

Université de Montréal

**Synthèse systématique des écrits :
Les instruments d'évaluation de la comorbidité psychiatrique
présentée par les adolescent(e)s.**

par

Nathalie Marchiori

École de criminologie

Faculté des Arts et des Sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maîtrise en criminologie (M.Sc.)

Décembre 2015

© Nathalie Marchiori, 2015

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

**Synthèse systématique des écrits : Les instruments d'évaluation de la comorbidité
psychiatrique présentée par les adolescent(e)s.**

présenté par
Nathalie Marchiori

a été évalué par un jury composé des personnes suivantes :

Frédéric Ouellet
président-rapporteur

Denis Lafortune
directeur de recherche

Serge Brochu
membre du jury

Mémoire accepté le : 29 avril 2016

Résumé

Ce mémoire présente les résultats d'une synthèse systématique (SS) des écrits traitant des instruments d'évaluation multidimensionnelle des troubles concomitants qui peuvent être présentés par les adolescent(e)s. La SS a permis d'identifier 11 instruments en mesure d'évaluer les troubles comorbides de l'Axe I du DSM-IV, incluant chaque fois les troubles liés à l'utilisation de substances psychoactives (TUS). Une fois les instruments répertoriés, une seconde recherche fut effectuée afin d'identifier les études les ayant mis à l'épreuve du point de vue de leur validité et de leur fidélité diagnostique : 57 études furent identifiées. La robustesse méthodologique de ces études fut analysée à l'aide de la grille du QUADAS-2 et 47 études furent retenues pour l'échantillon final. Les résultats sont présentés par diagnostics (troubles liés à l'utilisation des substances (TUS) (obligatoire), trouble d'anxiété généralisée (TAG), épisode dépressif majeur (ÉDM), troubles des conduites (TC), trouble du déficit de l'attention /hyperactivité (TDA/H), état de stress post-traumatique (ÉSPT) et par instrument retenu. Suite à l'analyse des données recueillies, il s'avère difficile de comparer les instruments les uns aux autres, étant donnée la très grande diversité des échelles qu'ils contiennent, ainsi que les devis fort différents des études qui les ont mis à l'épreuve. Par contre, deux instruments se distinguent par la robustesse méthodologique des études à leur sujet, ainsi que leur excellente performance globale. Il s'agit du ChIPS et du K-SADS.

Mots-clés : synthèse systématique, évaluation multidimensionnelle, comorbidité, psychiatrique, adolescent, adolescente, validité, fidélité

Abstract

This dissertation outlines the results of a systematic review of literature regarding comprehensive assessment of comorbid disorders with adolescents. The SR identified 11 instruments capable of assessing comorbid disorders of DSM-IV's Axis I, including every time substance related disorders. Once those instruments identified, a second review took place to identify research that had put those instruments to the test regarding validity and reliability of diagnostics produced: 57 studies were identified. Their methodological robustness was evaluated with the QUADAS-2 and 47 studies composed the final sample. The results are presented by diagnostics (Substance related disorder (mandatory), Generalized anxiety disorder, Major depressive disorder, Conduct disorder, Attention-Deficit Hyperactivity Disorder, Posttraumatic stress disorder) and by instruments. After analyzing the data collected, it appears that it is difficult to compare the instruments to each other, given the great diversity of scales they contain, as well as the various researchers' design when put to the test regarding their validity and reliability. However, two instruments distinguish themselves by the methodological robustness of the studies who have analyzed them and their excellent global result. They are the CHIPS and the K-SADS.

Keywords: systematic review, comprehensive assessment, comorbidity, psychiatric, youth, validity, reliability

Table des matières

Résumé.....	iii
Abstract.....	iv
Table des matières.....	v
Liste des tableaux.....	ix
Remerciements.....	xiii
Introduction	1
Chapitre 1 : Recension des écrits.....	6
1. La perspective de la psychopathologie développementale	7
2. L'évolution de la terminologie relative à la santé mentale	7
3. Comorbidité	10
4. Consommation de substances psychoactives (SPA) : abus et/ou dépendance.	11
5. Troubles avec comorbidité chez les adolescents et adolescentes	18
5.1 Taux de prévalence chez les adolescents et adolescentes.....	19
5.1.1 Au plan international	19
5.1.2 Aux États-Unis	20
5.1.3 Au Canada	21
5.1.4 Au Québec.....	23
6. L'évaluation.....	27
6.1 Les instruments d'évaluation.....	30
6.1.1 Entrevue basée sur l'intervieweur ou le participant	31
6.1.2 Souple, semi-structurée, structurée	33
6.1.3 Qualité psychométrique : fidélité et validité	34
7. La synthèse systématique.....	35
8. Problématique	36
Chapitre 2 : Méthodologie.....	39
1. La synthèse systématique.....	40
2. Stratégie de recherche.....	40

2.1 L'identification des instruments d'évaluation	40
2.1.1 Identification des bases de données	40
2.1.2 Les critères d'inclusion et d'exclusion des instruments d'évaluation	42
2.2 L'identification des études.....	45
2.2.1 Identification des bases de données	45
2.2.2 Les critères d'inclusion et d'exclusion des articles de recherche.....	47
3. Comité accompagnateur.....	49
4. Analyse de la qualité méthodologique des études retenues	50
4.1 Critères du QADAS-2.....	50
4.2 Question de départ.....	50
4.3 Les domaines.....	51
Chapitre 3 : Résultats.....	54
1. Présentation des instruments retenus	55
1.1 Adolescent Diagnostic Interview-ADI	55
1.2 Adolescent Psychopathology Scale-APS	56
1.3 The Child and Adolescent Psychiatric Assessment-CAPA.....	56
1.4 Children's Interview for Psychiatric Syndromes-ChIPS.....	57
1.5 Diagnostic Interview for Children and Adolescents-DICA	57
1.6 Diagnostic Interview Schedule for Children-DISC.....	58
1.7 Global Appraisal of Individual Needs-GAIN	59
1.8 Interview Schedule for Children and Adolescents-ISCA	60
1.9 Schedule of Affective Disorders and Schizophrenia for Children-K-SADS.....	61
1.10 Practical Adolescent Dual Diagnostic Interview-PADDI.....	62
1.11 Personality Assessment Inventory-Adolescent-PAI-A	62
2. Présentation des études recensées	64
2.1 Résultats de l'évaluation de qualité méthodologique (QUADAS-2).....	65
2.2 Type de validité et fidélité répertoriées	69
2.2.1 Mesures de la fidélité	70
2.2.2 Mesures de la validité	73

3.	Analyse par type de fidélité et de validité	74
4.	Analyse par diagnostic	78
4.1	Troubles liés à l'utilisation des substances psychoactives (TUS).....	79
4.2	Trouble anxiété généralisé (TAG)	82
4.3	Épisode dépressif majeur (ÉDM)	86
4.4	Troubles des conduites (TC).....	90
4.5	Troubles du déficit de l'attention avec ou sans hyperactivité (TDA/H)	95
4.6	État de stress post-traumatique (ESPT)	100
5.	Analyse par instrument.....	103
5.1	ADI.....	104
5.2	APS	104
5.3	CAPA.....	104
5.4	ChIPS	105
5.5	DICA.....	105
5.6	DISC	107
5.7	GAIN	108
5.8	K-SADS.....	108
5.9	PADDI	109
	Conclusion.....	111
1.	Discussion des résultats	113
2.	Les limites des études	119
3.	Limites et forces de la synthèse systématique	122
4.	Conclusion.....	124
	Bibliographie	130
	Annexes.....	150
	Annexe A : Légende des Tableaux diagnostiques 1 à 6 (annexes B à G)	151
	Annexe B : Tableau 1, Trouble lié à l'utilisation de substance psychoactive (TUS).....	152
	Annexe C : Tableau 2, Trouble anxiété généralisé (TAG).....	160
	Annexe D : Tableau 3, Épisode dépression majeure (ÉDM)	168

Annexe E : Tableau 4, Trouble des conduites (TC)	180
Annexe F : Tableau 5, Trouble déficitaires de l'attention avec ou sans hyperactivité (TDA/H)	192
Annexe G : Tableau 6, État de stress post-traumatique (ESPT).....	205
Annexe H : Retour sommaire dans les bases pour les années 2013 à 2016 en utilisant le titre de l'instrument seulement	209

Liste des tableaux

Tableau I : Critères d'abus et de dépendance à une substance selon le DSM-IV-TR	14
Tableau II : Prévalence de TM aux États-Unis (13 à 17 ans)	20
Tableau III : Prévalence TUS et TM au Canada	22
Tableau IV : Prévalence TUS et TM au Canada	22
Tableau V : Prévalence TUS et TM au Québec chez les 15-24 ans	24
Tableau VI : Consommation régulière ou quotidienne d'alcool au Québec (écoles secondaires)	25
Tableau VII : Consommation de drogues illégales au Québec (écoles secondaires).....	25
Tableau VIII : Consommation régulière de drogue au Québec (écoles secondaires).....	26
Tableau IX : Instruments d'évaluation retenus.....	43
Tableau X : Occurrences des instruments dans OVID, Web of Science et Proquest.....	48
Tableau XI : Données techniques des instruments retenus	63
Tableau XII : Analyse des études retenues à l'aide de la grille QUADAS-2.....	67
Tableau XIII : Convention des méthodes d'évaluation de la validité et de la fidélité	73
Tableau XIV : Nombre d'études par mesures de validité et de fidélité recensées par instrument	77
Tableau XV : Nombre d'études par diagnostics ayant été mis à l'épreuve par instrument.....	78
Tableau XVI : Nombre d'études par types de fidélité et validité pour les TUS	79
Tableau XVII : Nombre d'études par types de fidélité et validité pour les TAG	82
Tableau XVIII : Nombre d'études par types de fidélité et validité pour les EDM.....	86
Tableau XIX : Nombre d'études par types de fidélité et validité pour les TC.....	90
Tableau XX : Nombre d'études par types de fidélité et validité pour les TDA/H	95
Tableau XXI : Nombre d'études par types de fidélité et validité pour les ESPT	101
Tableau XXII : Valeurs inférieures et supérieures des indices de validité et de fidélité par instrument	117

Liste des abréviations

ADI : *Adolescent Diagnostic Interview*

APS : *Adolescent Psychopathology Scale*

ASAM : *American Association of Addiction Medicine*

CAF : *Clinical Assessment Form*

CAPA : *Child and Adolescent Psychiatric Assessment*

CAS : *Child Assessment Schedule*

CASI : *Child and Adolescent Symptom Inventory*

CBCL: *Child Behavior Checklist*

ChIPS : *Children's Interview for Psychiatric Syndromes*

CCLT : Centre canadien de lutte contre les toxicomanies

CPLT : Comité permanent de lutte à la toxicomanie

CRD : centre de réadaptation en dépendance

CSMC : Commission de la santé mentale au Canada

DAWBA : *Development and Well-Being Assessment*

DIR : *Daily Incident Report*

DICA : *Diagnostic Interview for Children and Adolescents*

DISC : *Diagnostic Interview Schedule for Children*

DSM : *Diagnostic and statistical manual*

EBA : *evidence-based assessment*

EBP : *evidence-based practices*

EBPP : *evidence-based practices in psychology*

EBT : *evidence based treatment*

EN : logiciel bibliographique EndNote

ÉDM : épisode dépressif majeur

ÉSPT : état de stress post-traumatique

ESCC : Enquête sur la santé dans les collectivités canadiennes-santé mentale

GAIN : *Global Appraisal of Individual Needs*

ISCA : *Interview Schedule for Children and Adolescents*

ISQ : Institut de la statistique du Québec
K-SADS : *Schedule of Affective Disorders and Schizophrenia for Children*
NSDUH : *National Survey on Drug Use and Health*
NIMH : *National Institute of Mental Health*
PADDI : *Practical Adolescent Dual Diagnostic Interview*
OCHS-R : *Revised Ontario Child Health Study scales*
OSM : Organisation mondiale de la santé
PAI-A : *Personality Assessment Inventory-Adolescent*
PPD : psychopathologie développementale
SM : santé mentale
SPA : substance psychoactive
SS : synthèse systématique
TAG : trouble anxiété généralisé
TC : trouble des conduites
TDA/H : trouble du déficit de l'attention/hyperactivité
TM : trouble mental
TOP : trouble oppositionnel avec provocation
TUS : trouble lié à l'utilisation de substance psychoactive

À Marielle. Merci de m'avoir ouvert le chemin de la recherche.

*À tous les adolescents et adolescentes aux prises avec des difficultés, ne pas cesser
d'espérer...*

Remerciements

Entreprendre des études supérieures à 40 ans, lorsqu'on a un emploi à temps plein, un mari et quatre enfants, ne se fait pas sans heurt ni sans soutien. Un immense merci à Stéphane, mon conjoint des 25 dernières années, qui a été d'une patience sans limites. Merci à mes enfants, qui ont dû partager leur mère avec l'avancement des connaissances : Hugo, Aude, Maëlle et Robin.

Un merci du fond du cœur à mes confidentes, amies et collègues de travail, avec qui j'ai pu évacuer les frustrations, les moments de découragement, les victoires et les succès. Julie, Karine, Sophie et Sylvie, sans vous je n'y serais peut-être pas arrivée.

Merci à l'École de criminologie pour l'octroi d'une bourse de rédaction qui m'a permis de prendre plaisir à terminer ce mémoire. Et merci à mon employeur qui m'a donné un congé pour ce faire. Je dois souligner au passage la compréhension de quelques-uns de mes collègues gestionnaires : Dominique, François et Audrey, un gros merci.

Une synthèse systématique est une aventure dont on ne connaît l'ampleur qu'une fois terminée. Je ne peux passer sous silence l'aide et le soutien technique des bibliothécaires de l'Université de Montréal : Dany Létourneau, France Nadeau et Dominic Desaulniers. Combien de fois m'ont-ils déniché la perle rare tant recherchée ou réglé le problème technique tant craint ? D'innombrables...

Je tiens aussi à remercier Julie R. pour son aide indispensable à la mise en page, ainsi que Mme L'Espérance pour son support avec tout le côté administratif relié à une maîtrise.

Finalement, merci à mon directeur, M. Denis Lafortune. Son expertise, mais surtout son ouverture, sa compréhension et sa confiance, m'ont permis d'aller au bout de mes idées.

Introduction

L'heure est aux « meilleures pratiques », aux « pratiques gagnantes », basées sur des données scientifiquement mises à l'épreuve, méthodologiquement quasi irréfutables. Ces pratiques éprouvées sont au cœur d'un revirement dans le domaine des sciences humaines ; elles seraient gages de qualité et d'efficacité des services rendus à l'utilisateur. Elles visent le transfert des connaissances des chercheurs aux cliniciens (du scientifique au clinique), mais l'inverse est aussi vrai : l'utilisation des connaissances cliniques en vue de faire progresser la science. *L'American Psychological Association Presidential Task Force on Evidence-Based Practice* définit ainsi les meilleures pratiques en psychologie (*evidence-based practices in psychology* (EBPP) : ...*the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences* (2006 : 273). Cette définition rejoint celle de *l'Institute of Medicine* sur les meilleures pratiques (*evidence-based practices* (EBP): ... *the integration of best research evidence with clinical expertise and patient values* (2001 : 147).

C'est dans ce courant que se situent les synthèses systématiques (SS), qui visent à identifier les méthodes les plus efficaces en se basant sur les études scientifiques, robustes méthodologiquement et disponibles. Elles ont accompagné l'arrivée de ce que l'on nomme *evidence based treatment* (EBT) ou encore *evidence based practice* (EBP). Deux organisations reconnues s'y sont dévouées, visant à fournir l'information pertinente au moment de prendre une décision ou choisir une intervention. La Collaboration Campbell (2015), qui a été créée en 2000, vise la préparation, le maintien à jour et la dissémination des SS dans les domaines de l'intervention sociale, de l'éducation, du développement international ainsi que du crime et de la justice. La deuxième organisation est la Collaboration Cochrane (2015). Fondée en 1993 qui se spécialise dans les SS en santé humaine : de l'intervention à la prévention, en passant par le traitement, la réhabilitation et les tests diagnostiques. Son objectif est aussi de fournir de l'information pertinente et à jour pour soutenir des décisions prises par les intervenants, décideurs et bailleurs de fonds.

Hunsley et Mash (2008) rapportent que les EBA (*evidence-based assessment*), ou meilleures pratiques en évaluation, se méritent peu d'attention en comparaison des EBP. Or, il s'agit

d'un non-sens : comment parler des meilleures pratiques de traitement sans parler de ce qui y mène, soit les instruments d'évaluation? Les auteurs diront qu'ignorer la qualité de ces instruments place toute l'entreprise des meilleures pratiques en péril. Selon Hunsley et Mash, trois aspects de la définition des EBA sont importants. D'abord, les résultats de recherche qui supportent les théories de la psychopathologie et du développement normal devraient guider la sélection des construits à être évalués et leur processus d'évaluation. Ensuite, les auteurs suggèrent l'utilisation d'instruments à valeur psychométrique forte pour évaluer les construits sélectionnés. Ces instruments devraient avoir démontré leurs validité, fidélité et utilité clinique.

Finalement, même avec un instrument solide au niveau psychométrique, les auteurs rappellent que l'évaluation repose sur le processus décisionnel d'un clinicien qui est constamment confronté à la formulation d'hypothèses, à partir de données souvent incomplètes et incohérentes. En conséquence, une approche basée sur les meilleures pratiques d'évaluation devrait s'intéresser également à ce processus complexe de prise de décision à la lumière des erreurs de synthèse des données collectées, de leur interprétation, du coût associé au processus, ainsi que de leur impact sur les résultats cliniques. Il n'y a pas que l'instrument qui devrait répondre aux critères d'un EBA, mais aussi la façon dont le processus d'évaluation est mené.

C'est à partir de cette conception des EBA que le présent mémoire fut élaboré. Il se propose d'abord de mener une synthèse systématique sur les instruments d'évaluation de la comorbidité de l'Axe I du DMS-IV, incluant un trouble lié à l'utilisation des substances psychoactives (TUS), chez les adolescent(e)s¹. Dans un deuxième temps, il s'agira de faire une synthèse systématique des études ayant mis à l'épreuve ces instruments du point de vue de leur validité et fidélité. Les instruments retenus devront répondre aux deux premiers critères de Hunsley et Mash (2008) : avoir été construits selon les principes de la psychopathologie développementale (Cicchetti, 1984) et avoir fait l'objet d'une évaluation de leurs propriétés

¹ Pour souligner l'importance du genre, entre autre dans l'évaluation et le traitement, le mot adolescent a été féminisé partout où cela était possible dans le présent texte.

psychométriques. Le second critère sera vérifié à l'aide de la deuxième partie de la SS : le recensement des études ayant mis à l'épreuve la validité et la fidélité de certains diagnostics générés par ces instruments.

Le premier chapitre présente d'abord le cadre conceptuel du présent projet, soit la psychopathologie développementale. Celle-ci permet de démontrer la nécessité d'utiliser des instruments conçus pour les adolescent(e)s et leurs besoins propres. Ensuite, une présentation des termes utilisés pour définir la maladie mentale vient justifier le choix de la classification basée sur le DMS-IV-TR (American Psychiatric Association, 2000). Ce manuel diagnostique, peu importe les versions, ne fait pas l'unanimité, mais il représente ce qu'il y a de mieux jusqu'à présent pour faciliter la compréhension et la communication entre chercheurs et cliniciens.

Ensuite, un survol de travaux épidémiologiques montre l'importante prévalence des troubles liés à l'utilisation de substances psychoactives (TUS), des troubles mentaux et de leur comorbidité. Les modalités d'évaluation de ces troubles, ainsi que les recommandations émises au fil des années pour arriver à de meilleurs résultats sont aussi présentées. On y détaille les types d'évaluations, pour ensuite aborder l'importance que revêt la mise à l'épreuve de leur validité et de leur fidélité. Finalement, la SS est présentée comme le moyen le plus adéquat pour arriver à l'objectif principal du présent mémoire. Puisqu'aucun consensus n'existe sur les instruments à utiliser pour évaluer la comorbidité entre un TUS et autre trouble mental (TM) chez un adolescent(e), la présente SS permet d'apporter un début de réponse.

Le second chapitre présente en détail les étapes de la SS, de la recherche dans différentes bases de données bibliographiques à l'aide de mots clés, aux choix des critères d'inclusion et d'exclusion des études repérées. Les premières étapes ont visé l'identification des instruments d'évaluation les plus pertinents. Ainsi, 11 instruments ont été recensés. Ensuite est présentée la démarche d'identification des études ayant mis à l'épreuve chacun des 11 instruments (selon un des six diagnostics retenus). De cette manière, 57 études ont été

recensées. Le chapitre se termine avec la présentation de la grille d'analyse du QUADAS-2, avec laquelle la robustesse méthodologique de chaque étude a été évaluée.

Le chapitre 3 se subdivise en six sections : 1) une présentation de chacun des 11 instruments recensés (un bref historique, ses sous-sections, ses particularités et ses différentes versions) ; 2) suit une présentation des 57 études retenues (leur provenance, leur format, les types d'échantillons utilisés et la langue de la version ayant été testée) ; 3) les résultats de l'analyse menée à l'aide de la grille QUADAS-2 (robustesse méthodologique) ; 4) les types de validité et de fidélité recensés, incluant les méthodes utilisées pour les évaluer ; 5) une analyse des résultats par diagnostic (trouble lié aux substances (TUS), trouble anxiété généralisée (TAG), épisode dépressif majeur (ÉDM), trouble des conduites (TC), trouble du déficit de l'attention/hyperactivité (TDAH) et état de stress post-traumatique (ÉSPT) et finalement, 6) une analyse de ces mêmes résultats, mais par instrument d'évaluation.

La discussion porte sur les constats en lien avec les résultats obtenus. Les limites des études selon ce qu'en disent leurs auteurs, ainsi que les limites et les forces de la présente SS sont ensuite présentées. Un résumé des éléments à prendre en considération pour le choix du meilleur instrument selon les besoins d'une organisation est également abordé. Quelques possibilités de recherches futures concluent le présent mémoire.

Chapitre 1 : Recension des écrits

Le premier chapitre trace un portrait des connaissances actuelles en lien avec les troubles concomitants chez les adolescent(e)s, chiffres à l'appui. Il présente également tous les concepts sous-tendant la présente étude.

1. La perspective de la psychopathologie développementale

Les deux objectifs premiers de la psychopathologie développementale (PPD) sont a) d'appréhender la psychopathologie chez un individu, à l'aide d'une meilleure description du développement psychologique normal et, inversement, b) de comprendre les changements normaux pour mieux identifier l'émergence du pathologique. Cette science intégrative utilise les découvertes faites dans nombre de disciplines pour en arriver à une meilleure compréhension des processus développementaux et de la manière dont ces derniers culminent dans des résultats « extrêmes », autant négatifs (pathologie) que positifs (résilience) (Cicchetti, 1984, 1987; 1993; Cicchetti et Toth, 2009). Un des éléments centraux de la PPD, c'est l'idée qu'on ne peut pas concevoir la psychopathologie des enfants et des adolescent(e)s de la même façon que celle des adultes (Cicchetti, 1987; Grisso, Vincent et Seagrave, 2005; Ingram et Price, 2010; Steinberg, 2002). Malgré l'augmentation des travaux et des études sur la psychopathologie des adolescent(e)s au cours des dernières années, il n'y a toujours pas de consensus sur les critères pour la définir (Ingram et Price, 2010). La PPD offre un cadre conceptuel tout indiqué pour explorer plus à fond des critères plus adaptés à cette période de mouvance de l'adolescence lorsque vient le temps d'évaluer divers éléments, dont des symptômes de santé mentale ou encore de TUS. C'est par ailleurs le cadre conceptuel du présent projet. La section suivante présente l'évolution de la terminologie de la santé mentale, d'où l'on semble effectivement avoir évacué l'aspect développementale.

2. L'évolution de la terminologie relative à la santé mentale

Note : durant la période de réalisation du présent mémoire, le DSM-5 (American Psychiatric Association, 2013) a fait son apparition. Évidemment, aucune des études répertoriées n'utilise les critères de cette nouvelle version. Les analyses seront donc effectuées en lien avec les versions antérieures du DSM.

L'Organisation mondiale de la santé (OMS) définit la santé mentale (SM) comme étant bien plus que l'absence de maladie mentale. «La santé est un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité.» (Organisation mondiale de la Santé, 1946 : 100). Cet état permet aux gens de se réaliser et de prendre part activement à la société.

Selon la Commission de la santé mentale au Canada (CSMC, 2009), la SM permet de faire face aux stress de la vie. Elle peut agir comme rempart contre la maladie physique et les troubles mentaux ou aider à y faire face. Elle peut aussi jouer de façon positive sur le rétablissement, ce processus qui permet de vivre, malgré la maladie, une vie satisfaisante et même enrichissante. En effet, sans qu'il n'y ait guérison, il peut y avoir un contrôle suffisant des symptômes.

Toujours selon la CSMC (2013 : 4), les termes «maladie mentale» et «troubles mentaux» comprennent « toute la gamme de comportements, de pensées et d'émotions qui peuvent être à l'origine d'une détresse, d'une souffrance ou d'une incapacité dont les effets peuvent se faire sentir à l'école, au travail ou dans les rapports sociaux et familiaux et peuvent empêcher une personne de vivre de façon autonome. » Leur étiologie repose sur un « ensemble complexe de facteurs sociaux, économiques, psychologiques, biologiques et génétiques ».

Depuis 1980, le DSM a introduit le terme de trouble mental, qui remplace celui de maladie. L'*American Psychiatric Association* (1980) le définit ainsi :

- Chaque trouble mental est conçu comme un modèle ou un syndrome comportemental ou psychologique cliniquement significatif, survenant chez un individu et associé à une détresse concomitante (p. ex. symptômes de souffrance), à un handicap (p. ex. altération d'un ou de plusieurs domaines de fonctionnement), ou à un risque significatif élevé de décès, de souffrance, de handicap ou de perte importante...
- Ce modèle ou syndrome ne doit pas être simplement la réponse attendue et culturellement admise à un événement particulier...

