 2005. 1488 p.
41. Lord C, Storoschuk S, Rutter M, Pickles A. Using the ADI–R to diagnose autism in preschool children. *Infant Ment Health J*. 1993 Sep 1;14(3):234–52.
42. Ventola PE, Kleinman J, Pandey J, Barton M, Allen S, Green J, et al. Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *J Autism Dev Disord*. 2006 Oct;36(7):839–47.
43. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *J Autism Dev Disord*. 2000;30(3):205–23.

44. Schopler E, Reichler RJ, Renner BR. The Childhood Autism Rating Scale (CARS). Western Psychological Services; 1980.
45. Wiggins LD, Robins DL. Brief Report: Excluding the ADI-R Behavioral Domain Improves Diagnostic Agreement in Toddlers. *J Autism Dev Disord.* 2007 Sep 19;38(5):972-6.
46. Le Couteur A, Haden G, Hammal D, McConachie H. Diagnosing autism spectrum disorders in pre-school children using two standardised assessment instruments: the ADI-R and the ADOS. *J Autism Dev Disord.* 2008 Feb;38(2):362-72.
47. Gotham K, Risi S, Pickles A, Lord C. The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *J Autism Dev Disord.* 2007 Apr;37(4):613-27.
48. Ward-King J, Cohen IL, Penning H, Holden JJA. Brief report: telephone administration of the autism diagnostic interview--revised: reliability and suitability for use in research. *J Autism Dev Disord.* 2010 Oct;40(10):1285-90.
49. Vrancic D, Nanclares V, Soares D, Kulesz A, Mordzinski C, Plebst C, et al. Sensitivity and specificity of the autism diagnostic inventory-telephone screening in Spanish. *J Autism Dev Disord.* 2002 Aug;32(4):313-20.
50. Vanegas SB, Magaña S, Morales M, McNamara E. Clinical Validity of the ADI-R in a US-Based Latino Population. *J Autism Dev Disord.* 2016 May;46(5):1623-35.
51. PubMed entry [Internet]. [cited 2016 Aug 7]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16926617,%2016926617>
52. Chawarska K, Klin A, Paul R, Volkmar F. Autism spectrum disorder in the second year: stability and change in syndrome expression. *J Child Psychol Psychiatry.* 2007 Feb 1;48(2):128-38.
53. Cox A, Klein K, Charman T, Baird G, Baron-Cohen S, Swettenham J, et al. Autism spectrum disorders at 20 and 42 months of age: stability of clinical and ADI-R diagnosis. *J Child Psychol Psychiatry.* 1999 Jul;40(5):719-32.
54. Falkmer T, Anderson K, Falkmer M, Horlin C. Diagnostic procedures in autism spectrum disorders: a systematic literature review. *Eur Child Adolesc Psychiatry.* 2013 Jan 16;22(6):329-40.
55. Lord C, Pickles A, McLennan J, Rutter M, Bregman J, Folstein S, et al. Diagnosing autism: analyses of data from the Autism Diagnostic Interview. *J Autism Dev Disord.* 1997 Oct;27(5):501-17.