- Il doit être considéré comme la manifestation d'un dysfonctionnement comportemental, psychologique ou biologique de l'individu...

La conceptualisation du DSM-IV-TR (American Psychiatric Association, 2000) est basée sur des catégories (*clusters*). Le diagnostic est alors déterminé par la présence ou l'absence de critères précis. Selon Grisso et Vincent (2005), il s'agit là d'une classification utile, voire la meilleure disponible à ce jour, mais qui est parfois mal adaptée aux besoins des adolescent(e)s. Malgré l'introduction d'une section « troubles pouvant apparaître durant l'enfance », par ses instruments d'évaluation, diagnostics et classifications, le DSM véhicule essentiellement un modèle adulte de la psychopathologie infantile. Mais le sujet demeure d'intérêt : arriver à utiliser les critères avec les adultes tout comme avec les adolescent(e)s est une demande de longue date de la part des intervenants œuvrant dans ce domaine (Grisso et al., 2005). Malgré leurs limites, ces catégories continuent de baliser la recherche et l'application clinique de traitement.

Toujours selon Grisso et Vincent, trois autres approches de classification peuvent compléter l'information diagnostique (sans relier cette information à un diagnostic) et broser un portrait plus complet des problèmes de santé mentale chez les adolescent(e)s, permettant un plan de traitement plus efficace. Il est d'abord possible d'utiliser la condition symptomatique (la présence et la sévérité de symptômes). Cette méthode, souvent utilisée par les outils de détection, ne permet pas une évaluation en profondeur. La seconde approche suppose le regroupement de comportements, sentiments et pensées regroupés par « dimension conceptuelle » de la psychopathologie en utilisant un score sur une échelle. Cette approche a été développée par Achenbach (1993; Achenbach et Edelbrock, 1984) et comprend toute une famille d'instruments, dont le CBCL². Chaque dimension (huit au total), par exemple « anxieux-dépressif » ou encore « problème d'attention », est évaluée en fonction de l'atteinte du fonctionnement. On peut en tirer deux syndromes bien connus, soit les troubles internalisés et externalisés. Finalement, la troisième approche est dite « orientée vers les problèmes » (*problem oriented*). Elle consiste à identifier les forces et les difficultés de

² *Child Behaviour Checklist* (Achenbach, 1991b)

fonctionnement dans les sphères courantes de la vie d'un adolescent(e) : à l'école, avec la famille, les amis, dans les loisirs, etc.

Le choix de la présente SS s'est arrêté sur les instruments en mesure de produire des diagnostics, ce qui est différent des approches de classification mentionnées ci-haut, qui peuvent alors devenir complémentaires.

La sous-section suivante démontre par contre la complexité de l'évaluation de ces critères menant à un diagnostic : rarement évalue-t-on une situation exempte de recouvrements de symptômes. La comorbidité semble plutôt être un phénomène très courant (incluant un trouble de SM et un TUS), comme le démontre les chiffres qui suivent.

3. Comorbidité

Le terme de comorbidité désigne la présence de plusieurs maux de façon simultanée. La personne reçoit alors deux diagnostics ou plus. Il peut s'agir d'une combinaison de troubles de la santé physique et/ou de la santé psychologique ou psychiatrique. Dans le domaine de la psychiatrie, la comorbidité désigne le fait de présenter des signes et symptômes qui répondent à plusieurs critères et ainsi correspondent à plus d'une catégorie diagnostique. Selon la construction du DSM, il peut y avoir cooccurrence de deux diagnostics (ou plus) sur un même axe ou alors sur deux et même trois axes différents. Selon nombre d'auteurs, la comorbidité psychiatrique semble être la norme plutôt que l'exception autant chez les adultes que chez les adolescent(e)s (Chan, Dennis et Funk, 2008; Diamond et al., 2006; Perepletchikova, Krystal et Kaufman, 2008).

Les jeunes garçons judiciairisés ne font pas exception (Grisso, 2004, 2008; Vincent, Grisso, Terry et Banks, 2008) ni les jeunes filles judiciairisées (Goldstein et al., 2003). Il est en de même pour les adolescent(e)s aux prises avec des difficultés reliées à la consommation de SPA (Costello, Erkanli, Federman et Angold, 1999; Diamond et al., 2006; Federman, Costello, Angold, Farmer et Erkanli, 1997; Kaminer et Bukstein, 2008; Marmorstein, 2010; Saban et Flisher, 2010; Sung, Erkanli, Angold et Costello, 2004; Vitaro, Rorive, Zoccolillo, Élisabeth et Tremblay, 2001; Wu et al., 2010). C'est d'ailleurs ce type de comorbidité à l'Axe I, comprenant

un trouble mental (TM) et un trouble lié à l'utilisation d'une substance (TUS) qui fait l'objet du présent mémoire.

Différentes interactions sont possibles entre les TUS et les TM (Bolton, Robinson et Sareen, 2009; Tupker, 2004) :

- La consommation de psychotropes peut *créer* des symptômes psychiatriques,
- La consommation de psychotropes peut *déclencher* certains troubles mentaux s'il y a prédisposition,
- La consommation de psychotropes peut *exacerber* la symptomatologie des troubles mentaux,
- Les symptômes reliés à la consommation de psychotropes peuvent *imiter* ceux des troubles mentaux,
- Les symptômes reliés à la consommation de psychotropes peuvent *masquer* ceux reliés aux troubles mentaux,
- Enfin, la consommation de substances psychotropes peut être utilisée pour *gérer la détresse* associée à des troubles mentaux (automédication).

Ce survol rapide permet de comprendre les subtilités de l'évaluation des troubles comorbides, surtout lorsqu'un TUS est identifié. La sous-section suivante apportera des précisions sur ce que comporte un TUS, en distinguant l'abus et les différents types de dépendance, mais également sur le manque d'éléments adaptés à la réalité des adolescent(e)s.

4. Consommation de substances psychoactives (SPA) : abus et/ou dépendance.

Différents critères sont utilisés pour identifier les problèmes de consommation de SPA chez les adolescent(e)s. Mais d'abord, il convient de préciser que les dépendances englobent non seulement la consommation de psychotropes (drogues, médicaments (prescrits ou non), alcool), mais aussi les dépendances sans substance, telles que le jeu pathologique (Desrodières et Jacques, 2009; Gagnon, Hickey, Hugron et Racine, 2006; Griffiths et Sutherland, 1998;

Kairouz et al., 2005; Ladouceur, 1996; Tremblay, Allaire, Dufresne, Lecavalier et Neveu, 2004) ou encore la cyberdépendance (Dufour, Nadeau et Gagnon, 2014; Goyette et Nadeau, 2008; Griffiths, Wood, Gupta et Derevensky, 2000; Griffiths, Davies et Chappell, 2004; Kairouz et al., 2005; Wood, Griffiths, Derevensky et Gupta, 2002). Ces phénomènes touchent autant les adultes et les adolescent(e)s.

Le jeu pathologique a été inclus dans le DSM-IV-TR (American Psychiatric Association, 2000), mais classé dans la catégorie des « troubles du contrôle des impulsions non classés ailleurs ». Suite à la parution du DSM-5 (American Psychiatric Association, 2013), le jeu pathologique a été inclus dans les *Troubles reliés à une substance et troubles addictifs*, puisque les études montrent que le jeu active le système cérébral dit «de récompense», ce qui produit des effets similaires aux drogues. La thèse doctorale de Aulakh (2010), supervisé par deux sommités en jeu pathologique (Dr Jeffrey Derevensky et Dre Rina Gupta), résume bien les avancés en lien avec le jeu pathologique et la neuroscience. Les conclusions de Aulakh supportent la thèse de la présence du phénomène de la neuroadaptation (relié entre autre à la tolérance) que l'on retrouve chez les consommateurs de SPA. Chez les joueurs pathologiques, ce phénomène se développerait également au fils des séances de jeu et se manifesterait par une augmentation des paris ou de la fréquence de jeu pour atteindre un niveau d'excitation comparable d'un épisode de jeu à un autre.

Au Québec, le jeu pathologique est considéré comme une dépendance et traité dans les Centres de réadaptation en dépendances (CRD) (Desrodiers et Jacques, 2009). Selon les derniers chiffres de l'Institut de la statistique du Québec (ISQ) dans *l'Enquête québécoise sur le tabac, l'alcool, la drogue et le jeu chez les élèves du secondaire* (Traoré et al., 2014), les taux de prévalence chez les jeunes de la première secondaire à la cinquième seraient à la baisse. Ils se situeraient à 2,7% pour les joueurs à risque et 0,9% pour les joueurs pathologiques probables³.

³ À risque : deux à trois critères diagnostiques sur un maximum de neuf ; pathologique probable : quatre critères ou plus (Fisher, S. (2000). Developing the DSM-IV-MR-J criteria to identify adolescent problem gambling in non-clinical populations. *Journal of Gambling Studies*, 16(2-3), 253-273).

La cyberdépendance est également considérée sous l'angle des dépendances. Actuellement, 13 CRD au Québec, ainsi qu'un centre de traitement interne pour jeunes, l'étudient chez les adolescent(e)s⁴. Ce champ de recherche est en émergence comme le démontrent Dufour, Nadeau et Gagnon (2014). Suite à des travaux de recherche actuels et futurs, il est possible qu'éventuellement, la cyberdépendance se retrouve dans la même catégorie que la dépendance aux SPA et au jeu pathologique.

Cela dit, aux fins du présent projet, seules les dépendances aux SPA seront retenues. Dans le DSM-IV-TR, elles sont décrites à l'Axe I sous le vocable des « troubles liés à l'utilisation d'une substance » (TUS). Les critères d'identification de cette problématique chez les adolescent(e)s varient selon l'instrument de détection ou d'évaluation utilisé.

La détection est une première étape, nécessaire à la réalisation d'une évaluation plus complète. Pour la permettre, dans plusieurs régions du Québec, on utilise la Grille de dépistage de consommation problématique d'alcool et de drogues chez les adolescents et les adolescentes ou « DEP-ADO » (RISQ, 2007). Elle est basée sur la fréquence et la diversité de la consommation et les conséquences négatives engendrées par cette dernière. Le score obtenu donne un aperçu de la gravité de la consommation et permet d'orienter l'adolescent(e) vers les ressources adéquates (services de première ou deuxième ligne).

Les critères du DSM-IV-TR sont les plus utilisés pour effectuer un diagnostic d'abus ou de dépendance aux SPA. Le Tableau qui suit présente d'abord une définition des concepts d'abus et de dépendance, ainsi que les critères servant à poser ces diagnostics (American Psychiatric Association, 2000).

⁴ L'étude Virtuado a comme objectif de documenter le profil clinique des adolescent(e)s, âgé(e)s de 14 à 17 ans, aux prises avec un usage problématique d'internet et inscrit(e)s dans un des CRD du Québec, ainsi qu'au Grand Chemin (centre offrant des services de réadaptation en interne).

Tableau I : Critères d'abus et de dépendance à une substance selon le DSM-IV-TR

Abus d'une substance selon le DSM-IV

Mode d'utilisation inadéquat d'une substance, conduisant à une altération du fonctionnement ou à une souffrance cliniquement significative, comme en témoignent au moins une des manifestations suivantes, survenant à n'importe quel moment sur la même période de 12 mois :

- 1 Utilisation répétée d'une substance conduisant à l'incapacité de remplir des obligations majeures, au travail, à l'école ou à la maison ;
- 2 Utilisation répétée d'une substance dans des situations où cela peut être physiquement dangereux ;
- 3 Problèmes judiciaires répétés liés à l'utilisation d'une substance ;
- 4 Utilisation de la substance malgré des problèmes interpersonnels ou sociaux, persistants ou récurrents.

Dépendance à une substance selon le DSM-IV

Mode d'utilisation inadapté d'une substance, conduisant à une altération du fonctionnement ou à une souffrance cliniquement significative, comme en témoignent trois (ou plus) des manifestations suivantes, survenant à n'importe quel moment sur la même période de 12 mois :

- 1 Tolérance (besoin de plus grandes quantités ou effet diminué en cas d'utilisation continue) ;
- 2 Sevrage (syndrome ou soulagement avec consommation de la substance (ou substance de la même famille) ;
- 3 Substance souvent prise en quantité plus importante ou plus prolongée que prévu ;
- 4 Désir persistant de cesser ou efforts infructueux ;
- 5 Beaucoup de temps passé à des activités nécessaires pour obtenir la substance, à l'utiliser et à en récupérer ;
- 6 Activités sociales, professionnelles ou de loisir importantes abandonnées ;
- 7 Utilisation de la substance poursuivie bien que la personne sache avoir un problème psychologique ou physique persistant ou récurrent, causé ou exacerbé par la substance.

Ces critères, conçus et identifiés pour diagnostiquer les adultes, sont aussi ceux utilisés avec les adolescent(e)s. Sont-ils adéquats ? La question semble pertinente, puisque plusieurs écrits ont été publiés à ce sujet.

Chassin et ses collaborateurs (Chassin, Beltran, Lee, Haller et Villalta, 2010) montrent bien que les critères d'abus et dépendance du DSM-IV-TR ne collent pas suffisamment à la réalité développementale des adolescent(e)s, ni même à la réalité des adultes. La validité d'une telle subdivision (abus versus dépendance) a été remise en question, les auteurs suggérant plutôt un continuum⁵ de problèmes, une proposition valide autant pour les adultes que pour les adolescent(e)s (Harthman et al., 2008; Langenbucher et al., 2004; Martin, Chung, Kirisci et Langenbucher, 2006). Chassin et ses collaborateurs rappellent que l'application des critères du DSM-IV-TR aux adolescent(e)s génère souvent des « diagnostics orphelins », soit des résultats qui se situent juste sous le seuil diagnostique. Cela représente, à leur avis, un mauvais ajustement des critères au stade développemental des adolescent(e)s. Il en irait ainsi des surestimations de la tolérance et des symptômes de sevrage chez les adolescent(e)s : plusieurs jeunes se retrouvent avec un diagnostic de dépendance malgré un niveau de consommation relativement faible.

En utilisant les critères du DSM-IV-TR, il faudrait donc tenir compte du niveau développemental pour que le diagnostic soit adapté à la réalité des adolescent(e)s et ce, à plusieurs égards (neurobiologique, psychobiologique, physiologique, génétique, etc.). Une bonne connaissance des étapes de développement des adolescent(e)s, mais aussi de leurs trajectoires de consommation en fonction de l'âge est nécessaire pour qu'une évaluation faite à l'aide des critères du DMS soit adaptée. La PPD prend ici tout son sens : comprendre le développement normal pour comprendre « l'anormalité » de la consommation des adolescent(e)s et être en mesure d'ajuster les critères diagnostiques d'un TUS.

En 1998 a été réalisée une des premières études évaluant l'utilisation des critères abus/dépendance du DSM-IV auprès d'une population représentative des adolescent(e)s aux États-Unis. Il s'agissait de 74 008 jeunes d'écoles publiques, des niveaux 9 à 12, soit

⁵ Concernant le continuum, c'est chose faite dans le DSM-5.

l'équivalent de la troisième secondaire à la première année de CÉGEP au Québec, (Harrison, Fulkerson et Beebe, 1998). Les auteurs, utilisant les données de ce sondage (*Minnesota Student Survey*), rappellent d'abord que les critères diagnostiques sont incontournables, que ce soit pour permettre une communication efficace entre chercheurs et cliniciens, effectuer des références vers les services et traitements appropriés ou permettre l'accès au traitement lorsqu'il est remboursé par des compagnies d'assurance. Ils ajoutent, que d'un point de vue épidémiologique, les énoncés de certains critères sont peu clairs et parfois peu adaptés à la réalité des adolescent(e)s (ex. : « en quantité plus grande », « plus prolongé que prévu » ou « obligations majeures »). De plus, on souligne les difficultés soulevées par les critères de l'abus et de la dépendance : un seul critère est nécessaire pour poser un diagnostic d'abus, tandis qu'il en faut trois pour celui de dépendance. Selon Harrison, Fulkerson et Beebe (1998) l'objectif pour les adolescent(e)s étant de cesser la consommation, une telle exigence n'a pas de valeur clinique. Probablement qu'un continuum répondrait mieux aux besoins d'évaluation à cette période de la vie.

Winters, Latimer et Stinchfield (1999) ont pour leur part évalué la validité des critères du DSM-IV auprès d'un échantillon constitué de 772 adolescent(e)s étasunien(ne)s traité(e)s dans des cliniques pour leur consommation. Ils rapportent qu'au moment où ont été apportés les changements de critères entre le DSM-III-R et le DSM-IV, très peu de données sur la consommation des adolescent(e)s étaient disponibles. Il était donc difficile de juger de l'adéquation des critères du DSM-III-R. Par ailleurs, peu d'études sur le terrain ayant été réalisées avec la version subséquente, il est ardu de savoir s'il y a eu depuis amélioration ou recul. Les auteurs en viennent malgré tout à la conclusion que les critères du DSM-IV sont valides. Ils recommandent de revoir les énoncés des critères en lien avec les quantités et fréquences de l'abus pour les rendre moins flous, tout comme la définition du concept de tolérance chez les adolescent(e)s. Cette position, qui rejoint précisément celle de Chassin et ses collaborateurs (2010) et de Harrison, Fulkerson et Beebe (1998), est bien représentée par le passage suivant:

« Indeed, the favorable evidence for DSM-IV criteria does not preclude the need for continuing the search on the definition of adolescent substance abuse disorders. The extent to which the DSM-IV system is optimal for adolescent is still worthy of further examination. We believe that efforts to design developmentally appropriate taxonomies, such as the DSM-PC for primary care physicians (Wolraich, 1996, cité dans Winters, 1997), are a step in the right direction. Another potentially fruitful research avenue would be to explore the clinical significance of individuals who do not meet diagnostic criteria yet report several negative consequences related to drug use. Also, future diagnostic investigation studies would benefit from the inclusion of nonclinical adolescent populations. » (Winters et al., 1999: 343)

Chung et Martin (2005) réanalysent la validité du critère de tolérance (un élément central de la dépendance) en lien avec la consommation d'alcool chez les adolescent(e)s. Malgré la petite taille de leur échantillon (50 adolescent(e)s étasunien(ne)s traité(e)s pour leur consommation), les auteurs arrivent aussi à la conclusion que la définition opérationnelle de la tolérance, ainsi que les expressions « en plus grande quantité » et « plus prolongé que prévu » font en sorte que plusieurs symptômes sont mal évalués chez les adolescent(e)s. Ces derniers ne saisissent pas toujours toute la portée de ces concepts, qui sont peu en lien avec leur réalité. Une amélioration des définitions devrait tenir compte de la PPD. On suggère que les intervieweurs ou cliniciens soient en mesure de bien comprendre ces concepts pour mieux les expliquer aux adolescent(e)s. On suggère également l'utilisation de données supplémentaires, biologiques notamment. L'utilisation du taux d'alcool dans le sang (*blood alcohol concentration* ou BAC) est d'ailleurs suggérée dans le DSM-IV-TR.

Une équipe de chercheurs du *National Institute of Mental Health* (NIMH,2010) suggère de revoir les catégories diagnostiques à la lumière des avancées en neuroscience et en génétique pour tenter d'améliorer les résultats des traitements. Le RDoC : *Research Domain Criteria project* est un projet visant à mettre sur pied un cadre de recherche fondé sur la pathophysiologie, la génomique et la neuroscience. Cela pourrait ultimement donner lieu à une classification diagnostique mieux adaptée : « *Our expectations [...] is that identifying*

syndromes based on pathophysiology will eventually be able to improve outcomes. »
(NIMH,2010: 748)

Suite à la diffusion des changements prévus par le DSM-5 par rapport aux critères des TUS, Winters, Martin et Chung (2011) ont fait des suggestions au comité de travail du DSM-5, en ce qui concerne la situation des adolescent(e)s. En effet, les études qu'ils ont menées suggèrent l'importance d'ajuster les concepts de tolérance et de sevrage, ainsi que d'utiliser un continuum de symptômes sans subdivision (abus versus dépendance). Les auteurs y sont aussi allés de suggestions par rapport aux définitions des conduites de risque et des envies de consommer (*craving*), ainsi que par rapport au seuil diagnostique d'un TUS. À la lecture de la section *Troubles reliés à une substance et troubles addictifs* du DMS-5 (2013), on constate que certaines de ces suggestions ont été retenues. Les diagnostics d'abus et de dépendance ont été fusionnés en un seul trouble et le seuil diagnostique est passé à deux critères, mais avec la possibilité de préciser la sévérité : deux à trois critères dénotant un trouble léger, quatre à cinq, un trouble modéré, six et plus, un trouble sévère.

Il est donc clair que l'évaluation des critères d'un TUS est complexe et présentement peu adaptée à la réalité des adolescent(e)s, même si encore très utiles et utilisés. Par contre, avec l'arrivée du DSM-5, de nouvelles possibilités d'évaluation se pointent à l'horizon et permettront de mieux rendre compte des particularités de ces jeunes. Mais qu'en est-il vraiment des taux de comorbidité (incluant un TUS) chez les adolescent(e)s ? La sous-section suivante en trace le portrait pour une compréhension encore plus poussée.

5. Troubles avec comorbidité chez les adolescents et adolescentes

Bien que moins nombreuses que celles portant sur la comorbidité des adultes, plusieurs études font état d'un taux élevé de comorbidité chez les adolescent(e)s (Bender, Springer et Kim, 2006; Chan et al., 2008; Merikangas, Avenevoli, Costello, Doreen et Kessler, 2009; Vincent et al., 2008). Au Québec, même si Bertrand, Ménard et Plourde (Bertrand et al., 2009; 2004), ainsi que Laventure (2008) se sont intéressés au phénomène, on constate un manque de publications en ce domaine (Arial, 2009), ce qui peut limiter l'orientation des jeunes vers des services appropriés. Les chiffres qui suivent portent sur la prévalence des troubles

mentaux chez les adolescent(e)s et la comorbidité et confirment que le phénomène est important. Ils appuient par le fait même la nécessité de mieux le comprendre et l'évaluer pour bien l'identifier et le traiter.

5.1 Taux de prévalence chez les adolescents et adolescentes

5.1.1 Au plan international

Dans le cadre du projet *World Health Organization (WHO)'s World Mental Health (WMH) Surveys*, Kessler et ses collaborateurs (2007), présentent des données provenant de 17 pays d'Afrique, des Amériques, de l'Asie et du Pacifique, ainsi que de l'Europe et du Moyen-Orient (N = 85 052). Bien que la collecte de données ait été effectuée principalement auprès d'adultes de plus de 18 ans (une seule étude inclut des individus de plus de 16 ans, une autre, de plus de 20 ans et une dernière, de plus de 21 ans), on y rapporte que certains troubles émergent durant l'enfance et l'adolescence. Par exemple, des troubles anxieux apparaissent à partir de 7 à 14 ans, tandis que des troubles reliés à l'impulsivité et au contrôle le font à partir de 7 à 15 ans. Pour les TUS, l'âge d'apparition se situe à la « mi-adolescence » et la prévalence tend à augmenter jusqu'au début de l'âge adulte.

Une SS et une méta analyse menée par Fazel, Doll et Långström (2008), à partir de 25 études qui regroupent 16 750 jeunes de moins de 19 ans incarcérés provenant de huit pays⁶ différents, arrivent à la conclusion suivante: 3% des jeunes délinquants présente une psychose, 11% des garçons et 29% des filles un épisode dépressif majeur (ÉDM) et un peu plus de 50% un trouble des conduites (TC). La prévalence moyenne du trouble déficitaire de l'attention avec ou sans hyperactivité (TDA/H) est de 11,7% (IC 95% = 4,1% à 19,2%) chez les garçons et de 18,5% chez les filles (IC = 9,7% à 27,7%). Pour ces mêmes troubles, on a pu démontrer que la prévalence était supérieure chez les jeunes incarcérés que dans la population générale du même âge.

⁶ 15 des États-Unis, quatre du Royaume Unis et une de l'Australie, de la Russie, de la Hollande, du Danemark, du Canada et de l'Espagne.

5.1.2 Aux États-Unis

La première étude à porter sur la forte prévalence des troubles cooccurrents entre SPA et autres troubles mentaux, fut celle de Regier et ses collaborateurs (Regier et al., 1990) qui analysèrent les données de *l'Epidemiological Catchment Area study* aux États-Unis. Plusieurs autres suivirent.

Aux États-Unis, le *National Comorbidity Survey-Adolescents Supplement (NCS-AS)* est une source de renseignements importante (10 123 adolescent(e)s de 13 à 17 ans) qui, malgré certaines limites, permet d'estimer la prévalence et la concomitance des troubles mentaux chez les adolescent(e)s. Voici les taux que Merikangas, Avenevoli, Costello, Doreen et Kessler (2009) ont obtenus (prévalence à vie) :

Tableau II : Prévalence de TM aux États-Unis (13 à 17 ans)

Troubles	Taux
Troubles anxieux	31,9%
Troubles du comportement	19,1%
Troubles de l'humeur	14,3%
Troubles liés aux substances	11,4%
Alcool	6,4%
Drogues	8,9%
Troubles alimentaires	2,7%
Occurrence de comorbidité	42%

(Merikangas et al., 2009)

Une méta analyse (Chan et al., 2008) confirme ces données, concluant que 42,9% des adolescent(e)s de 15 à 17 ans répondent positivement aux critères de comorbidité. Perepletchikova, Krystal et Kaufman résumant en ces mots la situation: «*co morbidity is the rule, not the exception in adolescent alcohol use disorders*» (2008 : 1136).

Des chiffres récents tirés de la *National Survey on Drug Use and Health (NSDUH)* (Substance Abuse and Mental Health Services Administration et Center for Behavioral Health Statistics

Quality, 2014), viennent appuyer ces résultats. Le NSDUH porte sur la consommation de tabac, alcool, drogues illicites et utilisation de médicaments non prescrits aux États-Unis. Il s'agit d'un échantillon de civils, représentatif du point de vue de chaque État, mais aussi du pays en entier (excluant les prisons et les patients hospitalisés à long terme). Cela implique 150 000 adresses et 70 000 personnes interviewées chaque année, avec une surreprésentation des adolescent(e)s et jeunes adultes. Selon les chiffres publiés en 2014, 10,7% des jeunes de 12 à 17 ans ont souffert d'un EDM au cours de la dernière année. Cela représente 2,6 millions d'adolescent(e)s aux États-Unis. Plus préoccupant encore, 7,7% ont connu un EDM avec dysfonctionnement sévère. En ce qui concerne la consommation de drogues, 8,8% des jeunes de 12 à 17 ans avaient consommé des drogues au cours du dernier mois. Le cannabis est de loin la substance illicite préférée des jeunes États-Uniens, 7,1% en ayant consommé au cours du mois précédent le sondage. En regard de la consommation d'alcool, 11,6% rapporte en avoir consommé au cours du dernier mois, dont 6,2% de façon excessive et 1,2% de façon excessive et répétitive (*heavy drinkers*). Il est estimé que 5,2% des jeunes ont présenté un TUS au cours de la dernière année. Finalement, en lien avec les troubles cooccurrents, 1,4% (soit 350 000 adolescent(e)s) rapportent un TUS et un EDM au cours de la dernière année.

5.1.3 Au Canada

Dans une analyse des données de l'Enquête sur la santé dans les collectivités canadiennes-santé mentale (ESCC) (données recueillies en 2002), Rush et ses collaborateurs (2008) brossent un portrait de la cooccurrence des troubles mentaux et des TUS au Canada. Selon eux, la catégorie des personnes âgées de 15 à 24 ans a le plus haut taux de cooccurrence des TUS avec un trouble de l'humeur ou anxieux, soit 3,8%, comme le démontre le Tableau III. Chez les individus souffrant d'un TM, le taux de cooccurrence avec les TUS chez les 15-24 se situe à 32,4% et décroît avec l'âge. Chez les individus avec des TUS, la prévalence de cooccurrence avec un autre TM est plus stable, soit autour de 15%. Tous âges confondus, les Canadiens avec des TUS ont plus de deux fois plus de risque de cooccurrence et ceux avec une dépendance aux SPA, de cinq à six fois plus de risque.

Tableau III : Prévalence TUS et TM au Canada

Cooccurrence	Âge				
	15-24	25-34	35-44	45-54	+55
	12 ms	12 ms	12 ms	12 ms	12 ms
TUS + Humeur/Anx.	3,88	2,6	2	1,1	0,2
TM + TUS	32,4	26,4	20,8	13,2	5,1
TUS + TM	15,6	15,6	15,6	15,7	10

Taux le plus élevé (ESCC, 2002)

Les données obtenues en 2012 par l'ESCC (Tableau IV), bien qu'elles ne soient pas analysées de la même façon, montrent des tendances semblables (Pearson, Janz et Ali., 2013).

Tableau IV : Prévalence TUS et TM au Canada

Trouble	Âge							
	15-24		25-44		45-64		+65	
	vie	12 ms	vie	12 ms	vie	12 ms	vie	12 ms
Tout trouble*	30,7	18,5	34,8	11,4	37,4	8,3	23,2	3,2
EDM	10,7	7,1	11,9	5,4	13,1	4,5	7,2	1,6
TAG	6,3	2,4	8,9	2,9	10,8	3	6,4	1,2
Troubles liés aux SPA	11,9	21,3	23,7	4,6	23,6	2,5	13,8	0,7**

Taux le plus élevé 12 derniers mois (ESCC 2012)

* Épisode dépressif majeur, trouble bipolaire, trouble d'anxiété généralisée, dépendance à l'alcool, abus d'alcool, dépendance au cannabis, abus de cannabis, dépendance aux autres drogues (excluant cannabis), et d'abus d'autres drogues (excluant cannabis).

** Utiliser avec prudence

L'Enquête de surveillance canadienne de la consommation d'alcool et de drogues de 2012 (Santé Canada, 2012) confirme que les taux de consommation de drogues sont plus élevés pour les jeunes que les adultes. Le taux de consommation de cannabis (dernier 12 mois) est de 20,3% chez les 15-25 ans comparativement à 8,4% chez les adultes de plus de 25 ans. Il en va de même pour l'ensemble des drogues illicites: 22,2% versus 9,3%. Au chapitre de l'alcool, les chiffres sont inversés, mais élevés dans tous les cas : 70% versus 80%.

Toujours en regard de la santé mentale des Canadiens, dans une étude de la Commission de la santé mentale au Canada (2013), on estimait que plus de 6,7 million de Canadiens vivent avec un TM, dont plus de 1 million des jeunes âgés de 9 à 19 ans. Les troubles de l'humeur et de l'anxiété sont les plus répandus, avec un taux de 11,75% de la population, tous âges confondus. Chez les jeunes de 9 à 19 ans, ce taux était plutôt de 12,1%. En regard des TUS les taux avancés sont de 5,9% tous âges confondus, comparativement à 6,8% pour les jeunes. Chez les mineurs, les troubles les plus présents sont (en ordre croissant) : le trouble de l'opposition (TOP), le TC, le TDA/H, les TUS et les troubles de l'humeur et d'anxiété (taux combiné). Enfin, les études montrent que le fait d'être affecté d'une de ces conditions dans l'enfance ou l'adolescence, augmente les risques d'être affecté de l'une ou l'autre de ces conditions durant l'adolescence ou l'âge adulte.

5.1.4 Au Québec

Les chiffres rapportés par Kairouz, Boyer, Nadeau, Perreault et Fiset (2008) pour le Québec, toujours pour la catégorie des 15-24 ans, permettent de faire certaines comparaisons.

Tableau V : Prévalence TUS et TM au Québec chez les 15-24 ans

PRÉVALENCE TUS ET TM	
Au moins un trouble de l'humeur	13,7%
Au moins un trouble anxieux	13,6%
Buveurs réguliers*	30,4%
Boire excessif	27,9%
Dépendance	5,4%
Troubles causés par alcool	2,3%
Usage de cannabis 12 mois	38,9%
Usage substances illicites, excluant cannabis, 12 mois	9,9%

*Une corrélation est présente entre troubles reliés à l'alcool et certains troubles mentaux. (Kairouz et al., 2008)

De nouveau, pour la quasi-totalité des items mesurés, les pourcentages sont plus élevés pour la catégorie des 15-24 ans. En somme, les personnes qui présentent une comorbidité sont plus susceptibles de se retrouver dans les services et de manifester plus de détresse (Kairouz et al., 2008).

Au niveau de la consommation de SPA, un rapport récent de l'ISQ (Traoré et al., 2014) présente des résultats comparables. Malgré une baisse observable au cours des dernières années de la consommation régulière⁷ et excessive⁸ d'alcool, les chiffres demeurent préoccupants pour les adolescents et adolescentes du niveau secondaire⁹. En effet, 16.8% des jeunes du secondaire ont eu une consommation excessive et répétitive¹⁰ au cours de la dernière année, ce taux grimant à 30.3% chez les adolescent(e)s de la cinquième secondaire.

⁷ Une fois par semaine pendant au moins un mois.

⁸ Cinq consommations standards au cours d'une même occasion.

⁹ Chiffres émanant des jeunes de la première à la cinquième secondaire, d'environ 12 à 17 ans.

¹⁰ Plus de cinq occasions.

Tableau VI : Consommation régulière ou quotidienne d'alcool au Québec (écoles secondaires)

Consommation régulière ou quotidienne d'alcool (%)				
	2000	2004	2008	2013
Total	20,3	16,8	14,5	9,9
Garçons	22,4	18,8	16,2	10,8
Filles	18,2	14,7	12,7	8,9
Consommation excessive				
Total	45,6	43,1	40,0	33,7
Troisième secondaire	51,3	40,4	39,4	35,5
Quatrième secondaire	54,9	59,5	58,9	47,1
Cinquième secondaire	70,9	70,8	68,3	63,3

(Traoré et al., 2014)

Au chapitre de la consommation de drogues, les chiffres obtenus par l'ISQ confirment que la substance la plus consommée par les adolescent(e)s est le cannabis, avec un taux de prévalence de 22,9%. Pour toutes les autres drogues illégales, le taux demeure en deçà de 4,2%. La diminution observée pour l'alcool s'observe aussi pour les drogues comme le démontrent les chiffres du Tableau VII.

Tableau VII : Consommation de drogues illégales au Québec (écoles secondaires)

Consommation de drogues illégales (12 mois) %				
Année	2000	2004	2008	2013
Première secondaire	18,6	16,0	8,3	5,5
Deuxième secondaire	39,1	29,9	18,8	14,8
Troisième secondaire	50,8	37,7	29,8	26,6
Quatrième secondaire	49,3	50,4	38,7	33,0
Cinquième secondaire	61,7	59,3	47,3	44,0

(Traoré et al., 2014)

Par contre, le taux de consommation *régulière*¹¹ de drogues est élevé chez les jeunes ayant consommé au cours de la dernière année, comme le démontre le Tableau VIII.

¹¹ Les fins de semaine ou une à deux fois par semaine, mais moins de sept jours par semaine.

Tableau VIII : Consommation régulière de drogue au Québec (écoles secondaires)

Consommation régulière de drogues		
	Ensemble des élèves	Élèves ayant déjà consommé de la drogue (24,1%)
Garçons	11,2	42,5
Filles	9,5	41,8
Première secondaire	1,8	38,8
Deuxième secondaire	5,7	39,6
Troisième secondaire	10,9	40,7
Quatrième secondaire	13,9	40,3
Cinquième secondaire	21,1	45,9

(Traoré et al., 2014)

Précisons que 12,4% et 3,4% des jeunes de la cinquième secondaire sont respectivement des consommateurs réguliers et quotidiens, toutes drogues confondues. Pour l'ensemble des jeunes, de la première à la cinquième secondaire, on trouve 6,4% de consommateurs réguliers et 1,4% de quotidiens.

Sous l'angle de la santé mentale, un éclairage est apporté par l'Enquête québécoise sur la santé des jeunes du secondaire 2010-2011 (Pica et al., 2013) qui s'appuie sur l'indice de détresse psychologique de Santé Québec. Cet indice est basé sur 14 questions mesurant la fréquence des symptômes liés à la dépression, à l'anxiété, à l'irritabilité et aux troubles cognitifs. On retrouve un niveau élevé de détresse psychologique chez 13,6% des garçons et 28,2% des filles. À ce niveau, une différence peut être notée entre les jeunes de la cinquième secondaire (24,7%) versus ceux de la première secondaire (15,5%). Voici les troubles confirmés par un médecin pour ces jeunes :

- Anxiété : 8,6 % (fille : 11,0% ; garçon : 6,2%)
- Dépression : 4,9% (fille : 5,9% ; garçon : 3,9%)
- Troubles alimentaires : 1,8% (fille : 2,5% ; garçon : 1,1%)
- TDA/H : 12,6 % (fille : 9,3% ; garçon : 15,9%)

Ce survol des taux de prévalence et de comorbidité montre bien l'importance que revêt l'évaluation permettant de maximiser l'efficacité des interventions auprès de ces jeunes, de les orienter vers des traitements efficaces, d'éviter une détérioration de leur état et de leur fonctionnement. La sous-section qui suit met en perspective les éléments importants de cette évaluation.

6. L'évaluation

En 2001, Santé Canada (Currie), soulignait la nécessité d'effectuer une évaluation globale à partir d'expertises multiples. Plus récemment encore, le Centre canadien de lutte contre les toxicomanies (CCLT) (Comité directeur du projet collaboratif en toxicomanie et en santé mentale, 2014 :27) évoquait la nécessité d'une collaboration au long cours entre fournisseurs de soins, plus précisément d'«une démarche collaborative longitudinale à l'égard de l'évaluation des troubles cooccurrents complexes».

Pour acquérir une meilleure compréhension de la chronologie des symptômes, l'échange d'informations entre fournisseurs de soins devient incontournable. Dans le domaine de la consommation de substances psychoactives, on souligne depuis un certain temps la nécessité d'évaluer les problématiques associées à l'aide d'instruments diagnostiques structurés (Bertrand et al., 2004). Cela permettrait de mettre en place des traitements plus appropriés et individualisés, d'autant plus qu'on sait que les adolescent(e)s présentant des troubles internalisés autant qu'externalisés ont des difficultés plus marquées 12 mois après le traitement.

De l'évaluation découle logiquement l'appariement, ce qui signifie que l'individu puisse recevoir le bon traitement (meilleures pratiques), au bon moment (disponibilité physique, psychologique, etc.) et au bon endroit (niveau de soin requis). Brown et Wood (2002) qui ont analysé les données de l'étude MATCH (*Matching Alcoholism Treatment to Client Heterogeneity*) proposent des conclusions qui sont éclairantes, même si leur étude a été effectuée auprès d'une clientèle adulte. Le projet MATCH a été l'une des plus importantes études aux États-Unis, portant sur les effets différentiels des traitements et sur l'efficacité du traitement en général. Ses résultats montrent que certaines approches mènent à de meilleurs

résultats que l'absence de traitement, mais qu'en bout de piste, aucune ne démontre de meilleurs résultats pour l'ensemble des usagers. Et en ce qui concerne l'efficacité du traitement et du pronostic des troubles de SPA ? La réponse apportée est « ... un « traitement fructueux » ne survient qu'un peu plus fréquemment que la rémission spontanée en l'absence de tout traitement. » (Brown et Wood, 2002 : 8). Cette importante recherche tend à conclure que l'appariement est une notion incontournable.

Chaque fois qu'il est question des caractéristiques particulières des jeunes, que ce soit en terme développemental ou de besoins particuliers (jeunes de la rue, jeunes itinérants, jeunes avec des problèmes concomitants ou judiciaires), la question de l'appariement se pose. En son absence, on constate ce que Charles et Alexander (2007) nomment « la tendance aux programmes uniformisés (*one size fits all*) », soit des interventions faites sans égard aux besoins spécifiques des jeunes et de leur famille.

Laventure (2008) rappelle la nécessité d'évaluer les besoins des jeunes garçons et filles affichant à la fois des TC et une consommation de SPA pour mieux les orienter. Sans élaborer, on peut mentionner que l'appariement doit se faire en fonction des différences entre individus, du genre ou de l'âge, ce qui est cohérent avec les principes de la PPD.

Nombre d'études ont été réalisées sur le traitement des adultes consommateurs, leurs besoins particuliers et les approches efficaces (Bertrand et al., 2004). Bien moins de travaux l'ont été sur les traitements (Chassin et al., 2010; Leslie, 2008; Morral, McCaffrey, Ridgeway, Mukherji et Beighley, 2006; Vaughn et Howard, 2004), leur efficacité (Brochu, 2007; Morral et al., 2006) et les besoins spécifiques des jeunes (Charles et Alexander, 2007; Comité permanent de lutte à la toxicomanie (CPLT), 2001).

La notion d'appariement considérée dans sa globalité et appliquée à la PPD, consiste à dire que les traitements pour les adultes ne conviennent pas nécessairement aux jeunes, pas plus que les instruments d'évaluation pour adultes ne sont nécessairement adaptés pour eux.

Au cours des dernières années, dans le réseau des CRD du Québec, les critères de l'ASAM¹², provenant des États-Unis, ont fait leur apparition dans les plans d'intervention des usagers, tant adultes qu'adolescent(e)s. Or, ces critères sont au cœur même de l'évaluation et de l'appariement. Dans leur troisième version (Mee-Lee, Shulman, Fishman, Gastfriend et Miller, 2013), ils visent les troubles de SPA et les conditions concomitantes qui leur sont associées. Reposant sur l'expertise des cliniciens, ces lignes directrices permettent d'organiser et d'analyser les éléments recueillis lors de l'évaluation pour ensuite orienter les personnes selon leur niveau de besoin et les soins qu'ils requièrent. Deux critères sont évalués : le premier (A) concerne la consommation de SPA et le second (B) six dimensions couvrant les sphères de vie de l'utilisateur (intoxication aiguë/risque de sevrage, condition biomédicale, condition émotionnelle/psychologique/comportementale, motivation au changement, potentiel de rechute/poursuite de la consommation et environnement). Peu importe l'instrument d'évaluation utilisé, cette grille d'analyse sert à prioriser et orienter selon les besoins propres et spécifiques de l'adulte, autant que de l'adolescent(e). Autant à l'échelle nord-américaine que québécoise, un débat (Chauvet, Kamgang, Ngamini Ngui, Fleury et Simoneau, 2015) porte actuellement sur l'approche de l'ASAM qui définit les TUS comme suit (American Society of Addiction Medicine (ASAM), 2001):

« Addiction^{xi} is a primary, chronic disease of brain reward, motivation, memory and related circuitry. Dysfunction in these circuits leads to characteristic biological, psychological, social and spiritual manifestations. This is reflected in an individual pathologically pursuing reward and/or relief by substance use and other behaviors »

Selon Chauvet et ses collaborateurs (2015) , les tenants de l'approche médicale estiment que cette nouvelle définition était nécessaire, mais d'autres lui reprochent de réduire les TUS à un « dysfonctionnement du cerveau » et arguent qu'on privilégie les aspects biologiques au détriment de la part du psychosocial. Au Québec, où les TUS sont généralement considérés comme un phénomène multidimensionnel, qu'il faut aborder avec une approche biopsychosociale, la définition de l'ASAM a suscité des questionnements. Par exemple, Landry

¹² American Society of Addiction Medicine

et Brochu (2012) craignent un retour vers le modèle médical et une explication neurobiologique des TUS. Ils appellent à une réflexion approfondie des enjeux soulevés par la définition de l'ASAM, autant au niveau scientifique que clinique. Entre temps, les intervenants du réseau de CRD du Québec apprivoisent les critères de l'ASAM dans leur travail journalier d'évaluation des besoins des individus se présentant à leur porte.

6.1 Les instruments d'évaluation

Puisque plusieurs formes d'évaluation existent, il importe de distinguer la détection de l'évaluation proprement dite. La première consiste à identifier si un adolescent(e) a besoin de certains services. À partir de sa symptomatologie, il peut s'agir d'évaluer le niveau d'urgence (suicidaire par exemple), le développement d'un trouble ou la présence d'un trouble avéré (dépression ou consommation). La détection est d'ordinaire un exercice de courte durée auquel on peut se prêter avec tou(te)s les adolescent(e)s. Le résultat peut indiquer la nécessité d'une évaluation plus complète (multidimensionnelle) ou plus ciblée (un trouble en particulier) ou encore la nécessité d'une intervention immédiate sinon à court terme.

Le présent projet s'intéresse plutôt aux évaluations dites complètes (*comprehensive assessment instrument*). Elles contrastent avec les évaluations qui ne cernent qu'un type de symptomatologie (troubles dépressifs par exemple). Voici ce qu'en disent Winters, Stinchfield et Bulkstein (2008 : 71) en lien avec les TUS :

« These instruments yield information that can more definitively assess the nature and severity of the drug involvement, to assign a substance use disorder, and to identify the psychosocial factors that may predispose, perpetuate and maintain the drug involvement. »

Les auteurs identifient trois groupes d'évaluations de ce genre : les questionnaires multiéchelles (*multiscales questionnaire*), les entrevues orientées vers les problèmes (*problem-focused interviews*) et les entrevues diagnostiques (*diagnostic interviews*). Les questionnaires multiéchelles sont d'ordinaire autoadministrés. Variant en nombre de questions, donc en temps de passation, ils se subdivisent en plusieurs échelles. Ils couvrent nombre d'aspects sur un sujet donné, incluant la sévérité et les facteurs de risque et ils ont la capacité de détecter certaines distorsions dans les réponses fournies par le sujet. Les

entrevues orientées vers les problèmes mesurent un certain nombre de variables en lien avec une problématique spécifique et ont la capacité d'en évaluer la sévérité. Quant aux entrevues diagnostiques, elles sont structurées ou semi structurées et construites pour produire des diagnostics en lien avec le DSM. Dans le cas du présent projet, ces dernières sont visées, puisque les instruments retenus doivent être en mesure de produire certains diagnostics spécifiques (voir le chapitre Méthodologie).

6.1.1 Entrevue basée sur l'intervieweur ou le participant

L'apparition des entrevues diagnostiques structurées fait suite au constat qu'avec une même formation, dans une même discipline et parfois dans un même établissement, les cliniciens sont souvent incapables d'en arriver aux mêmes conclusions devant des informations identiques (Angold et Fisher, 1999). Il s'agit essentiellement de créer des instruments qui permettent de contrôler différents biais. Angold et Fisher (1999) en répertorient cinq en lien avec l'intervieweur : tendance à poser le diagnostic avant d'avoir obtenu toute l'information pertinente, tendance à ne recueillir que l'information confirmant ce diagnostic, tendance à ignorer les informations qui infirment ce diagnostic, tendance à combiner l'information de façon idiosyncrasique et finalement, tendance à baser son jugement sur les patrons cognitifs les plus probables. Un des documents de formation à l'instrument d'évaluation GAIN (Association des centres de réadaptation du Québec (ACRDQ) et Chestnut Health Systems (GAIN Coordinating Center), 2012) regroupe en trois catégories les biais pouvant être reliés à un instrument d'évaluation :

- Interprétatif (cognitif) : survient lorsque l'interprétation des données cliniques est influencée par les expériences, allégeances et croyances des intervieweurs. L'éducation ou la formation ne parvient généralement pas à éliminer ce type de biais. Une fois qu'on est sensibilisé à cet état de fait, des mesures correctives et compensatoires appropriées peuvent être mises en place.
- Instrumental : survient lorsque l'instrument utilisé a le potentiel d'engendrer des erreurs d'ordre culturel, c'est-à-dire qu'il est non sensible aux particularités ethnoculturelles.

- Structurel : survient lorsque le processus montre une tendance inhérente à favoriser des résultats particuliers.

Pour tenter de contrôler ces biais et obtenir les diagnostics les plus objectifs possible, les caractéristiques de toute entrevue diagnostique structurée ont été définies (Angold et Fisher, 1999) :

- Mettre en place une structure permettant de collecter toute l'information pertinente auprès de tous les acteurs pertinents.
- Définir comment cette information devrait être collectée.
- Poser un diagnostic seulement une fois que toute l'information qui pourrait confirmer ou infirmer ce diagnostic a été obtenue.
- Structurer le processus de traitement de l'information confirmant ou infirmant le diagnostic final.

Outre ces quatre caractéristiques, deux structures distinctes existent : l'une « centrée » sur l'intervieweur (*interviewer-based*), l'autre sur le participant (*respondent-based*). Les instruments centrés sur l'intervieweur soutiennent ce dernier, en lui servant de guide pour déterminer si les symptômes sont présents. Son expertise et ses connaissances sont nécessaires pour décider, sur la base des informations fournies par le participant, si le symptôme exploré répond à la définition et peut être considéré comme « présent ». Les instruments basés sur le participant ont fait leur apparition suite à la constatation que les instruments «intervieweurs» pouvaient difficilement être utilisés par des gens sans expérience clinique (*lay-interviewers*), comme c'est souvent le cas dans les études épidémiologiques de grande envergure. Les stratégies de sondage politique ou de marketing furent alors mises à profit : l'intervieweur n'a pas de décision à prendre sur la présence d'un symptôme. Il n'a qu'à poser les questions telles quelles. Ainsi on minimise la variance d'erreur due à l'intervieweur.

6.1.2 Souple, semi-structurée, structurée

La structure de l'entrevue varie aussi selon son « degré de rigidité » : souple, semi-structurée ou structurée. L'évaluation souple comprend notamment l'entrevue clinique au cours de laquelle l'intervenant, généralement un clinicien, n'a aucun plan précis à suivre et permet à l'usager de « diriger » en grande partie un entretien interactif. L'intervieweur ne dispose d'aucune directive, restriction, question prédéterminée ou liste d'options. Il pose quelques questions de nature générale pour engager une discussion ouverte et spontanée avec le participant. Il peut poser des questions supplémentaires pour approfondir certains sujets, explorer des contradictions et recueillir des renseignements plus détaillés.

Les évaluations rigoureusement structurées sont régies par un protocole strict qui ne laisse place à aucun écart ni improvisation et qui doit être appliqué de façon conforme. Aucun dialogue n'est possible avec le participant qui doit tenter de répondre aux questions au meilleur de ses connaissances et avec peu d'explications supplémentaires. L'intervieweur pose une série de questions déterminées à l'avance sur des sujets particuliers, dans un ordre précis. Les participants choisissent les réponses parmi une liste d'options. Les tests d'intelligence, tels les échelles de Wechsler (Wechsler, 1995) en sont un exemple.

Entre ces deux options se trouve l'évaluation semi-structurée. L'intervieweur y pose une série de questions prédéterminées auxquelles les gens répondent dans leurs propres mots. Toutefois, des choix de réponse sont souvent fournis pour leur faciliter la tâche. Certains intervieweurs utilisent un guide tel qu'une liste de contrôle pour veiller à ce que tous les participants fournissent des renseignements sur les mêmes sujets. L'intervieweur peut approfondir certains points ou poser des questions supplémentaires dans le but de clarifier les réponses obtenues. Les entrevues semi-structurées sont utilisées dans l'optique de recueillir une information détaillée de manière systématique auprès de nombreux participants ou interviewés. Elles sont basées sur l'intervieweur.

Les instruments d'évaluation tentent de cerner les comportements, états, traits, symptômes et/ou problèmes psychosociaux, permettant de poser un diagnostic. Plusieurs moyens et mesures ont été instaurés pour contrôler la variance due à l'intervieweur, à l'instrument lui-

même et à l'interviewé. Mais encore faut-il que ces instruments mesurent bien ce qu'ils sont censés mesurer, et ce de façon stable dans le temps. Les principales qualités psychométriques d'un instrument d'évaluation sont la validité et la fidélité qui sont présentées dans la sous-section suivante.