56. De Bildt A, Sytema S, Kraijer D, Minderaa R. Prevalence of pervasive developmental disorders in children and adolescents with mental retardation. *J Child Psychol Psychiatry*. 2005 Mar 1;46(3):275–86.
57. Sappok T, Diefenbacher A, Budczies J, Schade C, Grubich C, Bergmann T, et al. Diagnosing autism in a clinical sample of adults with intellectual disabilities: how useful are the ADOS and the ADI–R? *Res Dev Disabil*. 2013 May;34(5):1642–55.
58. Magaña S, Smith LE. The Use of the Autism Diagnostic Interview–Revised with a Latino Population of Adolescents and Adults with Autism. *J Autism Dev Disord*. 2013 May;43(5):1098.
59. Papanikolaou K, Paliokosta E, Houliaras G, Vgenopoulou S, Giouroukou E, Pehlivanidis A, et al. Using the Autism Diagnostic Interview–Revised and the Autism Diagnostic Observation Schedule–Generic for the diagnosis of autism spectrum disorders in a Greek sample with a wide range of intellectual abilities. *J Autism Dev Disord*. 2009 Mar;39(3):414–20.
60. Albores–Gallo L, Roldán–Ceballos O, Villarreal–Valdes G, Betanzos–Cruz BX, Santos–Sánchez C, Martínez–Jaime MM, et al. M–CHAT Mexican Version Validity and Reliability and Some Cultural Considerations. *ISRN Neurol [Internet]*. 2012 Jul 1 [cited 2016 Aug 8];2012. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3395208/>
61. Grinker RR, Kang–Yi CD, Ahmann C, Beidas RS, Lagman A, Mandell DS. Cultural Adaptation and Translation of Outreach Materials on Autism Spectrum Disorder. *J Autism Dev Disord*. 2015 Aug;45(8):2329–36.
62. García SB, Pérez AM, Ortiz AA. Mexican American Mothers’ Beliefs About Disabilities Implications for Early Childhood Intervention. *Remedial Spec Educ*. 2000 Mar 1;21(2):90–120.
63. Hilton A, Skrutkowski M. Translating instruments into other languages: development and testing processes. *Cancer Nurs*. 2002 Feb;25(1):1–7.
64. Sperber AD. Translation and validation of study instruments for cross–cultural research. *Gastroenterology*. 2004 Jan;126(1 Suppl 1):S124–128.
65. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross–cultural health care research: a clear and user–friendly guideline. *J Eval Clin Pract*. 2011 Apr;17(2):268–74.
66. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross–cultural adaptation of self–report measures. *Spine*. 2000 Dec 15;25(24):3186–91.

67. Sousa VD, Zauszniewski JA, Mendes IAC, Zanetti ML. Cross-cultural equivalence and psychometric properties of the Portuguese version of the Depressive Cognition Scale. *J Nurs Meas.* 2005;13(2):87-99.
68. Bracken BA, Barona A. State of the Art Procedures for Translating, Validating and Using Psychoeducational Tests in Cross-Cultural Assessment. *Sch Psychol Int.* 1991 Apr 1;12(1-2):119-32.
69. Brislin RW. Back-Translation for Cross-Cultural Research. *J Cross-Cult Psychol.* 1970 Sep 1;1(3):185-216.
70. Brislin RW. The wording and translation of research instruments. In: Lonner WJ, Berry JW, editors. *Field methods in cross-cultural research.* Thousand Oaks, CA, US: Sage Publications, Inc; 1986. p. 137-64. (Cross-cultural research and methodology series, Vol. 8.).
71. Chapman DW, Carter JF. Translation Procedures for the Cross Cultural Use of Measurement Instruments. *Educ Eval Policy Anal.* 1979 May 1;1(3):71-6.
72. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993 Dec;46(12):1417-32.
73. Jones E. Translation of quantitative measures for use in cross-cultural research. *Nurs Res.* 1987 Oct;36(5):324-7.
74. Jones EG, Kay M. Instrumentation in cross-cultural research. *Nurs Res.* 1992 Jun;41(3):186-8.
75. McDermott MAN, Palchanes K. A Literature Review of the Critical Elements in Translation Theory. *Image J Nurs Sch.* 1994 Jun 1;26(2):113-8.
76. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health J Int Soc Pharmacoeconomics Outcomes Res.* 2005 Apr;8(2):94-104.
77. Maneesriwongul W, Dixon JK. Instrument translation process: a methods review. *J Adv Nurs.* 2004 Oct;48(2):175-86.
78. Werner O, Campbell DT. Translating, working through interpreters, and the problem of decentering. In: *A handbook of method in cultural anthropology.* Naroll & R. Cohen. New York, NY: American Museum of Natural History; p. 398-420.

79. Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin's translation model for cross-cultural research. *Nurs Res.* 2001 Oct;50(5):300-4.
80. Beck CT, Bernal H, Froman RD. Methods to document semantic equivalence of a translated scale. *Res Nurs Health.* 2003 Feb;26(1):64-73.
81. Sousa VD, Hartman SW, Miller EH, Carroll MA. New measures of diabetes self-care agency, diabetes self-efficacy, and diabetes self-management for insulin-treated individuals with type 2 diabetes. *J Clin Nurs.* 2009 May;18(9):1305-12.
82. Topf M. Three estimates of interrater reliability for nominal data. *Nurs Res.* 1986 Aug;35(4):253-5.
83. Lynn MR. Determination and quantification of content validity. *Nurs Res.* 1986 Dec;35(6):382-5.
84. Feher Waltz C, Strickland OL, R. Lenz E. *Measurement in Nursing and Health Research.* 3rd ed. New York: Springer Publishing Company; 2005. 460 p.
85. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health.* 2006 Oct;29(5):489-97.
86. Wynd CA, Schmidt B, Schaefer MA. Two quantitative approaches for estimating content validity. *West J Nurs Res.* 2003 Aug;25(5):508-18.
87. Streiner DL, Norman GR. *Health Measurement Scales: A practical guide to their development and use.* 4 edition. Oxford ; New York: Oxford University Press; 2008. 450 p.
88. Nunnally JC, Bernstein IH. *Psychometric theory.* McGraw-Hill; 1994. 786 p.
89. Hair JF, Tatham RL, Anderson RE, Black W. *Multivariate Data Analysis.* 5th edition. Upper Saddle River, N.J: Prentice Hall; 1998. 768 p.
90. Stevens JP. *Applied Multivariate Statistics for the Social Sciences, Fourth Edition.* 4 edition. Mahwah, N.J: Psychology Press; 2001. 718 p.
91. Tabachnick BG, Fidell LS. *Using Multivariate Statistics.* 4 edition. Boston, MA: Allyn & Bacon; 2000. 966 p.
92. Gonzalez R, Griffin D. Testing parameters in structural equation modeling: every "one" matters. *Psychol Methods.* 2001 Sep;6(3):258-69.
93. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods.* 1996;1(2):130-49.

94. MacCallum RC, Browne MW, Cai L. Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychol Methods*. 2006 Mar;11(1):19–35.
95. Norbury CF, Sparks A. Difference or disorder? Cultural issues in understanding neurodevelopmental disorders. *Dev Psychol*. 2013;49(1):45–58.
96. Dyches TT. Assessing Diverse Students With Autism Spectrum Disorders. *ASHA Lead*. 2011 Jan 1;16(1):12–5.
97. Begeer S, Bouk SE, Boussaid W, Terwogt MM, Koot HM. Underdiagnosis and referral bias of autism in ethnic minorities. *J Autism Dev Disord*. 2009 Jan;39(1):142–8.
98. Tina Taylor Dyches, Lynn K. Wilder, Festus E. Obiakor. Autism: Multicultural perspectives. In: *Autistic Spectrum Disorders: Educational and Clinical Interventions* [Internet]. Emerald Group Publishing Limited; 2001 [cited 2016 Aug 25]. p. 151–77. (Advances in Special Education; vol. 14). Available from: <http://www.emeraldinsight.com/doi/abs/10.1016/S0270-4013%2801%2980012-X>
99. Wilder LK, Dyches TT, Obiakor FE, Algozzine B. Multicultural Perspectives on Teaching Students With Autism. *Focus Autism Dev Disabil*. 2004 May 1;19(2):105–13.
100. Cicchetti DV, Lord C, Koenig K, Klin A, Volkmar FR. Reliability of the ADI–R: multiple examiners evaluate a single case. *J Autism Dev Disord*. 2008 Apr;38(4):764–70.
101. Hill A, Bölte S, Petrova G, Beltcheva D, Tacheva S, Poustka F. Stability and interpersonal agreement of the interview–based diagnosis of autism. *Psychopathology*. 2001 Aug;34(4):187–91.
102. de Bildt A, Sytema S, Ketelaars C, Kraijer D, Mulder E, Volkmar F, et al. Interrelationship between Autism Diagnostic Observation Schedule–Generic (ADOS–G), Autism Diagnostic Interview–Revised (ADI–R), and the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV–TR) classification in children and adolescents with mental retardation. *J Autism Dev Disord*. 2004 Apr;34(2):129–37.
103. Fleiss JL, Cohen J. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educ Psychol Meas*. 1973 Oct 1;33(3):613–9.
104. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic*. 1981 Sep;86(2):127–37.
105. Cicchetti DV, Volkmar F, Klin A, Showalter D. Diagnosing Autism using ICD–10 criteria: A comparison of neural networks and standard multivariate procedures. *Child Neuropsychol*. 1995 Apr;1(1):26–37.