6.1.3 Qualité psychométrique : fidélité et validité

La psychométrie s'intéresse à l'élaboration et la validation des techniques de mesure utilisées en psychologie et dans les sciences sociales. Bernier et Pietrulewicz (1997) rappellent que Legendre (1993, cité dans Bernier et Pietrulewicz : VII) l'a définie comme étant « [...] la mesure et le traitement statistique des faits psychiques, principalement par la méthode des tests ». En d'autres mots, la psychométrie passe par la mesure des phénomènes et attributs psychiques des individus, mais déborde le cadre des tests et de leur construction.

Ce sont donc à l'aide de ces concepts de validité et fidélité, qui se déclinent en plusieurs sous-types, que les études, portant sur les instruments d'évaluation de la comorbidité, seront scrutées.

L'étude de la validité se centre sur les erreurs systématiques. Elle concerne donc la qualité de ce qui est mesuré. On la définit comme étant la portion de la variance vraie qui est pertinente aux buts de l'utilisation d'un instrument. Selon Legendre (1993, cité dans Bernier et Pietrulewicz, 1997 : 179), la validité est « la capacité d'un instrument à mesurer réellement ce qu'il doit mesurer, selon l'utilisation que l'on veut en faire ».

Le concept de fidélité pour sa part, renvoie à la constance des scores obtenus à l'aide d'un instrument (Bernier et Pietrulewicz, 1997). Les mesures psychologiques sont reconnues, comparativement aux mesures physiques, comme moins stables, parce qu'elles mesurent des comportements humains qui peuvent varier d'une fois à l'autre et d'une personne à l'autre. La fidélité d'un instrument renvoie à la capacité de ce dernier de mesurer un item en arrivant à un même résultat d'une fois à l'autre. Elle concerne la précision avec laquelle un test mesure certaines caractéristiques. Selon ces mêmes auteurs, on définit la fidélité comme un rapport de variance. L'étude de la fidélité concerne les erreurs dues au hasard ou aléatoires et pose la

question suivante : mesure-t-on, d'une fois à l'autre, la même chose, et ce de façon cohérente ?

Il existe plusieurs types de validité et de fidélité. Ceux utilisés pour mettre à l'épreuve les instruments d'évaluation dans les études recensées par la SS seront présentés au chapitre de présentation des résultats. On y recensera également les différentes mesures utilisées pour tester ces concepts.

7. La synthèse systématique

Dans le domaine des sciences sociales, les synthèses systématiques (SS) portent souvent sur l'efficacité des traitements ou interventions psychosociales. Elles identifient, analysent et résument les résultats de toutes les recherches publiées durant un certain nombre d'années et visent à faciliter ainsi la prise de décisions des décideurs, administrateurs et cliniciens. La SS diffère des recensions narratives par son approche scientifique, objective et transparente qui vise à minimiser les biais. Elle cherche à faire état des connaissances réelles et non de la vision qu'en ont les auteurs. La SS fournit des réponses empiriques à des questions précises, tout en identifiant ce qui est connu et ce qui reste à découvrir pour orienter la recherche future (Khan, Riet, Glanville, Sowden et Kleijnen, 2001).

Un des buts principaux de la SS dans plusieurs domaines des sciences sociales est d'identifier les meilleures connaissances disponibles face à une intervention ou un instrument. Dans le cas du présent projet, il s'agira de la validité d'instruments d'évaluation de la comorbidité présentée par les adolescent(e)s. Un autre but de la SS est d'identifier les limites des connaissances actuelles et ainsi d'orienter les études subséquentes ou proposer des améliorations possibles au niveau de la méthodologie. Farrington et Petrosino (2001) mettent l'emphase sur le fait que les SS contiennent une description plus explicite de leur méthodologie, une revue de littérature plus extensive et « internationale », procurent plus d'informations sur les critères d'inclusion et d'exclusion et sur les sources de données. Elles visent, ultimement à procurer de l'information de qualité à ceux qui en ont besoin (cliniciens, décideurs, administrateurs et même grand public). Pour Khan, Riet, Glanville, Sowden, Amanda et Kleijnen (2001), le but ultime est la dissémination de l'information ainsi générée.

Le manuel de la Collaboration Cochrane sur les SS en regard de l'intervention propose la définition suivante (Higgins et Green, 2009) :

...a clearly stated set of objectives with pre-defined eligibility criteria for studies;

an explicit, reproducible methodology;

a systematic search that attempts to identify all studies that would meet the eligibility criteria;

an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias; and

a systematic presentation, and synthesis, of the characteristics and findings of the included studies.

La présente SS ne prétend pas être en mesure d'atteindre ces objectifs pleinement, mais pourrait à tout le moins déblayer le terrain en regard des instruments d'évaluation de la comorbidité disponibles et leur validité et fidélité, en regard d'éléments bien précis.

En l'absence de consensus, la SS s'avère tout indiquée puisqu'elle permet d'identifier ce qui a déjà été publié sur un sujet donné et ensuite de se prononcer sur l'efficacité d'un traitement, ou, comme dans le cas présent, sur les qualités psychométriques d'instruments d'évaluation.

8. Problématique

Plusieurs travaux traitent de la comorbidité chez les adultes, mais peu chez les adolescent(e)s (Drabick et Kendall, 2010). Par ailleurs, la revue de littérature de Saban et Flisher (2010) rapporte que la plupart des études ciblent, soit un trouble en particulier ou une substance en particulier. Peu d'instruments sont validés auprès des adolescent(e)s, que ce soit pour évaluer leur santé mentale ou leur consommation (Perepletchikova et al., 2008). On peut donc conclure que les connaissances actuelles sont limitées tant au niveau de la comorbidité chez les adolescent(e)s que de la manière de bien l'évaluer.

Aucune recommandation ou ligne directrice clinique ne se trouve sur le site de l'Académie canadienne de psychiatrie de l'enfant et de l'adolescent (2011) en regard de l'évaluation des troubles comorbides ou de troubles mentaux spécifiques. Il n'y en a pas davantage sur le site

de la revue de l'Association canadienne de psychiatrie (2011). Le site de l'*American Academy of Child and Adolescent Psychiatry* (2011) propose plusieurs directives cliniques (*clinical guidelines*), mais chacune porte sur l'évaluation d'un trouble en particulier et non de la comorbidité. Le site de la *National Institute on Drug Abuse* présente des recommandations sur l'évaluation des troubles comorbides (2010) et précise l'importance d'une évaluation multidimensionnelle, mais sans qu'il n'y ait d'avis sur tel ou tel instrument précisément. Pour sa part, le *Center for Substance Abuse Treatment* de la *Substance Abuse and Mental Health Services Administration* a publié dans son *Treatment Improvement Protocol no 42* (2005), des recommandations portant sur l'évaluation des troubles comorbides, mais un seul des instruments proposés est à la fois multidimensionnel et recommandé avec les adolescent(e)s et il s'agit du GAIN. S'il semble y avoir consensus sur la nécessité d'instruments « complets » (*comprehensive*) et multidimensionnels, peu d'exemples sont fournis.

La première question posée par cette SS porte donc sur la disponibilité d'instrument d'évaluation des troubles comorbides, incluant un TUS, conçus pour les adolescent(e)s.

Comme il semble y avoir un consensus sur la nécessité d'utiliser des instruments qui soient à la fois multidimensionnels et validés, mais que la littérature ne les identifie pas clairement, la seconde question posée par ce mémoire de maîtrise porte sur la fidélité et la validité des différentes méthodes possibles et ce, en lien avec certains des diagnostics les plus prévalents chez les adolescent(e)s.

L'objectif principal de la présente recension consiste donc à répertorier les instruments d'évaluation multidimensionnelle des troubles comorbides de l'Axe 1 du DSM-IV-TR¹³, incluant les diagnostics liés à un TUS et ce, dans un même instrument, qui soit conçu pour les adolescents et les adolescentes.

Le deuxième objectif est d'identifier les études qui ont mis à l'épreuve ces instruments sous l'angle de leur fidélité et validité. Le plus grand nombre possible d'études répondant à

¹³ La plupart des versions actuelles des instruments ont été ajustées à la version du DSM-IV. Par contre, certaines études remontant avant sa parution, porte naturellement sur des diagnostics du DSM-III ou DSM-III-R.

certains critères d'inclusion et d'exclusion (voir chapitre Méthodologie) devra être rassemblé. Évidemment, il y a des limites à la possibilité d'être exhaustif, entre autres la langue dans laquelle sont publiées ces études (voir chapitre Méthodologie).

En date du 10 mars 2011, les sites de la Collaboration Cochrane (2011a, 2011b, 2011c) et de la Collaboration Campbell (2011) ont été vérifiés. À ce moment, il ne semblait y avoir aucune recension systématique sur la validité des instruments d'évaluation des troubles comorbides chez les adolescent(e)s. Pas plus qu'il ne semblait y avoir de synthèse répertoriant tout simplement les instruments disponibles. Plusieurs ouvrages (Center for Substance Abuse Treatment, 1999, 2005; Grisso et al., 2005; Kaminer et Bukstein, 2008; National Institute on Drug Abuse (NIDA), 2010; Samet, Waxman, Hatzenbuehler et Hasin, 2007) en identifient selon les besoins d'une population particulière, mais aucun ne les répertorie tous. Les instruments conçus d'abord pour les adolescent(e)s seront ciblés. S'y ajouteront ceux conçus au départ pour les adultes, puis ajustés et validés auprès des jeunes.

Un troisième objectif est d'évaluer la rigueur méthodologique des études ayant mis à l'épreuve ces instruments. Les critères de qualité méthodologique sont depuis longtemps au cœur des discussions entre les membres du *Campbell Crime and Justice Coordinating Group* (Farrington et Petrosino, 2001) : quel est le protocole de recherche ? Quel est l'échantillon ? Quels étaient les résultats attendus et quels sont ceux obtenus ? C'est pour tenter de répondre à ces questions que la grille d'analyse QUADAS-2 (Whiting et al., 2011), conçue spécifiquement pour évaluer la qualité méthodologique des études portant sur l'exactitude des diagnostics, sera utilisée (présentée au chapitre Méthodologie). Si le propre des SS est la transparence de la méthode dans la sélection des études, des éléments bien précis doivent être présents dans celles qui sont retenues. La grille QUADAS-2 est ainsi tout indiquée, car le principe même de la SS suppose de la rigueur et la capacité d'identifier les études robustes pour en tirer des conclusions généralisables.

Chapitre 2 : Méthodologie

Ce chapitre présente la méthode utilisée pour réaliser la présente recherche, soit la synthèse systématique et les détails de son utilisation. D'abord en lien avec l'identification des instruments d'évaluation, ensuite avec les études les ayant mis à l'épreuve en regard de leur validité ou fidélité. La méthode d'analyse de ces dernières y est également présentée, soit la grille du QUADAS-2.

1. La synthèse systématique

Comme il fut annoncé dans la recension des écrits, la méthode utilisée pour atteindre les objectifs poursuivis par ce mémoire de maîtrise est la synthèse systématique (SS) des écrits. Dans un premier temps, cette méthode permettra de répertorier les instruments d'évaluation multidimensionnelle existants. Dans un second temps, elle sera utilisée pour identifier les écrits portant sur la validité ainsi que la fidélité de chacun des instruments retenus. Une analyse de la méthodologie des études ayant mis à l'épreuve la fidélité ou la validité de certaines échelles retenues de ces instruments sera par la suite effectuée à l'aide de la grille du QUADAS-2 (Whiting et al., 2011).

2. Stratégie de recherche

2.1 L'identification des instruments d'évaluation

La première étape de la présente SS a été d'identifier les instruments d'évaluation existants tout en départageant les outils de détection ou de repérage pour ne conserver que les instruments d'évaluation multidimensionnelle.

2.1.1 Identification des bases de données

Plusieurs bases de données ont été interrogées pour identifier les instruments susceptibles de répondre aux critères d'inclusion qui seront présentés ultérieurement. Les plateformes OVID, WEB of Science, ainsi que Proquest, ont été utilisées. Dans le cadre d'un SS, la pertinence de la littérature « grise » (*grey littérature*), c'est-à-dire non publiée dans des livres ou périodiques scientifiques, ne fait généralement aucun doute (il s'agit entre autres des résumés de conférence). Par contre, dans le cas présent, aucune recherche n'a été effectuée en ce sens par manque de temps. Sans constituer la majorité des textes constituant le corpus

d'une SS, le fait de ne pas considérer ces écrits pourrait en affecter les résultats (Higgins et Green, 2009) ce qui sera pris en considération au moment de l'analyse des données.

Pour tenter d'identifier les instruments existants, en utilisant la plateforme OVID, les bases de données suivantes ont été explorées : PSYINFO, EMBASE, MEDLINE, CENTRAL (*The Cochrane Central Register of Controlled Trials*), EBM Reviews – Health Technology assessment et Health and Psychosocial Instruments. Les mots-clés utilisés pour cette première recherche furent : *assessment, evaluation, diagnostic, instrument, appraisal, scale, psychometric, psychiatric, mental-health, mental disorder, comorbidity, substance use, dependence, youth, adolescent, teen, validity, reliability*.

Les mots-clés suivant furent combinés, en y incluant la troncature (*) : «*assessment OR evaluation OR diagnostic OR instrument OR appraisal OR scale*» AND «*psychiatric OR mental health OR mental disorder OR comorbidity*» AND «*psychometric OR validity OR reliability*» AND «*teen OR youth OR adolescent*» AND «*substance use*».

Il existe des centaines d'instruments d'évaluation des problèmes de santé mentale. Un premier tri fut donc effectué à partir de la description générale des instruments. Ceux destinés aux adolescent(e)s et décrits comme des évaluations complètes (*comprehensive assessment*) en santé mentale furent retenus. Au terme de ce travail, 26 instruments furent répertoriés. La liste fut complétée au fil des lectures. Au final, le nombre d'instruments retenus fut de 34.

Les publications relatives à chacun de ces instruments furent explorées à l'aide de Google et Google Scholar afin d'obtenir des informations de base sur leurs composantes telles que : le nombre de questions posées, l'âge du public cible, le temps d'administration, les construits mesurés (les échelles) et les diagnostics produits.

Ces informations permirent d'identifier les composantes présentes dans les différents instruments. Or, autant il existe un grand nombre d'instruments d'évaluation différents, autant ils diffèrent les uns des autres en regard de leurs composantes. Sous l'angle des possibilités de comparaison, la grande disparité des items et des échelles contenus dans ces instruments posait problème.

L'objectif principal du présent projet est de comparer la validité et la fidélité d'instruments d'évaluation, mais pour ce faire, il faut être en mesure de comparer des échelles semblables. C'est pourquoi l'identification d'échelles communes fut une étape incontournable et que les diagnostics générés par l'instrument furent analysés. Pour choisir ces diagnostics sur la base de leur pertinence, un retour vers la recension des écrits fut d'abord effectué. Rappelons que les diagnostics de l'Axe I (American Psychiatric Association, 1994) les plus souvent posés chez la clientèle visée sont les suivants (voir la sous-section Taux de prévalence chez les adolescents et adolescentes au chapitre 1) :

- Troubles liés à l'utilisation des substances (TUS)
- Troubles de l'humeur (dépression)
- Troubles anxieux (anxiété généralisée et ÉSPT)
- Troubles des conduites (TC)
- Trouble du déficit de l'attention /hyperactivité (TDA/H)

Les diagnostics précis qui furent retenus au terme de la revue de littérature sont :

- Troubles liés à l'utilisation des substances (TUS) (obligatoire)
- Trouble d'anxiété généralisée (TAG)
- Épisode dépressif majeur (ÉDM)
- Troubles des conduites (TC)
- Trouble du déficit de l'attention /hyperactivité (TDA/H)
- État de stress post-traumatique (ÉSPT)

Les instruments répertoriés furent revus sous l'angle suivant : chacun devait être en mesure d'évaluer et de fournir au moins cinq des six diagnostics retenus, dont celui des TUS.

2.1.2 Les critères d'inclusion et d'exclusion des instruments d'évaluation

Les critères d'inclusion et d'exclusion, qui sont d'une importance capitale dans une SS (Higgins et Green, 2009), furent établis de la manière suivante :

Inclusion

- Population : adolescent(e)s 13-14 à 17-18 ans.
- Instrument créé spécifiquement pour les adolescent(e)s.
- Évaluation multidimensionnelle.
- Instrument générant au moins cinq des six diagnostics retenus au terme de la recension des écrits, incluant obligatoirement celui d'abus/dépendances aux SPA.

Exclusion

- Population : enfant jusqu'à 12-13 ans et jeunes adultes de 18 et plus.
- Instruments de repérage et de détection.
- Instruments conçus pour la recherche exclusivement (sans possibilité de poser un diagnostic).

Au terme de cette démarche, 11 instruments furent retenus. Ils sont présentés dans le Tableau IX. Cette liste fut régulièrement revue et corrigée au fil de la recherche. Chaque instrument sera présenté sommairement au chapitre 3 (Résultats).

Tableau IX : Instruments d'évaluation retenus

No	Titre	Abr.	Auteurs/ année/	Pays	Particularité	Âge des jeunes	Identification Dx recherchés
1	Adolescent Diagnostic Interview	ADI	Winters, K.C et Henly, G.A., 1993	EU	Cible TUS. Entrevue structurée. Passation possible par un personnel paraclinique ¹⁴ .	12-18	TUS-TAG-ÉDM-TC-TDA/H

¹⁴ Fait référence aux gens sans expérience clinique (*lay-interviewers*).

2	Adolescent Psychopathology Scale	APS	M. Reynolds, W., 1998	EU	Évalue la présence et sévérité des Sx ¹⁵ psychopathologiques. Instrument d'auto-évaluation. Interprétation des résultats par des cliniciens expérimentés.	12-19	TUS-TAG-ÉDM-TC-TDA/H-ÉSPT
3	The Child and Adolescent Psychiatric Assessment	CAPA	Angold, A. et ses collaborateurs, 1995	EU-Ang.	Cible un large éventail de troubles. Entrevue basée sur l'intervieweur. Passation par des cliniciens expérimentés.	9-17	TUS-TAG-ÉDM-TC-TDA/H-ÉSPT
4	Children's Interview for Psychiatric Syndromes	ChIPS	Weller, E.B. et ses collaborateurs, 1999	EU	Cible un large éventail de troubles. Entrevue diagnostique structurée. Passation par un personnel paraclinique formé.	6-18	TUS-TAG-ÉDM-TC-TDA/H-ÉSPT
5	Diagnostic Interview for Children and Adolescents	DICA	Herjanic, B. et Reich W., 1982; Reich, W., 2000	EU	Cible un large éventail de troubles. Entrevue diagnostique structurée. Passation par un personnel paraclinique formé.	6-17	TUS-TAG-ÉDM-TC-TDA/H-ÉSPT
6	Diagnostic Interview Schedule for Children-IV	DISC	Costello, A.J., et ses collaborateurs, 1984 ; Shaffer, D. et ses collaborateurs 1993, 2000 ; Jensen, P. et ses collaborateurs, 1995	EU	Cible un large éventail de troubles. Entrevue diagnostique structurée. Passation par des non cliniciens formés.	9-17	TUS-TAG-ÉDM-TC-TDA/H-ÉSPT
7	Global Appraisal of Individual Needs	GAIN	Dennis, M. et ses collaborateurs, 1993	EU	Cible l'abus-dépendance de SPA. Entrevue diagnostique semi-structurée. Passation par des cliniciens formés.	Adolescent(e)-adulte	TUS-TAG-ÉDM-TC-TDA/H-ÉSPT

¹⁵ Symptômes.

8	Interview Schedule for Children and Adolescents	ISCA	Kovacs, M., 1985	EU	Cible : la population normale dans le cadre d'études épidémiologiques. Entrevue diagnostique semi-structurée. Passation par des cliniciens expérimentés.	8-17	TUS-TAG-ÉDM-TC-TDA/H
9	Schedule of Affective Disorders and Schizophrenia for Children	K-SADS	Chambers, W.J., et ses collaborateurs, 1985 ; Kaufman, J. et ses collaborateurs, 1997	EU	Cible un large éventail de troubles. Entrevue diagnostique semi-structurée. Passation par des cliniciens expérimentés.	6-17	TUS-TAG-ÉDM-TC-TDA/H
10	Practical Adolescent Dual Diagnostic Interview	PADDI	Estroff, T.W. et Hoffmann, N.G., 2001	EU	Cible l'abus-dépendance de SPA et d'autres troubles à prévalence élevée. Entrevue diagnostique structurée. Passation par des cliniciens.	13-18	TUS-TAG-ÉDM-TDA-ÉSPT
11	Personality Assessment Inventory-Adolescent	PAI-A	Morey, L. C., 2007	EU	Cible un bon nombre de troubles. Instrument d'auto-évaluation. Interprétation par des cliniciens.	12-18	TUS-TAG-ÉDM-TC-ÉSPT

TUS : trouble lié à l'utilisation de substances psychoactives; TAG : trouble d'anxiété généralisé; EDM : épisode dépressif majeur; TC : trouble des conduites; TDA/H : trouble du déficit de l'attention/hyperactivité; ÉSPT : état de stress post-traumatique.

2.2 L'identification des études

Une fois les instruments identifiés, la seconde étape fut d'identifier les études portant sur la fidélité ou la validité des diagnostics communs identifiés. Une recherche en ce sens a donc été faite pour chacun des 11 instruments retenus et présentés au Tableau IX.

2.2.1 Identification des bases de données

Les trois plateformes de recherche suivantes furent utilisées : OVID, WEB of Science et Proquest entre le 8 et le 30 novembre 2012. Voici le détail des bases de données qu'elles regroupent.

Ovid :

EBM Reviews - Cochrane Database of Systematic Reviews 2005 to November 2012

EBM Reviews - Health Technology Assessment 4th Quarter 2012

Health and Psychosocial Instruments 1985 to October 2012

Ovid MEDLINE(R) 1946 to November Week 3 2012

PsycINFO 1806 to November Week 4 2012

Proquest:

Dissertation and theses database

Health and medicine database

Social science database

Web of science:

Science Citation Index Expanded (SCI-EXPANDED) --1945-present

Social Sciences Citation Index (SSCI) --1956-present

Arts & Humanities Citation Index (A&HCI) --1975-present

Conference Proceedings Citation Index- Science (CPCI-S) --1990-present

Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) --1990-present

Pour chacun des instruments, les mots clés ont été sensiblement les mêmes, à savoir :

*scale**, *psychometr**, *valid**, *reliab** dans « sujet »

Titre de l'instrument dans « sujet »

Titre de l'instrument dans « titre »

Croisement de 1+2+3

Pour Web of Science :

Titre dans « sujet » et dans « titre »

*scale**, *psychometr**, *valid**, *reliab** dans « sujet »

Croisement de 1+2

Finalement avec Proquest :

Titre + *scale**, *psychometr**, *valid**, *reliab**

2.2.2 Les critères d'inclusion et d'exclusion des articles de recherche

Parfois l'abréviation du nom de l'instrument était utilisée, mais la plupart du temps, cela générerait trop de publications non pertinentes. Parfois, l'indication « A » (pour adolescent) était enlevée parce qu'en la gardant, aucune étude n'était repérée. Le Tableau X (p.48) résume le nombre d'études repérées pour chaque instrument avec chacune des trois plateformes. Le critère principal pour l'inclusion des études répertoriées était le suivant : analyse de la validité ou la fidélité d'un des six diagnostics identifiés préalablement chez des jeunes âgés de 13 à 18 ans. À partir du Tableau X, on peut également identifier le nombre de publications non retenues. Les raisons principales de leur exclusion sont répertoriées ici-bas :

- Doublon, l'article ayant été produit à partir d'une thèse (si l'article était disponible, ce dernier était retenu);
- Étude de prévalence;
- Autres diagnostics que les six retenus;
- Diagnostics groupés (par exemple, combiner les TC et TOP);
- Autre objectif de recherche que l'évaluation de la validité et/ou fidélité ;
- Revue ou résumé de résultats d'autres études;
- Âge des sujets de l'échantillon ne correspondant pas aux critères;
- L'instrument à l'étude sert de référence (étalon) et n'est donc pas celui qui est évalué;
- Manuel, livre ou chapitre de livre ou critique de livre;

- Document non disponible via le service des bibliothèques de l'Université de Montréal;
- Étude datant d'avant 1975.

Des références clés ont été repérées à partir des bibliographies des études retenues dans un premier temps. Elles peuvent représenter des ouvrages de référence ou des études qui sont centrales par rapport au présent sujet d'étude. Si ces nouvelles références répondaient aux critères d'inclusion, elles étaient ajoutées aux études retenues.

Tableau X : Occurrences des instruments dans OVID, Web of Science et Proquest

	Abréviation	Total occurrences des trois plateformes	Total ds EN ¹⁶ (retenu sur titre ou résumé)	Retenus
1	ADI	Ovid : 24 Web of science : 7 Proquest : 35 Total : 66	19	1
2	APS	Ovid : 77 Web of Science : 7 Proquest : 57 Total : 141	18	1
3	CAPA	Ovid : 159 Web of Science : 38 Proquest : 144 ¹⁷ Total : 341	30	3
4	ChIPS	Ovid : 74 Web of Science : 10 Proquest : 34 Total: 118	21+ 1 sur réf.	7
5	DICA	Ovid : 441 Web of Science : 118 Proquest : 423 Total : 982	50 + 1 sur réf.	9

¹⁶ Total moins les doublons de l'une et l'autre des plateformes.

¹⁷ Lorsque + de 100 occurrences : vérification des 100 premières.

6	DISC	Ovid :	317		
		Web of Science :	33		
		Proquest :	1515		
		Total :	1865	108	22
7	GAIN	Ovid :	102		
		Web of Science :	3		
		Proquest :	97		
		Total :	202	31	1
8	ISCA	Ovid :	27		
		Web of Science :	1		
		Proquest :	24		
		Total :	52	10	0
9	K-SADS	Ovid :	348		
		Web of Science :	24		
		Proquest :	1393		
		Total :	1765	96 +2 sur réf.	13
10	PADDI	Ovid :	4		
		Web of Science :	0		
		Proquest :	5		
		Total :	9	8	2
11	PAI-A	Ovid :	92		
		Web of Science :	138		
		Proquest :	5		
		Total :	235	48	0
		Total :	6874	436 +5 sur réf.	59¹⁸

3. Comité accompagnateur

Quelques intervenants ou chercheurs clés ont été contactés afin de vérifier si le corpus couvrait bel et bien l'ensemble des études ou instruments qu'ils connaissaient. Ce groupe inclut le directeur du présent mémoire, des chercheurs du Light House Institute de l'Illinois¹⁹

¹⁸ Total d'études : 57, puisque deux études testent deux instruments chacune.

¹⁹ Division de recherche et de formation affiliée au Chesnut Health Systems de l'État de l'Illinois (dépendances, santé mentale et jeu). <http://www.chestnut.org/LI/>

et des intervenants et des psychiatres du CSSSNL²⁰. Aucune absence importante n'a alors été signalée.

4. Analyse de la qualité méthodologique des études retenues

4.1 Critères du QADAS-2

Pour évaluer la qualité méthodologique de chaque étude retenue, le QUADAS-2 (Whiting et al., 2011) fut choisi. Cette grille, conçue pour évaluer la qualité des études diagnostiques, a été révisée suite aux commentaires de ses premiers utilisateurs et aux suggestions de la Cochrane Collaboration. Elle permet d'identifier la possibilité de biais dans les devis de recherche. Un biais est une erreur systématique, une déviation, qui peut mener à une surestimation ou sous-estimation d'un objet d'étude. Il est donc important de chercher à les identifier (Higgins et Green, 2009). Le QUADAS-2 permet de le faire et ainsi d'avoir une meilleure idée de la robustesse méthodologique d'une étude. À partir de notre question de recherche, le corpus a été évalué sous quatre angles ou domaines : 1) la sélection des sujets, 2) l'instrument mis à l'épreuve, 3) l'instrument de référence (leurs conditions d'application et d'interprétation) et 4) la séquence des opérations et les délais caractérisant le devis de l'étude.

Tel que le suggèrent les auteurs du QUADAS-2, la grille d'analyse a été ajustée pour tenir compte des besoins de la présente recherche, ce qui est expliqué dans la section suivante.

4.2 Question de départ

La question à l'origine de la SS est reprise et c'est à partir d'elle qu'on évaluera si les objectifs des études lui correspondent. C'est ainsi que, dans un premier temps, on peut juger de la correspondance («*applicability concerns*») entre les objectifs des études retenues et le présent projet. Celles-ci doivent porter sur l'analyse de la validité ou de la fidélité d'un des 11 instruments retenus, auprès d'une population adolescente et en lien avec un des six diagnostics retenus. Certaines études sont retenues malgré le fait que leur objectif principal ne répond pas directement à la question de départ lorsqu'elles comportent des données sur

²⁰ Centre de santé et services sociaux du nord de Lanaudière, programme régional de pédopsychiatrie.

la validité ou la fidélité des instruments. Dans un deuxième temps, chacun des quatre domaines du QUADAS-2 est évalué, les critères étant adaptés en fonction de l'objectif principal de la SS.

4.3 Les domaines

Chaque domaine contient des questions pour aider à l'évaluation de possibles biais, et ce en lien avec quatre éléments bien précis des devis de recherche.

Le premier domaine couvre la sélection des sujets de l'étude pour tenter de cerner les biais possibles et en évaluer la pertinence en lien avec les objectifs poursuivis.

Domaine 1 : Sélection des sujets. Cible : jeunes de 13 à 18 ans.

Risque : la sélection des sujets a-t-elle pu introduire un biais ?

Questions de précision :

S'agit-il d'un échantillon recruté de manière aléatoire (*random*) ou consécutive?

A-t-on évité le recours à des groupes contrôles ²¹?

Y a-t-il eu exclusion inappropriée de certains participants ?

Correspondance : Y a-t-il des doutes quant à la possibilité que les sujets retenus ne correspondent pas à la question de départ ?

Le second domaine examine l'instrument qui est évalué. On analysera les conditions d'utilisation, de passation et d'interprétation, ainsi que l'impact de ces conditions sur les objectifs poursuivis.

Domaine 2 : Instrument à l'étude. Cible : un des 11 instruments d'évaluation retenus.

Risque : est-ce que la passation ou l'interprétation (des résultats) de l'instrument a pu introduire un biais ?

Questions de précision :

²¹ Selon les concepteurs du QUADAS, la comparaison de participants symptomatiques à des participants contrôles risque d'exagérer les qualités psychométriques d'un instrument. Cela n'est donc pas recommandé.

Est-ce que les résultats de l'instrument à l'étude ont été interprétés sans connaître les résultats de l'instrument de référence (validité) ? Dans le cas d'étude portant sur la fidélité (test-retest ou interjuge), les résultats de l'une des passations ont-ils été interprétés sans connaître les résultats de l'autre ?

Est-ce que les évaluateurs ont été formés à l'instrument ?

Correspondance : Y a-t-il des doutes quant à la possibilité que la passation ou l'interprétation de l'instrument à l'étude ne corresponde pas à la question de départ ?

Le troisième domaine est semblable au deuxième puisqu'il évalue l'instrument de comparaison utilisé pour tester celui à l'étude. On vérifiera donc également les conditions d'utilisation, de passation et d'interprétation, en plus de la pertinence du choix de cet instrument en lien avec la condition visée par l'étude.

Domaine 3 : Instrument de référence dans les études de validité. Cible : un autre instrument d'évaluation, des échelles (*checklists*) ou une évaluation clinique seront utilisés en tant qu'instrument de référence. Cible : données en lien avec un des six diagnostics retenus.

Note : dans le cas d'études portant sur la fidélité, le domaine 3 ne sera pas évalué.

Risque : est-ce que la passation ou l'interprétation (des résultats) de l'instrument de référence a pu introduire un biais ?

Questions de précision :

Est-ce que l'instrument de référence choisi est à même de bien identifier la condition ciblée ?

Est-ce que les résultats de l'instrument à l'étude ont été interprétés sans connaître ceux de l'instrument de référence (validité) ?

Correspondance : Y a-t-il des doutes quant à la possibilité que l'état ciblé, tel que défini par l'instrument de référence, ne corresponde pas à la question de départ ?

Le quatrième et dernier domaine s'attardera au déroulement de l'étude, incluant les délais. On notera également qui s'est vu administrer ou non les instruments choisis : y a-t-il présence d'attrition ?

Domaine 4 : Séquence et délais.

Risque : est-ce que la séquence des opérations ou les délais ont pu introduire un biais ?

Questions de précision :

Est-ce que le délai entre la passation de l'instrument à l'étude et l'instrument de référence (validité) ou entre la première passation et la deuxième (fidélité) était adéquat ?

Est-ce que tous les sujets se sont vu administrer un instrument de référence ? Le même ?

Note : dans le cas d'études portant sur la fidélité, ces questions seront analysées sous l'angle de la deuxième passation : est-ce que tous les sujets se sont vu administrer une deuxième passation ? Toujours avec l'instrument à l'étude ?

Est-ce que tous les sujets ont été inclus dans l'analyse finale ?

Quel est le risque de biais introduit par la séquence ou les délais ?

Suite au survol de la littérature du chapitre 1, en lien avec le sujet du présent mémoire, il apparaît évident que la présence de trouble de SM chez les adolescent(e)s, incluant les TUS, est une réalité qu'on ne peut remettre en question. Il appert également que ces jeunes ont des besoins propres en lien avec leurs stades de développement et que pour cette raison, l'évaluation de ces troubles doit être développée et adaptée à cette réalité et non calquée sur celle des adultes. Des instruments d'évaluation construits en fonction de ces prémisses doivent être rendus disponibles aux intervenants œuvrant auprès de cette clientèle pour qu'ils soient en mesure de leur offrir des traitements efficaces et efficients. Ces instruments existent-ils ? Quels sont-ils ? Ont-ils fait leurs preuves ? C'est avec la méthodologie présentée dans le présent chapitre que des réponses à ces questions seront apportées. Le chapitre suivant présente les résultats obtenus à ces questions. Tout d'abord les instruments identifiés seront sommairement présentés. Ensuite, ce sera les études recensées à leur sujet qui le seront. Suivra une analyse de plusieurs éléments de ces études : les types de validité et de fidélité utilisés, les taux recensés pour les six diagnostics retenus et un résumé global des résultats obtenus pour chacun des 11 instruments seront présentés.

Chapitre 3 : Résultats

1. Présentation des instruments retenus

Dans la présente section, chacun des instruments répertoriés fait l'objet d'une présentation sommaire, incluant un bref historique, les motifs de sa conception, ses spécificités, les échelles ou sections qui le composent, la version du DMS auquel il est rattaché et les versions qui sont disponibles. De plus, le Tableau XI (en p.63) résume l'âge des jeunes visés et le nombre de diagnostics possibles, le type d'entrevue à mener (basé sur l'intervieweur ou le participant), la structure de l'instrument, le temps d'administration prévu, la disponibilité de versions pour les parents (ou tiers significatif) ou adolescent(e)s et les langues dans lesquelles l'instrument est disponible. La dernière colonne précise si une formation est requise pour la passation et si cela requiert des cliniciens ou des non-cliniciens. Tous les instruments ont été conçus aux États-Unis, sauf le CAPA qui est le fruit d'une collaboration entre les États-Unis et l'Angleterre.

1.1 Adolescent Diagnostic Interview-ADI

(Winters et Henly, 1993)

L'ADI est un instrument d'évaluation multiaxiale qui cible principalement les TUS. Il peut être utile autant pour les chercheurs que les professionnels. Il identifie également les stressors psychosociaux, ainsi que les niveaux de fonctionnement scolaire, interpersonnel, occupationnel, légal et cognitif. De plus, l'ADI permet d'identifier la présence de troubles concomitants. Développé d'abord en lien avec le DSM-III, puis ajusté au moment de la parution du DSM-IV, l'ADI fait partie d'une famille d'instruments, le *Minnesota Chemical Dependency Adolescent(e) Assessment Package* (Winters et Henly, 1986), qui comprend un outil de détection et un inventaire multidimensionnel. Un questionnaire auto-administré accompagne également l'ADI pour arriver à une évaluation des plus complètes.

Suite au projet *Adolescent Assessment/Referral System (AARS)* réalisé en 1991 (Rahdert, National Institute on Drug Abuse, Pacific Institute for Research Evaluation et Westover Consultants, 1991) et au constat sur l'étendue des problèmes associés à la consommation de SPA des jeunes, le *National Institute on Drug Abuse* a identifié l'ADI comme l'instrument idéal pour évaluer les TUS chez les adolescent(e)s. L'objectif du projet AARS était d'identifier les

instruments existants et d'émettre des lignes directrices balisant leur utilisation dans les milieux cliniques, auprès des jeunes âgés de 12 à 19 ans.

1.2 Adolescent Psychopathology Scale-APS

(Reynolds, 1998)

L'APS est un instrument multidimensionnel en mesure d'orienter les jeunes vers les services ou traitements nécessaires. Conçu pour identifier les symptômes psychosociaux et psychiatriques, il regroupe 40 échelles (20 sur les troubles cliniques, cinq sur les troubles de personnalité, 11 sur les problèmes psychosociaux et quatre sur le style de réponses). Basé sur le DSM-IV, il génère des diagnostics de l'Axe I et II. Il existe aussi en version abrégée, celle-ci ayant été traduite en portugais. Une fois l'entrée de données effectuée, un programme informatique génère les résultats, incluant les diagnostics qu'il est possible de poser.

1.3 The Child and Adolescent Psychiatric Assessment-CAPA

(Angold et Costello, 1995)

Le CAPA est un instrument multidimensionnel composé de différents modules, conçu dans le cadre d'une collaboration entre chercheurs et cliniciens des États-Unis et de l'Angleterre. Les auteurs ont voulu en faire un instrument qui soit à la fois clinique et épidémiologique. Il se divise en trois sections (l'introduction, la revue des symptômes et les atteintes au fonctionnement psychosocial) qui regroupent les informations suivantes : les données sociodémographiques, la structure et le fonctionnement familial, les relations interpersonnelles, l'évaluation des événements marquants de la vie, les symptômes regroupés selon des « domaines » diagnostiques, ainsi que l'impact de la symptomatologie psychiatrique sur le fonctionnement psychosocial. On peut également évaluer l'utilisation de services en SM, les obstacles à l'amorce d'un traitement et l'impact de la maladie sur la famille. Il génère des diagnostics à partir des DMS-III et IV, mais répertorie également des symptômes d'intérêt psychodéveloppemental. Le CAPA provient d'une grande famille d'instruments incluant une version abrégée, pour les jumeaux, pour les jeunes adultes et

enfants du préscolaire. Une version informatique pour l'auto-administration par le jeune lui-même est aussi disponible.

1.4 Children's Interview for Psychiatric Syndromes-ChIPS

(Weller, Weller, Fristad et Rooney, 1999; Weller, Weller, Fristad, Rooney et Schecter, 2000)

Le ChIPS a été conçu au départ pour répondre aux besoins particuliers de jeunes enfants de 6 à 12 ans ayant des difficultés à comprendre les items des autres questionnaires. On porta une attention particulière à la longueur des questions, ainsi qu'au vocabulaire utilisé. Par la suite, l'instrument fut testé avec les adolescent(e)s de 13 à 18 ans et les résultats furent concluants. Depuis, il peut être utilisé lorsqu'une évaluation diagnostique est nécessaire, ainsi que pour mener des recherches cliniques. Le ChIPS est composé de 15 sections qui évaluent 20 diagnostics de l'Axe I du DMS-IV et de deux sections qui couvrent certains stressors psychosociaux, dont l'abus et la négligence. Les questions cardinales posées au début de chaque section permettent par la suite de n'explorer que celles où des symptômes sont rapportés, sauf pour les sections sur l'abus et la négligence où toutes les questions doivent être posées. La grille des résultats (diagnostics et symptômes) inclut également les informations démographiques, des observations sur le comportement et toute autre information pertinente. Une version informatique est disponible.

1.5 Diagnostic Interview for Children and Adolescents-DICA

(Herjanic et Reich, 1982; Reich, 2000)

Le DICA a été conçu suivant le modèle du *Diagnostic Interview Schedule* (DIS) (Robins, Helzer, Ratcliff et Seyfried, 1982), un instrument pour les adultes. La première version, élaborée à partir du DMS-III puis ajustée au DMS-III-R, avait pour objectif d'améliorer la validité de la collecte de données auprès des enfants. À partir de la version du DMS-III-R, deux instruments distincts furent construits, un pour les 6-12 ans et un pour les 13-18 ans. Autrement dit, une adaptation développementale a été effectuée pour obtenir des questions appropriées selon le groupe d'âge. À partir de 1992, l'instrument est devenu plus souple dans son format, tout

en demeurant plus structuré que les instruments semi-structurés. L'intervieweur devait en avoir une connaissance suffisante pour moduler son exploration selon les symptômes présents. La version de 1997, ajustée au DMS-III-R ou IV, génère des diagnostics se rapportant à la vie entière. Une première section contient les données démographiques, en lien avec les sphères scolaire et occupationnelle. Cette dernière contient également des questions que l'auteur (Reich, 2000) nomme « *happy questions* » (questions joyeuses). Elles servent à créer un lien non menaçant et amorcent la collecte d'information au niveau de l'atteinte du fonctionnement psychosocial. Dans la version parent, une section est consacrée à la période prénatale, à l'accouchement et aux complications s'il y a lieu. Des versions informatisées sont disponibles, dont une pour l'auto-administration. Trois dérivés sont aussi disponibles pour réaliser des études spécifiques :

- *Semi-Structured Assessment on the Genetics of Alcoholism (C-SSAGA)* (à utiliser lorsque l'objectif est de trouver des gènes prédisposant les individus à la dépendance à l'alcool).
- *Missouri Assessment of Genetics Interview for Children (MAGIC)* (conçu pour étudier le TDA/H et la dépression).
- *Structured Assessment Record of Alcoholic Homes (SARAH)* (conçu par l'équipe du DICA pour évaluer les différents aspects de l'impact de la consommation d'alcool des parents sur les enfants).

1.6 Diagnostic Interview Schedule for Children-DISC

(Costello A.J., Edelbrock C.S., Dulcan M.D., Kalas R. et Klaric S.H., 1984; Jensen, Roper, Fisher, Piacentini et al., 1995; Shaffer et al., 1996; Shaffer, Fisher, Lucas, Dulcan et Schwab-Stone, 2000; Shaffer et al., 1993)

Développé sous l'égide du *National Institute of Mental Health*, le DISC évolue depuis 1979. Tout comme le DICA, il a été conçu sur le modèle du *Diagnostic Interview Schedule (DIS)* (Robins et al., 1982) et est d'abord destiné aux grandes études épidémiologiques. En 1981, Costello et son équipe (1984) ont mis une nouvelle version à l'épreuve avec un échantillon

clinique, le DISC-1, qui génère des diagnostics du DSM-III. D'amélioration en amélioration (DISC-R (Shaffer et al., 1993), DISC-2.1 (Jensen et al., 1995), DISC-2.3 (Shaffer et al., 1996), la version du DISC-IV (Shaffer et al., 2000), relevant du DMS-IV, a vu le jour. Regroupant des questions courtes (pas plus de deux concepts par question) pouvant être répondues par oui/non, la première section débute avec des items qui interrogent les données démographiques, recueillent des informations générales et vérifient la bonne compréhension des questions par les participants. Par la suite, six modules couvrent l'ensemble des diagnostics possibles. Pour chacun d'eux, le DISC-IV collecte aussi des données en lien avec le degré d'atteinte au fonctionnement psychosocial. Malgré sa longueur, la structure de l'instrument fait en sorte que seules les questions pertinentes à la présence de symptômes sont posées. Le dernier module, optionnel, couvre l'information complémentaire aux diagnostics reçus depuis l'âge de cinq ans.

Des versions complémentaires du DISC-IV sont aussi disponibles : abrégée, «*present state*» (peut être utilisé comme instrument de détection, disponible en audio), pour jeunes adultes, pour enseignants ou prédictives (pouvant être utilisées comme instrument de détection). Plusieurs d'entre elles sont informatisées.

1.7 Global Appraisal of Individual Needs-GAIN

(Dennis, Rourke, Lennox, Campbell et Caddell, 1995; Dennis, Titus, White, Unsicker et Hodgkins, 2003)

Permettant une évaluation biopsychosociale complète, le GAIN est conçu pour collecter des données cliniques et de recherche grâce à un instrument unique. Tenant compte de l'avancement des connaissances sur les TUS, l'instrument est rigoureux et fournit de l'information servant au diagnostic, au placement en milieu thérapeutique selon le niveau de besoin, à la planification du traitement et au suivi. Tout comme le DISC, de version en version, il a été ajusté pour le rendre plus performant. La version longue (GAIN-I), comporte huit sections : données démographiques, TUS, santé physique, comportements à risque, santé mentale/émotionnelle/comportementale, environnement, sphères légale et occupationnelle, incluant les études et le travail. Une section est également intégrée pour collecter des

données d'un tiers significatif, comme un parent ou un conjoint. La présence de symptômes et de problèmes psychosociaux (ampleur, récurrence et fréquence), l'utilisation de services ainsi que les désirs et croyances de l'individu sont scrutés dans chacune des sections. En l'absence de symptôme ou problème, des questions ou sections peuvent être omises. Des diagnostics compatibles avec le DSM-IV sont générés, ainsi qu'un niveau de placement selon les critères de l'ASAM (Mee-Lee, Shulman, Gastfriend et Griffith, 2001). Une sous-section du GAIN-I permet la collecte de données auprès d'un tiers significatif. Le tout est informatisé (auto-administré disponible) et produit plusieurs types de rapports écrits (rapport complet avec recommandations de traitement, motivationnel, données brutes, incohérences).

Des versions abrégée (GAIN-Q), de détection (GAIN-SS), de suivi aux 90 jours (GAIN-M90) (incluant une version abrégée) et de base (essentiel du GAIN-I) sont aussi disponibles.

1.8 Interview Schedule for Children and Adolescents-ISCA

(Kovacs M., 1985; Sherrill et Kovacs, 2000)

D'abord conçu sous le nom d'*Interview Schedule for Children* (ISC) afin de répondre aux besoins d'une étude longitudinale, l'ISCA a subi des transformations pour devenir un instrument semi-structuré, orienté sur les symptômes, pouvant générer des diagnostics relevant des DSM-III et IV. Bien que conçu pour la recherche (complet, il est impossible d'y sauter des questions), l'instrument a déjà été utilisé dans des études cliniques. Il est composé d'items évaluant la sévérité de la condition actuelle, soit le fonctionnement global et le niveau d'atteinte des symptômes dans les différentes sphères sociales. Ses cinq sections explorent les signes et symptômes, l'état mental, les comportements, impressions cliniques et étapes développementales. L'instrument inclut des feuilles de codification des informations recueillies, selon qu'elles proviennent de l'entretien avec le parent ou avec l'enfant. Des versions abrégées, sous forme d'échelles de détection de symptômes dépressifs et anxieux pourraient être disponibles sous peu afin de simplifier leur utilisation dans les milieux cliniques (Sherrill et Kovacs, 2000).

Deux versions sont disponibles, soit pour : a) l'évaluation initiale (ISCA-C-L : symptômes actuels (*Current*) et à vie (*Lifetime*) ou b) une réévaluation de suivi (ISCA-C-I : actuel et

évaluation des progrès (*interim*). Une version pour jeunes adultes est aussi disponible (*Follow-up Interview Schedule for Adults-FISA*).

1.9 Schedule of Affective Disorders and Schizophrenia for Children-K-SADS

(Chambers et al., 1985; Kaufman et al., 1997)

La version pour enfants et adolescent(e)s de cet instrument est basée sur la version adulte, soit le *Schedule for Affective Disorders and Schizophrenia (SADS)* (Endicott J. et Spitzer R.L., 1978). D'abord conçu pour permettre l'évaluation diagnostique d'épisodes psychiatriques lors d'études portant sur les traitements des dépressions chez les enfants, il répond désormais aux exigences d'une évaluation tant clinique que de recherche. De la version K-SADS-P (*Present Episode*) à celle du K-SADS-PL (*Present and Lifetime version*), les améliorations suivantes ont été apportées : identification des diagnostics à vie, ajout de diagnostics importants (ex. : TDA/H ou ÉSPT), ajustement des questions visant à clarifier les symptômes, ajouts de cotations du fonctionnement psychosocial, sans oublier les ajustements en lien avec la parution du DSM-IV. Trois sections composent le K-SADS-PL : a) l'introduction (données démographiques, la demande actuelle, les traitements psychiatriques passés, le fonctionnement occupationnel [scolaire, loisirs] et les relations interpersonnelles), b) l'entrevue de détection (qui sert à éliminer les modules diagnostiques non nécessaires) et c) les cinq modules diagnostiques (troubles de l'humeur, psychotiques, anxieux, comportementaux et TUS, incluant les troubles alimentaires et les tics).

Il existe trois versions de cet instrument : a) une est épidémiologique, b) une autre est adaptée aux troubles bipolaires, le WASH-UKSADS (Geller, Williams, Zimmerman et Frazier, 1996), ainsi que c) le K-SADSPL-W (Lingler, Bedoya et Findling, 2007), qui combine le K-SADS-PL et la section sur les troubles de l'humeur du WASH-UKSADS, tout en étant ajusté au DSM-IV.

1.10 Practical Adolescent Dual Diagnostic Interview-PADDI

(Estroff et Hoffmann, 2001)

L'un des concepteurs du PADDI (Hoffmann, 2004) affirme que celui-ci a été créé pour répondre de façon pragmatique aux besoins d'évaluation des troubles concomitants chez les adolescent(e)s. Cet instrument a donc d'abord été prévu pour l'évaluation clinique et non pour la recherche. Par ailleurs, spécifiquement conçu pour les adolescent(e)s, il ne découle pas d'un instrument pour adulte. Il ne couvre que les troubles mentaux les plus prévalents chez les adolescent(e)s, dont les TUS. Outre la collecte de données sur les symptômes, on y retrouve de l'information sur le risque pour autrui et pour soi, ainsi que sur la victimisation. Il génère des diagnostics tirés du DSM-IV. Selon le site des fournisseurs de l'instrument, la version ajustée au DMS-5 est disponible (The Change Companies, 2014).

1.11 Personality Assessment Inventory-Adolescent-PAI-A

(Morey, 2007)

Le PAI-A provient d'un instrument d'abord utilisé avec les adultes (Morey, 1991). Relativement récente, la version pour adolescent(e)s a été conçue à la demande des professionnels. Pouvant être utilisé dans plusieurs contextes (scolaire, clinique, judiciaire), cet inventaire évalue la présence de symptômes de psychopathologie, d'un trouble de personnalité et des particularités de l'environnement psychosocial. Le PAI-A contient 22 échelles distinctes : quatre de validité, 11 cliniques, cinq pour le traitement et deux interpersonnelles. Elles regroupent toutes les informations nécessaires pour poser un diagnostic et un programme informatique de compilation des données en suggère.

Tableau XI : Données techniques des instruments retenus

Données techniques des instruments retenus									
Instrument	Âge	Dx	Version	Basé	Format	Items	Temps ¹ (min.)	Langue	Formation ²
ADI	12-18	8 groupes	Ado	Participant	Structuré	213	30-90 / 15-20/?	Anglais	Oui Non- clinicien
APS	12-19	>25	Ado	Auto- administré	-	346	45-60 / ?/?	Anglais Portugais	Non- clinicien
CAPA	9-17	>30	Ado Parent	Intervieweur	Semi- structuré	1400	±60 /45/25	Anglais Espagnol	Oui Non- clinicien
CHIPS	6-18	20	Ado Parent	Participant	Structuré	?	±40/?/?	Anglais	Oui Non- clinicien
DICA	13-18	>20	Ado Parent	Intervieweur- participant	Semi- structuré	?	60- 120/10/5	Anglais Espagnol	oui Non- clinicien
DISC-IV	6-17	>30	Ado Parent	Participant	Structure	3000	70 ³ /na/na	Anglais Français Espagnol Xhosa Chinois	oui Non- clinicien
GAIN	Ado- adulte	>30	Ado Tiers	Participant	Semi- structuré	500	90- 150/30/na	Anglais Espagnol Français	Oui Non- clinicien et clinicien
ISCA	8-17	>25	Ado ³	Intervieweur	Semi- structuré	?	120-150 ⁴ 45-135 ⁵ /inclus/na	Anglais Espagnol	Non Clinicien
K-SADS-PL	7-17	>30	Ado ³	Intervieweur	Semi- structuré	82/? ⁸	35-45 ⁶ 60-85 ⁷ / ?/?	Anglais Français	Oui Clinicien
PADDI	14-18	>10	Ado	Participant	Structuré	164	30-40/5/?	Anglais	Non Non- clinicien
PAI-A	12-18	?	Ado	Auto- administré	-	264	45/15-20/?	Anglais	Non Clinicien

1- Temps de passation/temps de codification-édition/
temps d'entrée des données dans le support informatique.

2- À l'instrument/par non-clinicien ou clinicien.

3- Feuille de cotation unique qui contient les
informations fournies par le parent, pour comparaison.

4- Parent.

5- Adolescent(e)

6- Temps pour le parent et pour l'enfant, groupe contrôle.

7- Temps pour le parent et pour l'enfant, groupe clinique.

8- 82 pour module de base, plus les cinq modules diagnostiques

2. Présentation des études recensées

Les études retenues, au nombre de 57 (Tableau X en p.48) proviennent en grande majorité (N=37) des États-Unis. Parmi les pays, l'Espagne suit avec quatre études (de la Osa, Ezpeleta, Domenech, Navarro et Losilla, 1997; Ezpeleta et al., 1997; Granero Perez, Ezpeleta Ascaso, Domenech Massons et de la Osa Chaparro, 1998; Moreno, Basurto, Becerra et San Pedro, 2006), le Canada avec trois (Boyle, Offord, Racine, Sanford et et al., 1993; Boyle et al., 1997; Breton, Bergeron, Valla, Berthiaume et St-Georges, 1998) et l'Iran avec deux (Ghanizadeh, Mohammadi et Yazdanshenas, 2006; Shahrivar, Kousha, Moallemi, Tehrani-Doost et Alaghband-Rad, 2010). L'Afrique (Flisher, Sorsdahl et Lund, 2012), le Brésil (Brasil et Bordin, 2010), la Grèce (Kolaitis, Korpa, Kolvin et Tsiantis, 2003), l'Islande (Lauth et al., 2010), le Portugal (Lemos, Faisca et Valadas, 2011), la Suisse (Rothen et al., 2009) et Puerto Rico (Bravo et al., 2001), chacun une. Une étude provient de la Chine, mais en collaboration avec des auteurs étasuniens (Ho et al., 2005) et deux dernières d'équipes réunissant Portoricains et Américains (Ribera et al., 1996; Rubio-Stipec et al., 1994).

La majorité des études ont été publiées dans des revues scientifiques (55), les deux autres étant des thèses (McGrath, 2000; Rathman, 2008). Il a été très difficile d'obtenir les manuels explicatifs, guides de passation ou d'interprétation des instruments, certains contenant des données psychométriques originales. Même en vérifiant avec le système de prêts entre les universités, plusieurs n'étaient pas disponibles. Une explication peut être le fait que ces documents sont vendus par les fournisseurs de ces instruments.

Les types d'échantillons sont d'origines diverses, mais principalement cliniques. Dans 48 études, les jeunes recrutés : a) sont suivis sur une base externe ou interne (hospitalisés ou incarcérés); b) reçoivent des soins en psychiatrie ou c) un traitement en lien avec un TUS ou un TC. Certains sont judiciairisés et incarcérés, d'autres placés en famille d'accueil. Une étude concerne des jeunes ayant reçu un diagnostic d'autisme (Witwer, Lecavalier et Norris, 2012) et deux autres (Dillon, Turner, Robbins et Szapocznik, 2005; Roberts, Solovitz, Chen et Casat, 1996) des jeunes issus des minorités ethniques. Au total, 17 de ces études utilisent un groupe contrôle issu de la communauté (cliniques pédiatriques, du milieu scolaire, à partir de

recensements, prenant déjà part à une autre étude ou en traitement, mais sans la condition recherchée ou encore à partir de publicités). Neuf études ont utilisé des échantillons de la communauté : cliniques pédiatriques, milieux scolaires, recensement ou provenant d'une autre étude. Deux de celles-ci (Boyle, Offord, Racine, Sanford, et al., 1993; Boyle et al., 1997) portent sur des jeunes avec une forte probabilité de diagnostics de TM et une autre (Schwab-Stone, Shaffer, Dulcan, Jensen et al., 1996) sur des jeunes avec un diagnostic confirmé de la condition ciblée.

Plusieurs études (18) mettent à l'épreuve une version traduite ou adaptée de l'instrument. Ainsi, on retrouve des études portant sur des versions espagnoles (Bravo et al., 2001; de la Osa et al., 1997; Ezpeleta et al., 1997; Granero Perez et al., 1998; Moreno et al., 2006; Ribera et al., 1996; Rubio-Stipec et al., 1994), française (Breton et al., 1998; Rothen et al., 2009), portugaise (Lemos et al., 2011), Xhosa (dialecte de l'Afrique du Sud) (Fisher et al., 2012), chinoise (Ho et al., 2005), hébreux (Shanee, Apter et Weizman, 1997), grecque (Kolaitis et al., 2003), islandaise (Lauth et al., 2010), brésilienne (Brasil et Bordin, 2010) et perse (Ghanizadeh et al., 2006; Shahrivar et al., 2010). Deux études ont testé une version informatique de l'instrument (Friman et al., 2000; Reich, Cottler, McCallum, Corwin et VanEerdewegh, 1995) et une le mode de passation par téléphone (Paing, Weller, Dixon et Weller, 2010). Le nombre d'individus de chaque échantillon varie de 10 (Rathman, 2008) à quelques centaines d'individus et même des milliers, comme pour l'échantillon de départ de 1751 individus de l'étude de Boyle et ses collaborateurs (1993).

Malgré que la période visée par cette recension débute en 1975, aucune étude n'a été identifiée entre 1975 et 1980. Sept études ont été publiées entre 1980 et 1990 et 25 pour chacune des deux périodes de 1990-2000 et 2000-2012.

2.1 Résultats de l'évaluation de qualité méthodologique (QUADAS-2)

Note : Dans cette partie du chapitre 3, les numéros du tableau XII en p.67 sont utilisés pour faire référence aux études qui sont souvent indiquées entre [...].

L'analyse à l'aide de la grille du QUADAS-2 permet une évaluation de la robustesse méthodologique des études retenues. Un premier constat est que pour un nombre significatif

d'études, des données importantes sont manquantes. En effet, pour 18 des 57 études, un manque d'information n'a pas permis d'évaluer la qualité méthodologique dans un ou plusieurs domaines. Il s'agit surtout d'informations en lien avec l'échantillon ou l'instrument à l'étude. Soit les critères d'exclusion ou le type d'échantillonnage ne sont pas précisés. Soit il est impossible de savoir si les intervieweurs ont été formés à l'instrument ou si les résultats de ce dernier ont été interprétés en tenant compte de ceux de l'instrument de référence (ou de la passation du temps 2 s'il s'agit d'une évaluation de la fidélité).

Tableau XII : Analyse des études retenues à l'aide de la grille QUADAS-2

Instrument	Étude	RISQUE de BIAS				Souci de correspondance	
		Sélection des patients	Instrument à l'étude	Instrument de référence	Séquence et délais	Sélection des patients	Instrument l'étude
ADI	1Winters, 1993	😊	😊	😊	😊	😊	😊
APS	2Lemos, 2011	😞	😞	X	😞	😊	😊
CAPA	3 Angold, 1995	😊	😊	X	😊	😊	😊
	4 Angold, 2012	😊	😊	?	😊	😊	😊
ChIPS	5 Costello, 1998	😞	😊	X	😊	😊	😊
	6 Fristad, 1998a	😊	😊	X	😊	😊	😊
	7 Fristad, 1998b	?	😊	😊	😊	😊	😊
	8 Moreno, 2006	😞	😊	😊	😊	😊	😊
	9 Paing, 2010	😞	?	X	😊	😊	😊
	10 Rathman, 2008	😊	😞	😞	😊	😊	😞
DICA	11 Swenson, 2007	😊	😊	😊	😞	😊	😊
	12 Witwer, 2012	😞	😊	😊	😊	😊	😊
	13 Boyle, 1993	😞	😊	X	😊	😊	😊
	14 Boyle 1997	😞	😊	😊	😞	😊	😊
	15 Dela Osa, 1997	😊	😊	😞	😊	😊	😊
	16 Ezpeleta, 1997	😊	😊	😊	😊	😊	😊
	17 Granero, 1998	?	😊	X	😊	😊	😊
	18 Hope, 1999	😞	😊	X	😊	😊	😊
	19 Reich, 1995	😞	?	X	😊	😊	😊
	20 Sylvester, 1987	😞	😊	😊	😊	😊	😊
DISC	21 Welner, 1987	😊	😊	😊	😞	😊	😊
	22 Aronen, 1993	😊	😊	😞	😞	😊	😊
	23 Bravo, 2001	😊	😊	X	😊	😊	😊
	24 Breton, 1998	😊	😊	X	😊	😊	😊
	25 Crowley, 2001	😊	?	?	😞	😊	😊
	26 Dillon, 2005	😊	?	?	😞	😊	😊
	27 Edelbrock, 1999	😞	😞	X	😊	😊	😊
	28 Flischer, 2012	?	😞	X	😊	😊	😞
	29 Friman, 2000	😊	😊	😊	😊	😊	😊
	30 Ho, 2005	😊	😊	X	😊	😊	😊
	31 Jensen, 1995	😞	😊	X	😊	😊	😊
	32 Jewell, 2004	😊	😊	😊	😞	😊	😊

	33 King, 1997						
	34 Lewczyk, 2003						
	35 McGrath, 2000						
	36 Piacentini, 1993						
	37 Ribera, 1996						
	38 Roberts, 1996						
	39 Rubio-S., 1994						
	40 Schwab-S., 1993						
	41 Schwab-S., 1996						
	42 Weinstein, 1989						
GAIN	43 Dennis, 2004						
K-SADS	44 Ambrosini, 1989						
	45 Brasil, 2010						
	46 Chambers, 1985						
	47 Ghanizadeh, 2006						
	48 Hodges, 1987						
	49 Kaufman, 1997						
	50 Kolaitis, 2003						
	51 Lauth, 2010						
	52 Rothens, 2009						
	53 Shahrivar, 2010						
	54 Shanee, 1997						
	55 Weissman, 1987						
PADDI	56 Abrantes, 2004						
	57 Hoffman, 2004						
ISCA	Aucune étude						
PAI-A	Aucune étude						

Risque faible

Risque élevé

Risque incertain (données manquantes) ne s'applique pas (fidélité)

Deuxième constatation, c'est au niveau de l'échantillonnage et de la séquence qu'il semble y avoir le plus de risques de biais. En effet, plusieurs études (N = 15) n'ont pas utilisé une stratégie d'échantillonnage impliquant une sélection consécutive ou randomisée ou n'ont pas évité l'utilisation d'un groupe contrôle. Selon les concepteurs du QUADAS-2, l'utilisation d'un groupe contrôle, regroupant des sujets sans la condition ciblée, pourrait exagérer les qualités psychométriques d'un instrument (Whiting et al., 2011). Au niveau de la séquence, on retrouve 12 études à risque élevé de biais. Des quatre éléments inclus dans ce domaine, c'est au niveau de l'inclusion de tous les sujets dans l'analyse qu'il semble y avoir plus de lacunes : dans 20 études, les chercheurs n'y sont pas arrivés. Par ailleurs, dans neuf études, les délais entre la passation de l'instrument à l'étude et l'instrument de référence (ou le délai entre les passations interjuge ou test-retest) étaient inadéquats c'est-à-dire de plus d'un mois.

De toutes les études retenues, peu présentaient un risque au niveau de la correspondance : les critères d'inclusion mentionnés au chapitre 2 ont assuré leur pertinence. Par contre, il est possible que la qualité des données recueillies sur la validité et la fidélité de ces études soit quelque peu faussée, les risques de biais se situant au niveau des domaines 1, 2 ou 3 demeurent. Les seules études qui ont été retirées des analyses (N = 10), l'ont été parce que dans trois domaines ou plus, un risque de biais élevé était clairement identifiable ou que les données étaient manquantes et ne permettaient pas de statuer sur le risque. Il s'agit des études 2, 10, 25, 26, 28, 33, 34, 35, 48 et 52 du Tableau XII. Identifiées en gris dans les Tableaux 1 à 6 des diagnostics ciblés (annexes B à G), elles seront exclues des analyses et des compilations qui suivent. L'échantillon final est donc constitué de 47 études.

2.2 Type de validité et fidélité répertoriées

Selon Blacker et Endicott (2008) la psychométrie s'intéresse à la performance de plusieurs types de mesures. Ils ajoutent que les deux principales propriétés psychométriques d'une mesure sont la validité et la fidélité. L'évaluation de la validité et de la fidélité détermine le potentiel d'une mesure en lien avec son objectif. Les deux concepts sont distincts, mais interreliés. La fidélité fait référence à la constance et la répétitivité dans le temps avec lesquelles les individus sont discriminés selon un critère particulier. La validité fait référence à

la conformité avec une valeur vraie, ou ce qui s'en approche le plus (*gold standard*) : l'instrument permet-il une classification exacte ? Blacker et Endicott (2008 :8) illustrent cette relation de la manière suivante:

“Reliability is sometimes said to put an upper limit on validity because of the “noise” of poor reliability makes it hard to detect a valid “signal.”

Dans les articles retenus et mettant à l'épreuve les instruments d'évaluation, plusieurs types de validité et de fidélité ont été utilisés. En voici un aperçu, ainsi que les méthodes retenues pour les évaluer.

2.2.1 Mesures de la fidélité

La cohérence interne est l'accord entre les différents items qui composent la mesure. Chacun des items est vu comme une mesure individuelle du construit ciblé. Ainsi, la cohérence des items signifie que chacun mesure le même construit. Dans les études retenues pour cette SS, la méthode d'évaluation généralement utilisée est le coefficient alpha de Cronbach (1951), bien qu'on retrouve la mesure Kuder-Richardson dans un cas [12].

L'accord interjuge est une mesure de la fidélité entre deux ou plusieurs évaluateurs qui se penchent sur un même individu, à l'aide des mêmes informations. Le test-retest mesure quant à lui l'accord entre deux évaluations, à deux moments différents dans le temps. Dans ce dernier cas, on ne mesure réellement la fidélité que lorsque la condition du participant demeure inchangée. L'intervalle entre le temps 1 et le temps 2 est donc primordial : juste assez court pour qu'il n'y pas de changement dans la condition évaluée, juste assez long pour que le participant ne se souvienne pas de ses réponses. Il demeurera toujours une variance due au fait, que malgré une condition stable, la compréhension et l'état de confort ou d'inconfort du participant fluctue inévitablement du temps 1 au temps 2.

La mesure du test-retest, plus près de la réalité clinique et ses difficultés (une série d'évaluations par plusieurs intervieweurs), tend à être un meilleur indicateur de la fidélité de l'instrument dans la situation où il sera effectivement utilisé (Blacker et Endicott, 2008). Reich (2000) mentionne par contre qu'il faut tenir compte de l'atténuation, phénomène qui

consiste à donner moins de réponses au temps 2 d'un test-retest. On n'en comprend pas toutes les facettes, mais Reich énumèrent quelques possibilités :

- Le participant peut vouloir réduire le temps de passation ;
- Il peut ne pas comprendre le but du test-retest et assumer que les réponses du temps 1 sont incorrectes, d'où la nécessité d'une deuxième passation ;
- Il peut approcher avec ennui une deuxième passation identique à la première ;
- Il peut avoir un état d'âme particulier du participant au temps 2, différent du temps 1 ;
- Un réel changement dans l'état clinique ou développemental du participant a pu se produire;
- Un effet « thérapeutique » a pu se produire lors de la passation du temps 1 ;
- Une meilleure appréciation de l'évaluateur est possible dans l'atteinte du seuil diagnostique ;
- Un effet statistique de régression vers la moyenne a pu jouer (Milich, Roberts, Loney et Caputo, 1980, cité dans Reich, 2000).

Plusieurs facteurs doivent donc être pris en considération lorsque vient le temps d'interpréter l'analyse de la fidélité à l'aide du test-retest.

L'accord interjuge et la fidélité test-retest sont évalués à l'aide des mêmes méthodes : le coefficient Kappa qui mesure l'accord en corrigeant le risque d'accord par chance (Fleiss, 1981) pour les données catégorielles et le K de Cohen (1968) pour les mesures ordinales. Dans certaines études moins récentes, le Yules Y, qui peut s'interpréter tel que le K, a aussi été recensé. A également été utilisée la mesure r de Pearson (1896) qui mesure le lien de corrélation entre deux variables, mais ne corrige pas le risque d'accord obtenu par chance. Finalement, le pourcentage d'accords est aussi fort utilisé, même s'il est moins précis selon Blacker et Endicott (2008), puisqu'il ne corrige pas le taux d'accord pouvant être obtenu par chance. Blacker et Endicott précisent aussi que lorsque ces statistiques sont analysées, des facteurs extérieurs à l'instrument lui-même doivent être pris en considération. D'abord,

l'estimation à partir de petits échantillons peut être erronée, comme pour toute statistique. Ensuite, les conditions dans lesquelles l'instrument est administré (endroit, bruit, chaleur, etc.), ainsi que l'état de l'intervieweur (enrhumé, distrait émotionnellement, peu formé, peu expérimenté, etc.) jouent sur la variance. Les conditions très contrôlées caractérisant les études scientifiques tendent donc à augmenter artificiellement les résultats de la fidélité. Des caractéristiques de l'échantillon jouent aussi sur cette dernière telles que : les cas difficiles (individu juste sous le seuil diagnostique ou peu enclin à être évalué) et les risques liés à la généralisation de la fidélité d'un résultat entre des groupes différents (groupe clinique versus groupe contrôle provenant de la population générale).

Le Tableau XIII présente les conventions permettant de qualifier les niveaux d'accord selon qu'on les considère faible, léger ou bon, jusqu'à presque parfait et parfait. Malheureusement, plusieurs études ne mentionnent pas la convention utilisée par les auteurs. Dans ces cas précis, une convention par défaut sera utilisée pour les analyses : celle d'Anastasi pour le coefficient alpha de Cronbach (cohérence interne) et celle de Landis et Koch pour le coefficient Kappa. Pour être en mesure de faire des comparaisons dans le présent projet, ces dernières seront utilisées, peu importe la convention mentionnée dans l'étude. Cela dit, les coefficients de Cronbach et Kappa se retrouvent dans la très grande majorité des études recensées.

Tableau XIII : Convention des méthodes d'évaluation de la validité et de la fidélité

INDICE	K			Y		r	α	K-R
CONVENTION	Landis et Koch (1977) Fleiss (1981)	Cicchetti (1994)	Bartko et Carpentier (1976)	Landis et Koch (1977)	Fleiss (1981)	Pearson (1896)	Anastasi (1986)	Cicchetti (1994)
FAIBLE	0.00-0.20-	Moins de 0.40	Moins de 0.40	Moins de 0.40	Moins de 0.40	0.10-0.30		Moins de 0.70 inacceptable
LÉGER	0.21-0.40							
CORRECT	0.41-0.59			0.41-0.59	0.41-0.59			0.71-0.79
BON	0.41-0.60	0.60-0.74	0.41-0.59	0.60-0.74	0.60-0.74	0.30-0.50		0.81-0.89
EXCELLENT	0.61-.80	Plus de 0.75	Plus de 0.60	Plus de 0.75	Plus de 0.75	0.50-1.0	Plus de 0.80	Plus de 0.90
PRESQUE PARFAIT	0.81-1.00						Plus de 0.90	

2.2.2 Mesures de la validité

Dans la présente SS, la validité désigne le degré d'accord entre le diagnostic posé et la «réalité». Selon Blacker et Endicott (2008), la validité est une notion en partie théorique (versus empirique pour la fidélité) puisque, pour plusieurs construits en psychiatrie, il n'y a pas de vérité absolue. Malgré tout, certaines évaluations de la validité apportent un éclairage dont on ne peut se passer. La validité peut être évaluée à l'aide d'un standard ou « étalon » (*gold standard*). Mais encore faut-il que l'on s'entende sur le choix de ce *gold standard*. De plus, Backer et Endicott précisent que les frontières des différents types de validité sont floues l'un par rapport à l'autre et que plusieurs auteurs peuvent les confondre. Dans les lignes qui suivent, seuls les types de validité effectivement recensés sont présentés.

Validité concordante : présence d'une forte corrélation avec d'autres tests (même type de mesure) mesurant le même phénomène (même caractéristique ou trait). Elle est en lien avec la validité de construit.

Validité concurrente : degré d'accord face à une mesure de référence ou étalon (*gold standard*). Par exemple, est-ce que l'utilisation d'un instrument par des non-cliniciens arrive au même résultat que l'évaluation de cliniciens expérimentés ? Ou encore, les résultats d'un instrument en lien avec un autre bien établi. Blacker et Endicott (2008) mentionnent que les versions adultes du K-SADS, le SADS (Endicott et Spitzer, 1978), ou encore le SCID (First, Spitzer, Gibbon M. et Williams, 2002) sont souvent les mesures étalons utilisées pour tester la validité de nouveaux instruments. Rathman (2008) rapporte qu'en 2008, le K-SADS était considéré comme une mesure étalon en regard de la recherche sur les troubles bipolaires chez les jeunes (Nottelmann, 2001). Malgré le fait qu'il n'y ait pas de consensus sur une mesure étalon pour les évaluations (Kazdin, 2005), l'entrevue par des cliniciens expérimentés a été retenue pour ce type de validité. Lorsque les instruments sont mis à l'épreuve avec un instrument semblable, les résultats sont classés sous le titre de la validité concordante.

Validité convergente : corrélation entre diverses mesures d'une même caractéristique (trait ou phénomène) par différentes méthodes (tests, observation, etc.). Blacker et Endicott (2008) donnent en exemple l'évolution de la maladie, les antécédents familiaux et la réponse au traitement. Elle est en lien avec la validité de construit.

Validité discriminante : capacité de l'instrument à différencier les individus ne répondant pas au construit recherché. Elle peut être vérifiée à l'aide d'indicateurs externes.

Les mêmes conventions permettant de qualifier les niveaux d'accord de la fidélité sont également utilisées pour évaluer le degré de validité (voir le Tableau XIII, p.73).

3. Analyse par type de fidélité et de validité

Il a déjà été mentionné que, selon Blacker et Endicott (2008), l'évaluation de la fidélité à l'aide du test-retest est plus près de la réalité sur «le terrain», où plus d'un intervieweur effectueront la passation et à divers moments. Il semble que plusieurs auteurs soient du

même avis puisque le Tableau XIV (p.77), qui résume les indices de la validité et de la fidélité recensés, montre que le test-retest est le plus utilisé. Il s'agit alors d'évaluer la stabilité dans le temps des résultats d'un instrument. Différentes variantes sont recensées. Plusieurs études [19, 23, 24, 30, 31, 38, 40, 41, 46, 53] utilisent un protocole à quatre intervieweurs travaillant à l'aveugle chaque fois; une méthode qui entraîne peu de biais. Six études [3, 5, 13, 16, 47, 49] utilisent deux intervieweurs différents, mais peu de précisions sont données. S'agit-il d'un intervieweur au temps 1 et d'un autre au temps 2? Ou d'un intervieweur pour le parent et d'un autre pour l'adolescent(e)? Une seule étude [49] est explicite à ce sujet et elle précise qu'il s'agit de la première option. Une étude [27] mentionne qu'un seul intervieweur a été utilisé tandis que trois [1, 9, 14] ne donnent aucune indication à ce sujet. Dix études [5, 13, 23, 24, 30, 31, 40, 41, 46, 49] ont évalué la fidélité test-retest pour le parent et l'adolescent(e), augmentant ainsi les possibilités de comparaison : les passations en dyade permettent d'analyser le degré d'accord entre l'adolescent(e) et son parent. Par ailleurs, sept études [1, 3, 17, 19, 38, 47, 53] n'ont interviewé que l'adolescent(e) et trois [9, 14, 27] seulement le parent.

Pour évaluer la fidélité d'un instrument, le test d'accord interjuge est également répandu et ce, avec plusieurs variantes. Cinq études [20, 21, 39, 47, 54] utilisent deux intervieweurs différents pour réaliser deux passations avec l'adolescent(e) et/ou le parent. Dans deux autres études [13, 37], les résultats d'intervieweurs non-cliniciens sont comparés avec ceux de pédopsychiatres. Pour quatre autres cas [1, 12, 18, 44], le premier clinicien effectue une entrevue, qui est enregistrée, et le second effectue son évaluation en visionnant la bande vidéo. Dans une autre étude [55], les résultats des parents sont comparés à ceux des adolescent(e)s, les entrevues ayant été réalisées par un même intervieweur (au risque d'introduire un biais, puisque l'information du premier entretien est connue). On peut donc comparer les résultats obtenus par deux intervieweurs ou entre deux participants.

La dernière mesure de la fidélité répertoriée, la cohérence interne, est la moins utilisée. Sept études [12, 24, 40, 44, 46, 56, 57] l'ont vérifiée au niveau d'échelles visant à produire un diagnostic en particulier.

En ce qui concerne la validité, la forme concurrente est la plus testée, une situation qui se retrouve dans 13 études. Il s'agit alors de confronter les résultats de l'instrument à l'étude à ceux d'un test étalon. Il a déjà été mentionné que l'entrevue diagnostique menée par un clinicien expérimenté est considérée comme une mesure étalon dans le présent projet. Encore une fois, plusieurs variantes sont recensées. La majorité des études [16, 21, 39, 41, 47, 53, 54] utilise de l'information provenant du parent et de l'adolescent(e). Certaines comparent le résultat de l'instrument au diagnostic reçu à la fin d'une hospitalisation [21, 22] ou à l'admission [42]. On cherche alors à savoir si un instrument administré par un non-spécialiste mène à des résultats équivalents donc «valides».

Vient ensuite la validité convergente qui a été étudiée à 10 reprises. Dans la quasi-totalité des études concernées, des échelles mesurant un type de comportements ou de sentiments ont été utilisées afin de vérifier si l'instrument à l'étude arrivait à des résultats semblables. Quelques-unes ont aussi utilisé des observations [32, 14], questionnaires autorévélés [1, 43] ou journaux de bord [1] pour répertorier les comportements et leur fréquence.

La validité concordante a été examinée dans six études [4, 6, 7, 11, 36,48] (dont l'étude 4 avec deux instruments différents), l'objectif étant de vérifier si deux instruments semblables arrivaient aux mêmes résultats. Cinq de ces études utilisent pour des fins de comparaison un des 11 instruments retenus pour le présent projet, soit le CAPA, le DICA, le DISC et le K-SADS.

Finalement, la validité discriminante, mise à l'épreuve dans six études [5, 8, 14, 15, 29, 51], désigne la capacité d'un instrument à distinguer les adolescent(e)s qui répondent à un diagnostic en particulier. Pour ce faire, on compare deux groupes, dont un groupe contrôle, sous l'angle du nombre de diagnostics, du nombre de symptômes et des taux de prévalence. Dans un cas [29], l'observation est utilisée.

Enfin, il est à noter que pour seulement trois instruments, soit le ChIPS, le DISC et le K-SADS, les sept formes de fidélité et validité ont été examinées.

Tableau XIV : Nombre d'études par mesures de validité et de fidélité recensées par instrument

Instrument	Fidélité			Validité			
	Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
ADI (1)	1	1	–	–	1	1	–
APS (0)	–	–	–	–	–	–	–
CAPA (3)	–	2	–	1	–	–	1
ChIPS (6)	1	1	1	3	3	1	1
DICA (9)	4	4	–	–	2	3	2
DISC (15)	2	8	2	2	4	1	1
GAIN (1)	–	–	–	–	–	1	–
ISCA (0)	–	–	–	–	–	–	–
K-SADS (10)	4	4	2	1	3	3	1
PADDI (2)	–	–	2	–	–	–	–
PAI-A (0)	–	–	–	–	–	–	–
Total (47)	12	20	7	6¹	13	10	6

Note : une même étude peut avoir utilisé plus d'une mesure.

(X) : nombre d'études recensées, retenues à la suite de l'analyse du QUADAS-2.

1- Une même étude a testé deux instruments.

4. Analyse par diagnostic

La prochaine section présente les résultats des études qui ont mis à l'épreuve certaines qualités psychométriques des instruments en fonction des différents diagnostics ciblés. Le Tableau XV présente les différents diagnostics mis à l'épreuve par instrument.

Il s'avère que les diagnostics de TC et TDA/H sont les plus souvent évalués tandis que l'ÉSPT est le moins souvent évalué. On remarque également qu'il n'y a que quatre instruments, soit le CAPA, le ChIPS, le DICA et le K-SADS, pour lesquels on dispose de données pour les six diagnostics retenus. Malgré le grand nombre d'études retenues impliquant le DISC, on note qu'elles n'ont pas inclus le diagnostic d'ÉSPT. Finalement, pour l'ADI et le GAIN, seul le diagnostic de TUS a été testé.

Tableau XV : Nombre d'études par diagnostics ayant été mis à l'épreuve par instrument

Instrument	TUS	TAG	ÉDM	TC	TDA/H	ÉSPT
ADI (1)	1	–	–	–	–	–
APS (0)	–	–	–	–	–	–
CAPA (3)	1	1	1	2	1	1
ChIPS (6)	5	6	6	6	6	4
DICA (9)	1	5	5	6	9	3
DISC (15)	2	7	9	13	13	–
GAIN (1)	1	–	–	–	–	–
ISCA (0)	–	–	–	–	–	–
K-SADS (10)	3	7	8	8	7	4
PADDI (2)	2	–	2	2	–	2
PAI-A (0)	–	–	–	–	–	–
Total(47)	16	26	31	37 ¹	36 ¹	14

Note : une même étude peut avoir mis à l'épreuve plus d'un instrument et plus d'un diagnostic.

(X) : nombre d'études recensées, retenues à la suite de l'analyse du QUADAS-2.

1- Angold (2012) a mis à l'épreuve deux instruments.

Chaque sous-section débute avec la référence au tableau en annexe contenant les données chiffrées des différentes études. Chaque diagnostic a donc son tableau (en annexe) où sont compilés les détails des études à leur sujet, incluant les résultats. Chacun sera ici présenté en débutant avec la mesure la plus utilisée pour l'évaluer. Ensuite, les résultats des autres

mesures recensées seront présentés suivant l'ordre du tableau résumé de chaque diagnostic, soit les tableaux XVI à XXI.

4.1 Troubles liés à l'utilisation des substances psychoactives (TUS)

Référence Tableau 1 : Troubles liés à l'utilisation des substances psychoactives (TUS) (annexe B)

Seize études ont mis à l'épreuve la fidélité ou la validité d'un diagnostic de TUS (Tableau XV, p.78) que ce soit à l'aide de l'ADI, du CAPA, du ChIPS, du DICA, du DISC, du GAIN, du K-SADS ou du PADDI. La particularité du diagnostic de TUS est la multitude des formes qu'il peut prendre, les études n'utilisant pas toutes les mêmes sous catégories (ex. : abus versus dépendance; type de substance visée). Par exemple, l'étude 1 teste l'ADI en lien avec les diagnostics d'abus d'alcool, dépendance à l'alcool, abus de cannabis et dépendance au cannabis. L'étude 54 ne teste le K-SADS qu'en lien avec les diagnostics d'abus de drogues. Quant à l'étude 16, elle porte sur le DICA et l'abus de colle ou de drogues de rue. De plus, la période de référence pour le diagnostic peut aussi être différente : à vie, actuel ou au cours de la dernière année.

Tableau XVI : Nombre d'études par types de fidélité et validité pour les TUS

Fidélité			Validité			
Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
3	6	2	3	6	2	1

Comme le démontre le Tableau XVI, la validité concurrente (ex æquo avec la mesure du test-retest de la fidélité) est la qualité psychométrique la plus utilisée pour tester le diagnostic de TUS. Il s'agit alors de savoir si l'instrument arrive au même diagnostic que celui posé par des cliniciens expérimentés. Selon les différents résultats des quatre études où on a été en mesure de produire une valeur Kappa²² pour la validité concurrente [1, 6, 54, 53], les résultats démontrent de très bons résultats pour l'ADI, le ChIPS et le K-SADS, les valeurs du

²² La valeur K ne peut être produite lorsqu'il y a 100% d'accord sur l'absence de diagnostic.

Kappa se situant entre 0,67 et 0,82. Les résultats plus faibles pour une étude [53] portant sur le K-SADS sont en lien avec l'obtention d'un diagnostic actuel et à vie, où les valeurs sont respectivement de 0,34 et 0,23.

La fidélité interjuge a été examinée à trois reprises [1, 54, 55], pour l'ADI et le K-SADS. On cherche, avec cette mesure, à connaître la stabilité de l'instrument d'un intervieweur à l'autre. Il va sans dire que la formation des intervieweurs à l'instrument est importante à ce niveau. Or, d'une étude à l'autre, le nombre d'intervieweurs et leur statut peuvent varier. L'étude 1 en a deux, le deuxième se référant à l'enregistrement vidéo du premier. L'étude 54 en utilise également deux, mais sans usage de la vidéo et à 14 jours d'intervalle. Dans l'étude 55, un seul évaluateur est mis à contribution à deux reprises le jour même puisqu'on cherche à comparer les résultats entre parent et adolescent(e) et non entre intervieweurs. Dans l'ensemble, les résultats Kappa varient entre 0,07 et 1,00 (accord très faible à parfait). L'ADI obtient des résultats allant d'excellent à presque parfait (K entre 0,68 et 0,86). Le K-SADS obtient des résultats allant d'un extrême à l'autre: 0,07 et 1,00.

Six études [1, 3, 9, 17, 23, 38] ont examiné la fidélité test-retest d'un des instruments suivants : ADI, CAPA, DICA et DISC. La stabilité de l'instrument a été mise à l'épreuve pour des intervalles variant de cinq à 14 jours, ce qui est très acceptable. Sauf pour les études 1 et 9, qui ne spécifient pas combien d'intervieweurs ont été mis à contribution, les autres études ont réduit au maximum les risques de biais en utilisant des intervieweurs différents au temps 1 et au temps 2. Les taux d'accord aux temps 1 et 2 sont particulièrement élevés lorsque l'information provient de l'adolescent(e), avec des valeurs Kappa entre 0,48 et 1,00 (et quelques valeurs Y entre 0,59 et 0,67). Lorsqu'elle provient d'un parent, les valeurs descendent sous la barre du 0,46. Peu importe l'instrument, les résultats sont considérés de bons à parfaits lorsque l'information provient de l'adolescent(e). Seule l'étude 9 s'est servi du pourcentage en testant la passation en face à face et ensuite au téléphone. L'accord fut alors parfait (100%).

Vient ensuite la cohérence interne, examinée dans deux études [56, 57] s'étant intéressées au PADDI. Les résultats sont excellents (α entre 0,79 et 0,89), ce qui signifie que les items composant les échelles générant les diagnostics de TUS sont cohérents entre eux.

Les trois dernières propriétés psychométriques concernent la validité. La première, la validité concordante, soit l'accord observé entre l'instrument étudié et un autre semblable, a été examinée dans trois études [6, 7, 11] portant toutes sur le ChIPS. Seules deux d'entre elles [6, 11] ont pu produire des valeurs K. Elles se situaient entre 0,63 et 0,66 avec l'information des adolescent(e)s, un taux d'accord excellent.

Vient ensuite la validité convergente, examinée lorsque les résultats de l'instrument à l'étude sont comparés à ceux de diverses échelles ou observations. Deux études [1, 43] l'ont évaluée : une pour l'ADI et l'autre pour le GAIN. Cette fois les méthodes sont semblables : les intervalles varient de sept à 14 jours, les informations proviennent de l'adolescent(e) et les instruments de référence sont semblables. Par contre, pour l'ADI, des mesures supplémentaires ont été reprises à six et 12 mois. Pour le GAIN, l'utilisation de l'instrument *Form 90*²³ a généré des résultats concluants ($p < 0.01$), nombre de jours de consommation d'alcool ($r_s=0.84$), nombre maximum de consommations d'alcool ($r_s=0.75$), nombre de jours d'utilisation de cannabis ($r_s=0.84$) et nombre maximum de consommations de joints de cannabis ($r_s=0.75$). Quatre mesures différentes ont été utilisées avec l'ADI, calculées à l'aide du *correlational (eta) analysis*²⁴. Des données ont été obtenues séparément pour l'alcool et le cannabis. À l'aide du DUF²⁵, les taux varient entre 0,28 et 0,68. Pour la cotation du clinicien²⁶, entre 0,60 et 0,71. Avec le Heilman's warning²⁷, entre 0,73 et 0,75. Finalement avec le PICS²⁸, entre 0,65 et 0,70. Les résultats montrent que l'ADI et le GAIN évaluent adéquatement des construits reliés aux TUS lorsqu'on les confronte à des mesures équivalentes.

²³ Questionnaire auto révélé de la consommation (Miller, 1996; Miller et Boca, 1994)

²⁴ *Correlational analysis* : indique le degré de corrélation entre variables.

²⁵ *Drug Use Frequency*

²⁶ Évaluation de la sévérité de la problématique selon le clinicien évaluateur.

²⁷ *Heilman's warning signs Checklist* (Heilman, 1973)

²⁸ *Personal Involvement with Chemicals Scale* dans PEI (Winters et Henly, 1989)

Finalement, une étude [8] a mis à l'épreuve la valeur discriminante du ChIPS et posé la question suivante : l'instrument est-il en mesure de distinguer les adolescent(e)s avec un TUS des autres ? Le nombre de diagnostics, ainsi que le taux de détérioration (famille, école, amis) sont utilisés pour le vérifier et ce, à l'aide d'un groupe contrôle. Par contre, l'étude démontre que les taux diagnostiques sont plus élevés pour le groupe contrôle que pour le groupe clinique (diagnostic 4,6% versus 0%). Les résultats ne sont donc pas concluants.

La qualité psychométrique la plus testée pour le diagnostic de TUS est donc la fidélité test-retest et la validité concurrente, avec six études chacune. La moins testée est la validité discriminante, avec une seule étude. Les meilleurs et les moins bons résultats se retrouvent avec la fidélité interjuge, puisque des taux extrêmes sont répertoriés : de 0,07 et 1,00. De façon globale, c'est la mesure de la fidélité de la cohérence interne qui obtient les meilleurs résultats.

4.2 Trouble anxieté généralisé (TAG)

Référence Tableau 2 : Trouble anxieté généralisé (TAG) (annexe C)

Vingt-six études ont mis à l'épreuve la validité ou la fidélité du diagnostic du TAG (Tableau XIV, p.77) à l'aide du CAPA, du ChIPS, du DICA, du DISC ou du K-SADS.

Tableau XVII : Nombre d'études par types de fidélité et validité pour les TAG

Fidélité			Validité			
Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
6	13	3	4	7	3	1

La propriété psychométrique la plus examinée est la fidélité test-retest, objet de 13 études [3, 9, 13, 14, 17, 19, 23, 24, 27, 38, 41, 46, 49]. Du temps 1 au temps 2, les intervalles varient alors d'un à 21 jours, ce qui est très acceptable. Deux études [41, 46] ont prévu un intervalle très court, soit un jour (pour une partie de l'échantillon de l'étude 41) ou trois (pour l'échantillon de l'étude 46). Lorsque l'intervalle est aussi court, les participants peuvent se souvenir de leurs réponses, ce qui fausse quelque peu les indices de fidélité. La moitié des

études [19, 23, 24, 38, 41, 46] a eu recours à quatre intervieweurs différents pour chaque passation (parent et adolescent(e); temps 1 et 2), ce qui réduit au maximum les risques de biais. Il en va de même avec les études [3, 17] qui ne font participer que les adolescent(e)s et qui utilisent deux intervieweurs (temps 1 et 2). Une étude [9] ne précise pas l'intervalle du temps 1 au temps 2, ni le nombre d'intervieweur utilisé.

Globalement, les valeurs K varient entre 0,24 et 0,79. Contrairement à ce qui a été constaté pour le diagnostic de TUS, la fidélité des informations provenant des adolescent(e)s n'est pas plus élevée que celle des parents. Au contraire, elle est plus faible. Les résultats des données recueillies auprès des adolescent(e)s varient entre 0,28 à 0,54, ce qui démontre une corrélation de légère à bonne. Un résultat de 0,79 (excellent) a été obtenu dans l'étude 3 s'intéressant au CAPA. Quant aux résultats provenant de l'information recueillie auprès des parents, ils varient de 0,39 à 0,79 et sont globalement plus élevés (bon à excellent). Lorsqu'il y a utilisation combinée de l'information provenant de plusieurs participants (parent, adolescent(e) et parfois intervieweur), les indices varient de 0,24 à 0,78. L'étude 9 obtient un pourcentage d'accord de 91,7% (passation en face à face vs au téléphone).

La fidélité interjuge a fait l'objet de six études [12, 13, 37, 39, 44, 55], qui se caractérisent par trois variantes intéressantes. Une première étude [12], réalisée avec le ChIPS, a évalué l'accord interjuge obtenu auprès d'un échantillon d'adolescent(e)s autistes et le résultat est quasi parfait (valeur k de 0,86). Deux études [13, 37] menées avec le DICA et le DISC, ayant eu recours à quatre intervieweurs, ont eu de moins bons résultats en comparant l'évaluation de cliniciens, non-cliniciens et spécialistes, tels que des pédopsychiatres. La première a porté sur un échantillon de la communauté et obtenu un coefficient de 0,06 aux temps 1 et 2, ce qui est faible. La deuxième, avec un échantillon clinique, a obtenu des résultats de bons à excellents (0,64 à 0,77) en comparant les données de deux non-cliniciens. Les taux d'accord baissent substantiellement lorsque l'on compare les données de non-cliniciens à celles de spécialistes, variant de 0,20 à 0,66. Une troisième étude [55] avec le K-SADS, a évalué l'accord avec l'adolescent(e) selon que le parent avait reçu un diagnostic d'ÉDM ou non (groupe contrôle). Les valeurs K étant de 0,17 et 0,24, l'accord est léger. De meilleurs

résultats sont obtenus pour le ChIPS et le K-SADS, lorsque la deuxième cotation se fait à l'aide de l'enregistrement vidéo ayant servi à la première [12, 44], l'indice K étant de 0,86 et 0,85.

Au niveau de la cohérence interne, trois études [12, 24, 44] ont testé les échelles diagnostiques en lien avec un TAG. Les auteurs de la première étude, portant sur le ChIPS, ont choisi la mesure K-R. L'indice étant de 0,65, il s'avère que la cohérence interne est sous les standards attendus. La deuxième étude menée avec le DISC a obtenu des résultats excellents (α entre 0,70 et 0,80). La troisième étude, s'intéressant au K-SADS, a testé deux échelles diagnostiques : la première en lien avec les symptômes au cours des 12 derniers mois et l'autre au cours de la dernière semaine. À nouveau, les résultats sont de très bons à excellents (α entre 0,76 et 0,81).

Les quatre dernières qualités psychométriques sont en lien avec la validité. Tout d'abord, la validité concordante a été vérifiée par quatre études [6, 7, 11, 12] et seulement pour le ChIPS. Les deux premières études l'ont comparé au DICA, la troisième au K-SADS et la quatrième au CASI²⁹. Les intervalles sont d'un à trois jours et l'information provient de l'adolescent(e), sauf pour la quatrième, où elle provient du parent. La valeur des indices varie de 0,29 à 0,53, ce qui indique un accord allant de léger à bon.

La validité concurrente a été mise à l'épreuve par sept études [6, 7, 8, 16, 41, 47, 53] et pour trois instruments (ChIPS, DICA, K-SADS). Les modalités sont assez diversifiées : les intervalles varient d'un jour à 30. Fait intéressant, l'étude [53] n'ayant prévu qu'un jour d'intervalle n'obtient pas de meilleurs résultats que celle [8] ayant planifié 30 jours : K de 0,28 pour la première (diagnostic à vie et actuel à partir de l'information combinée du parent et de l'adolescent(e)) et 0,48 pour la seconde (à partir de l'information du parent). Les résultats tirés de l'information obtenue auprès de l'adolescent(e) sont globalement légers (0,21, 0,23, 0,27 et 0,46) avec une exception à 0,74 (comparaison faite à partir de toute l'information contenue au dossier [7]). Les résultats sont globalement meilleurs lorsque l'information est obtenue auprès du parent : entre 0,42 et 0,60 (bons). Lorsqu'on combine l'information

²⁹ *Child and Adolescent Symptom Inventory* (Gadow et Sprafkin, 2002)

obtenue auprès du parent et de l'adolescent(e), les valeurs K varient de 0,26 à 0,51, à 0,78 (légers à excellents). Dans trois études, les auteurs n'ont pas précisé l'intervalle entre la passation de l'instrument à l'étude et l'évaluation clinique [6, 16, 47] : les résultats du K sont respectivement de 0,46, 0,21 et 0,78.

Les études 14, 45 et 49 ont pour leur part testé la validité convergente d'instruments d'évaluation du TAG : le DICA et le K-SADS ont vu leurs résultats confrontés à ceux de diverses échelles. L'intervalle de passation est de 15 jours dans un cas [45], tandis que deux études ne le mentionnent pas. Les résultats sont concluants pour le K-SADS lorsque confrontés au SCARED-P et C³⁰ (version parent : ANX/POS = 25.5 ± 14.6; ANX/NEG = 11.4 ± 11.4; $t_{63} = 4.32, P < .0001$, version enfant : ANX/POS = 26.1 ± 14.7; ANX/NEG = 12.6 ± 9.3; $t_{56.4} = 4.47, P < .0001$, [49]) et au CBCL³¹ (ANX/POS = 66.8 ± 11.5; ANX/NEG = 58.1 ± 13.7; $t_{63} = 2.76, P < .01$, [49] et léger avec r de 0,30 [45]). Pour le DICA, ils ne le sont pas lorsque comparés au OCHS-R³² : K de 0,31.

Une seule étude [8] a testé la validité discriminante d'un instrument d'évaluation du TAG. Les résultats sont concluants pour le ChIPS lorsqu'on compare le nombre de diagnostics (0% versus 9,7%) et le niveau de détérioration du fonctionnement familial, scolaire et social (2,3% versus 32,3%) du groupe contrôle de la communauté et de l'échantillon clinique. Le ChIPS serait donc en mesure de discriminer les individus avec un TAG et ceux sans.

La propriété psychométrique la plus examinée pour le diagnostic du TAG est la fidélité test-retest, mais c'est la mesure interjuge qui obtient les meilleurs taux d'accord avec une valeur K de 0,86. Par contre, cette dernière obtient également le taux le plus faible : 0,17. La propriété psychométrique la moins testée est celle de la validité discriminante. Globalement, c'est la mesure de la fidélité de la cohérence interne qui serait la plus robuste, exception faite d'un taux sous les standards attendus (K-R de 0,65).

³⁰ *Screen for Child Anxiety Related Emotional Disorders, parent and child versions* (Birmaher et al., 1997)

³¹ *Child Behavior Checklist Internalizing* (Achenbach et Edelbrock, 1983)

³² *Revised Ontario Child Health Study Scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

4.3 Épisode dépressif majeur (ÉDM)

Référence Tableau 3 : Épisode dépressif majeur (ÉDM) (annexe D)

Pour le diagnostic d'ÉDM, 31 études ont mis à l'épreuve, à l'aide de toutes les mesures de la fidélité et de la validité, les instruments suivants : CAPA, CHIPS, DICA, DISC et PADDI (Tableau XV, p.78).

Tableau XVIII : Nombre d'études par types de fidélité et validité pour les EDM

Fidélité			Validité			
Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
7	12	6	5	10	5	3

Tel que le démontre le Tableau XVIII, la propriété psychométrique la plus examinée est la fidélité test-retest, ayant fait l'objet de 12 études [3, 9, 13, 14, 19, 23, 30, 38, 40, 41, 46, 49]. Du temps 1 au temps 2, les intervalles varient de moins de trois jours à 21 jours, ce qui est très acceptable. Six études [19, 23, 30, 40, 41, 46] appliquent le protocole de quatre intervieweurs différents lorsque l'adolescent(e) et un parent sont rencontrés. Quatre études [3, 13, 38, 49] n'en prévoient que deux, seul l'adolescent(e) ayant été interviewé(e) ou l'évaluation reposant sur la combinaison des informations [49]. Pour le CAPA, les résultats sont quasi parfaits : K de 0,90 avec l'information provenant de l'adolescent(e) [3]. Deux études [13, 14] ont testé la fidélité du DICA avec l'information provenant de l'adolescent(e) (K de 0,45, ce qui est bon) et du parent (0,31 et 0,70, soit de léger à excellent). Cinq études [23, 30, 38, 40, 41] ont pour leur part analysé la fidélité du DISC. Les valeurs K obtenues à partir de l'information de l'adolescent(e) varient de 0,15 (léger) à 0,77 (excellent). Avec l'information provenant du parent, les valeurs passent de 0,31 à 0,82, donc de léger à excellent également. Les valeurs K provenant des informations combinées parent-adolescent(e) varient pour leur part de 0,48 à 0,71. Finalement deux études [46, 49] ont évalué la fidélité du K-SADS : les valeurs K produites à partir d'informations combinées sont de 0,54 pour l'étude 46, 0,90 (diagnostic actuel) et 1,00 (diagnostic à vie) pour l'étude 49.

La fidélité interjuge a été testée à sept reprises [12, 13, 20, 39, 44, 54, 55]. Dans une étude [12] portant sur le ChIPS, un premier intervieweur a interrogé l'adolescent(e) et le second a effectué son évaluation à l'aide de l'enregistrement vidéo. La valeur du K est bonne : 0,44. Deux études [13, 20] ont testé le DICA avec des intervalles très différents soit : pour la première, entre 10 et 20 jours et pour la seconde, le jour même. L'étude 13 a tenté de vérifier si le taux d'accord serait différent en faisant varier le statut de l'intervieweur. Au temps 1, l'accord entre deux cliniciens ayant interviewé le parent et l'adolescent(e) se caractérise par un K de -0,05, ce qui est négatif. Au temps 2, l'accord entre des cliniciens et spécialistes (pédopsychiatres) n'est guère meilleur : K de 0,13. Malgré le délai quasi inexistant prévu dans la seconde étude portant sur le DICA, le coefficient obtenu pour l'accord parent-adolescent(e) n'est que de 0,11. Une étude [39] a vérifié la fidélité du diagnostic d'ÉDM avec le DISC : avec un intervalle de moins de 15 jours et deux intervieweurs, l'accord entre les parents et adolescent(e)s du groupe contrôle est de 0,18, comparativement à 0,25 pour ceux du groupe clinique. Finalement, trois études [44, 54, 55] ont testé le K-SADS. La première, avec enregistrement vidéo, obtient un excellent accord lorsque l'information provient d'un parent et de l'adolescent(e) : K de 0,80. Il en va de même avec l'étude 54 : deux intervieweurs obtiennent un coefficient K de 0,84 en interviewant le parent et l'adolescent(e). Par contre, lorsque l'on compare l'accord parent-adolescent(e) de cette étude, la valeur du coefficient chute à 0,47. La dernière étude a pour sa part tenté de cerner le taux d'accord entre un parent (la mère) et adolescent(e) : le coefficient est alors de 0,30.

La cohérence interne des items composant les échelles diagnostiques a été testée par six études [12, 40, 44, 46, 56, 57]. De façon générale, peu importe le nombre d'items des échelles, les résultats sont excellents. Pour le ChIPS [12], le résultat K-R est de 0,83. Pour le DISC [40], les coefficients α sont de 0,88 lorsque l'information provient du parent et de 0,85 lorsque c'est l'adolescent(e) qui participe. Pour le K-SADS, l'étude 44 avec deux évaluateurs pour évaluer deux échelles diagnostiques, (la première en lien avec les symptômes au cours des 12 derniers mois et l'autre au cours de la dernière semaine), obtient des résultats qui varient de très bons à excellents (α entre 0,78 et 0,89). Tandis que les valeurs du coefficient α

de la seconde étude [46] varient entre 0,68 et 0,84. Pour le PADDI, les résultats sont aussi excellents : α de 0,84 [56] et 0,92 [57].

En tout 23 études ont testé une ou l'autre des formes de validité du diagnostic d'ÉDM. Les études 6, 7, 11, 12 et 36 ont mis à l'épreuve la validité concordante de ce dernier, mais seulement quatre ont pu produire des valeurs K. L'étude 6, avec l'information provenant de l'adolescent(e), a comparé le ChIPS au DICA à sept jours d'intervalle. Le résultat du K à 0,38 est considéré comme léger. Le ChIPS a aussi été comparé au K-SADS dans l'étude 11, toujours en utilisant l'information provenant de l'adolescent(e), à trois jours d'intervalle : K de 0,49. L'étude 12 le compare au CASI³³, avec l'information provenant du parent et de l'adolescent(e) au cours de la même journée : le résultat est bon (K de 0,44). Finalement, le DISC a été comparé au CAF³⁴ et les résultats sont forts semblables : la valeur K est de 0,36 à partir de l'information du parent, 0,39 à partir de celle de l'adolescent(e) et 0,41 en combinant les deux.

La validité concurrente a été testée par 10 études. Pour le ChIPS, une seule [6] des trois études a pu produire une valeur K : l'intervalle est inconnu, l'information provient de l'adolescent(e) et le résultat est de 0,39 (léger). L'étude 16 a testé la validité concurrente du DICA : l'intervalle est également inconnu et lorsque l'information provient de l'adolescent(e), le résultat est meilleur, bien que léger (K de 0,31). Lorsqu'elle provient du parent, le résultat est de 0,13 et si on combine les deux, de 0,25. Quatre études ont testé le DISC, dont deux en lien avec le diagnostic posé au moment du congé du traitement [22] et l'autre avec celui posé au moment de l'admission [42]. Les intervalles sont donc bien différents et passent de 135,6 à 15 jours. L'information provient de l'adolescent(e) et les valeurs du K sont respectivement de 0,19 et 0,17. Lorsque l'évaluation clinique se fait à partir des informations au dossier, les résultats grimpent quelque peu : de 0,33 (information du parent) à 0,54 (information de l'adolescent(e)). La dernière [41] étude a obtenu des valeurs K de 0,60 (information du parent) et de 0,79 (information de l'adolescent(e)). Finalement, deux études [53, 54] ont mis

³³ *Child and Adolescent Symptom Inventory* (Gadow et Sprafkin, 2002)

³⁴ *Clinical Assessment Form* (Piacentini, Shaffer, Fisher, Schwab-Stone et al., 1993)

à l'épreuve le K-SADS. À nouveau, les protocoles de recherche sont fort différents avec des intervalles de un jour à 2 mois après la fin du traitement. L'information combinée (parent-adolescent(e)) produit par contre des résultats excellent et quasi parfait (0,74 et 0,89) pour les deux études.

Les études 14, 20, 45, 49 et 51 ont pour leur part testé la validité convergente de quatre instruments. Les résultats ne sont pas concluants lorsque le DICA [14] est confronté au OCHS-R³⁵ (r de 0,05). Une corrélation légère est présente lorsque le DISC [20] est comparé au PIC³⁶ (K de 0,24). Trois études ont testé la validité convergente du K-SADS. Les deux premières [45, 49] obtiennent des résultats concluants lorsque comparés au CBCL³⁷ (léger avec un taux r de 0,30 et $DEP/POS = 68.7 \pm 10.0$; $DEP/NEG = 58.4 \pm 13.6$; $t_{62.4} = 3.54$, $P < .001$), avec les informations du parent et de l'adolescent(e). La troisième [51] utilise diverses informations provenant de plusieurs échelles : le tout est concluant autant avec l'information provenant de l'adolescent(e) (YSR *internalizing*³⁸, YSR *anxious depressed*¹⁴, SDQ-S *emotional symptoms*³⁹, CASS:L: *emotional problems*⁴⁰, BYI *depression*⁴¹, CDI *total*⁴²) que du parent (CBCL *anxious-depressed*⁴³).

La dernière propriété psychométrique est la validité discriminante. Trois études [8, 14, 51] l'ont mise à l'épreuve. Pour le CHIPS [8], les résultats sont concluants lorsqu'un groupe contrôle est comparé à un groupe clinique : le nombre de diagnostics (2,3% versus 6,4%) et le niveau de détérioration (famille, école, amis : 2,3% versus 32,3%) du groupe clinique sont plus élevés. Par contre, pour les deux études qui ont testé la validité discriminante du DICA [14] et du K-SADS [51], les résultats ne sont pas concluants (retard développemental, diagnostics d'anxiété et de TC).

³⁵ *Revised Ontario Child Health Study scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

³⁶ *Personality Inventory for Children* (Lacher et Gdowski, 1979; Wirt, Lacher, Klindedinst et Seat, 1977)

³⁷ *Child Behavior Checklist* (Achenbach et Edelbrock, 1983; Achenbach, 1991a)

³⁸ *Youth Self-Report* (Achenbach, 1991b)

³⁹ *Strengths and Difficulties Questionnaire* (Obel et al., 2004)

⁴⁰ *Conners/Wells Adolescent Self-report Scales Long Version* (Conners et al., 1997)

⁴¹ *Beck Youth Inventories of Emotional and Social Impairment* (Beck, Beck et Jolly, 2001)

⁴² *Children's Depression Inventory* (Kovacs M., 1992)

⁴³ *Child Behaviour Checklist* (Achenbach, 1991a)

La propriété psychométrique la plus examinée en ce qui concerne le diagnostic de TAG, est la fidélité du test-retest. C'est également cette mesure qui obtient les meilleurs résultats avec des taux presque parfaits et parfaits, exception faite de quelques accords considérés faibles. C'est la mesure de la fidélité interjuge qui obtient le moins bon score, il est d'ailleurs négatif (k de -0,05). Globalement, la mesure la plus performante est à nouveau celle de la cohérence interne (fidélité) et la moins utilisée, celle de la validité discriminante.

4.4 Troubles des conduites (TC)

Référence Tableau 4 : Trouble des conduites (TC) (annexe E)

Le trouble des conduites est le diagnostic concomitant dont l'évaluation a été mise à l'épreuve par le plus grand nombre d'études : 37 (Tableau XV, p.78). Et ce avec six des neuf instruments : CAPA, ChIPS, DICA, DISC, K-SADS et PADDI.

Tableau XIX : Nombre d'études par types de fidélité et validité pour les TC

Fidélité			Validité			
Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
7	14	6	7	12	4	2

Une fois de plus, la fidélité test-retest représente la propriété psychométrique la plus souvent examinée (N = 14 études). Une seule étude [3] a testé le CAPA : à intervalle de cinq jours, l'information provenant à deux reprises de l'adolescent(e) a généré une valeur K de 0,55 (bon). Pour le ChIPS [9], seul un pourcentage d'accord est disponible, soit 83,3% : les auteurs ont alors comparé la passation en face à face à celle faite au téléphone. Ils n'ont mentionné ni l'intervalle, ni le nombre d'intervieweurs. Quatre études ont mis à l'épreuve l'échelle du DICA permettant de poser un diagnostic de TC. Les intervalles varient de cinq à 20 jours ; deux évaluateurs effectuent les passations [17, 19], même lorsque le parent et l'adolescent(e) sont interviewés [13] (ce qui peut introduire un biais). La quatrième étude

[14] ne mentionne pas le nombre d'évaluateurs impliqués. Les valeurs du coefficient K varient beaucoup, allant de 0,17 à 0,92 (léger à presque parfait). Lorsque l'information est recueillie auprès des parents, elles fluctuent à nouveau entre des valeurs extrêmes (0,21 à 0,87). Six études ont vérifié le DISC [23, 24, 31, 38, 40, 41], à l'aide d'intervalles semblables, se situant en majorité entre sept et 21 jours. Les valeurs du coefficient K varient entre 0,49 (bon) et 0,86 (presque parfait) lorsque l'information provient de l'adolescent(e); 0,43 et 0,87 lorsqu'elle émane du parent; 0,50 et 0,66 (excellent) lorsque les deux types d'information sont combinés. Finalement, la fidélité du K-SADS a été mise à l'épreuve par deux études : les résultats des informations combinées (parent-adolescent(e) [49], parent-adolescent(e)-évaluateur [46]) sont respectivement de 0,83 et 0,63. Les intervalles diffèrent toutefois quelque peu : 17,9 jours en moyenne dans l'une et moins de trois jours dans l'autre.

Sept études ont examiné la fidélité interjuge d'échelles évaluant le TC. L'étude 12, portant sur le ChIPS, compare un enregistrement vidéo à l'information provenant de l'adolescent(e). La valeur du coefficient K est excellente : 0,77. Deux études ont évalué la fidélité du diagnostic TC posé à l'aide du DICA [13, 21]. La première, avec des résultats dans la négative, tente de cerner la différence entre l'évaluation faite par des cliniciens (K de -0,02) et des pédopsychiatres (K de -0,02), en combinant l'information du parent et de l'adolescent(e), à intervalle de 10 à 20 jours. Les résultats sont peu concluants. La deuxième, impliquant deux intervieweurs et des intervalles d'un à sept jours, arrive à des résultats fort différents : un coefficient K de 1,00 (parfait) lorsque l'information est collectée auprès de l'adolescent(e) et 0,80 (quasi parfait) lorsque l'on compare les résultats du parent et ceux de l'adolescent(e). Avec un protocole semblable à celui de l'étude 21, l'étude 39 teste la fidélité du DISC et arrive à des taux forts différents elle aussi : avec un intervalle de moins de 15 jours, la comparaison entre les résultats du parent et ceux de l'adolescent(e) sont de 0,01 pour le groupe contrôle et 0,24 pour le groupe clinique. Les trois dernières études ont analysé la fidélité du K-SADS. Les études 47 et 54 (intervalle inconnu dans un cas et de 14 jours dans l'autre), qui reposent sur deux intervieweurs combinant l'information du parent et de l'adolescent(e) arrivent à des valeurs K de 0,81 et 0,92, ce qui est quasi parfait. Lorsque l'on compare les résultats du parent à ceux de l'adolescent(e), la valeur du coefficient demeure

excellente, mais descend à 0,65. L'étude 55 reprenant ce protocole, arrive à un taux de 0,36, même avec un intervalle très court, soit la journée même.

La cohérence interne des échelles diagnostiques du TC a été analysée pour cinq instruments. Mais contrairement à ce qui a été observé pour les échelles diagnostiques de l'ÉDM, les résultats pour le TC se situent généralement entre 0,56 et 0,62 (α) pour le DISC et le PADDI. Deux exceptions peuvent être faites : dans l'étude 46, les chercheurs obtiennent un coefficient de 0,86 pour le K-SADS et, dans l'étude 57, 0,71 pour le PADDI. L'étude 12, testant le ChIPS, s'intéresse pour sa part au K-R et trouve une valeur de 0,65. Ces résultats sont, de façon générale, sous les normes acceptables.

La validité concordante des instruments a été examinée par sept études, dont une qui en a testé deux [4]. Celle-ci a porté sur le CAPA en le comparant à deux autres instruments, le DAWBA⁴⁴ et le DISC. Les indices K peuvent être qualifiés «légers» à «bons» avec des scores respectifs de 0,39 et 0,41. Quatre études ont analysé le ChIPS. Les études 6 et 7 l'ont comparé au DICA : les intervalles sont de moins de sept jours dans un cas et le jour même dans l'autre, avec l'information de l'adolescent(e). Les résultats présentés sont quasi parfaits ou parfaits, soient 0,80 et 1,00. L'étude 11 a comparé le DICA au K-SADS (intervalle de trois jours et information provenant de l'adolescent(e) avec un résultat de la valeur K de 0,44 (bon). Finalement, dans l'étude 12, le ChIPS a été comparé au CASI⁴⁵ avec un intervalle court (même journée) et en combinant l'information du parent à celle de l'adolescent(e). Le coefficient K obtenu est de 0,57, ce qui est bon. Deux études ont testé le diagnostic de TC posé à l'aide du DISC. La première [4], avec un intervalle de sept jours, avec l'information du parent et de l'adolescent(e), l'a comparé au DAWBA²⁸ et au CAPA. Les valeurs K obtenues sont respectivement de 0,26 et 0,41, donc de légers à bons. La seconde étude [36] a comparé le DISC au CAF⁴⁶. L'intervalle est court (même jour) et les valeurs K sont les suivantes : 0,21 lorsque l'information est recueillie auprès de l'adolescent(e), 0,60 lorsqu'elle est fournie par un parent et 0,46 lorsqu'elle est combinée (de léger à bon). Enfin, une étude [48] a vérifié le

⁴⁴ *Development and Well-Being Assessment* (Goodman, Ford, Richards, Gatward et Meltzer, 2000)

⁴⁵ *Child and Adolescent Symptom Inventory*(Gadow et Sprafkin, 2002)

⁴⁶ *Clinical Assessment Form* (Piacentini et al., 1993)

K-SADS en le comparant au CAS⁴⁷, avec un intervalle d'un jour. Lorsque l'information provient de l'adolescent(e), le coefficient K prend une valeur de 0,50, comparativement à 0,60 lorsqu'elle est du parent et 0,59 lorsqu'elle est combinée, ce qui est bon.

La validité concurrente est souvent examinée lorsqu'il est question du TC. En effet, 12 études ont cherché à savoir si l'instrument obtenait des résultats comparables à une évaluation clinique (*gold standard*). Tout d'abord, trois études [6, 7, 8] ont examiné le ChIPS. Les deux premières, avec l'information recueillie auprès de l'adolescent(e) (intervalle inconnu et sur dossier), obtiennent des résultats de 0,42 et 1,00 (K). La troisième, avec un intervalle de 30 jours et reposant sur l'information fournie par un parent, obtient un score parfait de 1,00. Deux études ont analysé le DICA. La première [16] (dont l'intervalle est inconnu) a comparé le résultat de l'évaluation clinique à l'information provenant de l'adolescent(e) : K de 0,79. Les résultats sont plus faibles lorsque l'information est recueillie auprès du parent (0,57) ou qu'elle est combinée (0,30). La seconde étude (dont l'intervalle est aussi inconnu), utilise également l'information combinée et obtient un résultat K de 0,43. Quatre études ont testé le DISC. Les intervalles qu'elles prévoient varient considérablement : de un à 15 jours [41, 42], à 135,6 jours [22] (moment du congé de l'hôpital) et une étude sur dossier [39]. Si on compare les résultats obtenus lorsque l'information provient de l'adolescent(e) (ce que présentent les quatre études) on retrouve aussi une grande variation, les valeurs du coefficient K étant de 0,06, 0,09, 0,15, et 0,77 (de faibles à excellents). Les études 39 et 41 fournissent également des résultats lorsque l'information est celle du parent (0,15 et 0,74) et lorsqu'elle est combinée (0,34 et 0,80). Les écarts sont à nouveau notables. Finalement, trois études ont pour leur part vérifié la validité concurrente du K-SADS. Encore une fois, les intervalles diffèrent : inconnu pour la première [47], le même jour pour la seconde [53] et deux mois après la fin du traitement pour la troisième [54]. L'information est combinée pour les trois études, ce qui génère des coefficients K de 0,92 pour l'étude 47 et de 0,86 pour l'étude 54. L'étude 53 comporte deux résultats : 0,46 et 0,52 (diagnostic actuel et à vie).

⁴⁷ *Child Assessment Schedule* (Hodges, Stern, Cytryn et McKnew, 1982)

La validité convergente des instruments d'évaluation du TC a été analysée par quatre études. D'abord, l'étude 14 compare le DICA au OCHS-R⁴⁸ (valeur K de 0,45). L'intervalle est inconnu, l'information a été recueillie auprès du parent et du professeur : les résultats sont concluants lorsque comparés au fonctionnement social détérioré, aux pauvres performances scolaires et aux problèmes disciplinaires. L'étude 32 a comparé le DISC à une mesure comportementale, le DIR⁴⁹. L'intervalle est très long, 164 jours. Des observations sont colligées sur une longue période pour les comparer avec le résultat fourni par l'instrument avant le traitement : lorsqu'un diagnostic de TC est généré par le DISC, il est généralement corroboré par les observations colligées à l'aide du DIR qui est plus élevé tout au long du traitement que ce soit au terme du premier trimestre ($F(3,530)=6.92, p < .01, \eta^2=.04$) ou du second ($F(3,413) = 4.25, p < .01, \eta^2 = .03$). Les deux dernières études ont comparé le K-SADS au CBCL⁵⁰, l'intervalle étant de 15 jours pour la première [45] et inconnu pour la deuxième [49]. L'information provient à la fois du parent et de l'adolescent(e) et, dans les deux cas, elle est significativement en lien avec les résultats du CBCL : r de 0,61 [45] et $BEH/POS = 68.7 \pm 8.6$; $BEH/NEG = 53.6 \pm 10.7$; $t_{63} = 6.29, P < .0001$ [49]⁵¹).

Pour terminer, deux études ont analysé la validité discriminante du ChIPS [8] et du DISC [29]. Les intervalles de l'étude portant sur le ChIPS sont de sept jours pour le groupe contrôle et 30 jours pour le groupe clinique. L'information provient du parent et les résultats sont concluants lorsque l'on compare le pourcentage de diagnostics des deux groupes : 6,9% pour le premier versus 16% pour le second. Il en va de même au niveau de la détérioration en lien avec trois indices : famille, école et amis : 2,3% contre 32,3%. La deuxième étude implique une collecte d'observations sur une période de 12 mois avec le DIR et obtient également des résultats concluants. Il est démontré que les jeunes avec un diagnostic de TC au début du traitement avaient 38% plus de risque d'obtenir un taux accru d'incidents au DIR. Les deux

⁴⁸ *Revised Ontario Child Health Study scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

⁴⁹ *Daily Incident Report* (Larzelere, 1996)

⁵⁰ *Child Behavior Checklist* (Achenbach et Edelbrock, 1983; Achenbach, 1991a)

⁵¹ Résultats du CBCL échelle troubles externalisés.

instruments sont donc en mesure de discriminer entre les adolescent(e)s avec et sans diagnostic de TC.

Pour le diagnostic de TC, c'est la mesure de fidélité du test-retest qui est la qualité psychométrique la plus utilisée, tandis que la validité discriminante est la moins répertoriée. Au niveau de la robustesse, on retrouve des taux parfaits (1,00) avec les mesures suivantes : interjuge, concordante et concurrente. Par contre, la fidélité interjuge fait également piètre figure avec un taux négatif de -0,02. Aucune mesure ne se démarque par des taux globalement plus élevés, plusieurs résultats se retrouvant dans les extrêmes.

4.5 Troubles du déficit de l'attention avec ou sans hyperactivité (TDA/H)

Référence Tableau 5 : Trouble déficitaire de l'attention avec ou sans hyperactivité (TDA/H) (annexe F)

Après le TC, le diagnostic de TDA/H est celui qui a été le plus mis à l'épreuve avec 36 études. Le CAPA, ChIPS, DICA, DISC et le K-SADS sont les instruments analysés.

Tableau XX : Nombre d'études par types de fidélité et validité pour les TDA/H

Fidélité			Validité			
Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
10	16	4	8	12	4	3

La fidélité test-retest a encore une fois été la plus souvent examinée (16 études). Une seule étude [9] a porté sur le ChIPS, visant à vérifier la stabilité de l'instrument lors d'une administration en face à face avec le parent et ensuite au téléphone. L'intervalle et le nombre d'intervieweurs sont inconnus et seul un pourcentage d'accord est disponible : 100%. Quatre études [13, 14, 17, 19] ont examiné le DICA. Les intervalles varient cette fois de cinq à 20 jours et deux intervieweurs sont généralement prévus (inconnu dans le cas de l'étude 14). Lorsque l'information provient de l'adolescent(e) [13, 17, 19], les valeurs du coefficient K sont de 0,24, 0,51 et 0,29 (léger à bon). Lorsqu'elle est recueillie auprès du

parent [13, 14], elles sont de 0,86 et 0,59 (quasi parfait et bon). Par ailleurs, huit études ont mis à l'épreuve le DISC. Les intervalles varient de façon générale de sept à 21 jours, exception faite de l'étude 47 (28 jours). Le nombre d'intervieweurs est de deux (adolescent(e)s seulement) ou quatre, sauf dans l'étude 27 où un seul agent de recherche a été impliqué. Lorsque l'information provient de l'adolescent(e) [30, 31 (groupe clinique et groupe contrôle), 38, 41], les valeurs du coefficient K sont respectivement de 0,25, 0,59 (groupe clinique) et 0,43 (groupe contrôle), 0,47 et 0,10. Cinq études [23, 24, 27, 30, 31 (groupe clinique et groupe contrôle), 40, 41] disposent des informations fournies par un parent. L'étude 23 a tenté de vérifier s'il y avait une différence selon qu'on tente d'identifier le diagnostic au cours de la dernière année ou du dernier mois : les résultats sont de 0,50 et 0,58. Les coefficients K présentés dans les autres études sont respectivement de 0,60, 0,91, 0,81, 0,69 (groupe clinique) et 0,57 (groupe contrôle), 0,55 et 0,60. Globalement, il s'agit de meilleurs taux. Finalement, lorsque l'information est combinée (parent-adolescent(e)) [30, 31 (groupe clinique et groupe contrôle), 40, 41], les valeurs K sont de 0,75, 0,68 (groupe clinique) et 0,62 (groupe contrôle), 0,56, 0,48. Trois études ont testé le K-SADS [47, 49, 53] sous l'angle test-retest. Dans le protocole de chacune, deux intervieweurs ont été mis à profit. Les études 47 et 53 reposent sur l'information de l'adolescent(e) et concluent à des valeurs K de 0,81 et 0,87 pour le diagnostic actuel et 0,75 pour le diagnostic à vie. La troisième utilise l'information combinée, avec un résultat de 0,63 pour le diagnostic actuel et 0,55 pour celui à vie. Ces taux sont considérés comme presque parfaits, excellents et bons.

La mesure de la fidélité interjuge a été vérifiée par 10 études pour le TDA/H. Les intervalles entre les deux entrevues sont variés : sur la base d'un enregistrement vidéo, la même journée ou entre sept et 20 jours. Une étude ne le mentionne pas [47]. Seule l'étude 12 a analysé la fidélité interjuge du diagnostic de TDA/H posé à l'aide du ChIPS. Les auteurs ont tenté d'identifier s'il y avait des différences selon le type de TDA. Ainsi, ils ont obtenu des taux d'accord (K) de 0,23 pour tous les types, de 0,43 pour le type hyperactif, de 0,60 pour le type inattentif et de 0,27 pour le type mixte. Les taux sont meilleurs lorsque l'on précise le type, soit hyperactif ou inattentif. Quatre études [13, 18, 20, 21] ont évalué le DICA sous l'angle de la fidélité interjuge. L'étude 13 l'a fait de trois façons. D'abord, l'information

provenant du parent a été comparée à celle fournie par l'adolescent(e), des résultats négatifs (K) de -0,02 étant constatés au temps 1 et -0,01 au temps 2. Ensuite, les résultats entre deux non-cliniciens ont été comparés, les coefficients étant de 0,75 avec les informations du parent et 0,34 avec celle de l'adolescent(e). Finalement, les résultats des non-cliniciens ont été comparés à ceux de pédopsychiatres, les valeurs de K étant de 0,71 avec l'information du parent et 0,78 avec celle de l'adolescent(e). Une étude [21] a confronté les conclusions de deux intervieweurs travaillant à partir des informations de l'adolescent(e), et obtenu un accord parfait (1,00). Trois études [18, 20, 21] comparant l'information fournie par le parent et celle recueillie auprès de l'adolescent(e) ont présenté les résultats suivants : 0,30, 0,42 et 0,66. Une seule étude a examiné le DISC [39] en comparant l'information du parent à celle de l'adolescent(e). Les valeurs K sont de 0,21 pour le groupe clinique et 0,07 pour le groupe contrôle. Les études 44, 47, 54 et 55 ont testé le diagnostic du TDA/H posé à l'aide du K-SADS. Les études 44, 47 et 54 comparent les taux d'accord entre deux intervieweurs combinant l'information du parent et de l'adolescent(e). Les taux sont excellent (0,60), presque parfait (0,88) et parfait (1,00). L'accord entre le parent et l'adolescent(e) des études 54 et 55 donnent des coefficients K de 0,94 et 0,34.

La cohérence interne des échelles diagnostiques du TDA/H a été vérifiée dans quatre études. La vérification a d'abord été faite pour le CHIPS [12] avec de très bons résultats (K-R) : 0,89 pour le type hyperactif, 0,84 pour le type inattentif et 0,89 de façon combinée. Pour le DISC, deux études [24, 40] ont utilisé l'information obtenue auprès du parent et ensuite celle fournie par l'adolescent(e). Avec l'information du parent, les coefficients α sont de plus de 0,90 pour l'étude 24 et 0,87 pour l'étude 40. En utilisant l'information de l'adolescent(e), les résultats sont aussi très bons : entre 0,80 et 0,90 et 0,83. Finalement, l'étude 44 a analysé deux échelles diagnostiques du K-SADS (symptômes au cours des 12 derniers mois et au cours de la dernière semaine). Les résultats obtenus à partir de l'évaluation de deux intervieweurs varient entre 0,75 et 0,84.

La validité concordante a été mise à l'épreuve par sept études, l'une d'entre elles ayant examiné deux instruments. Les intervalles prévus par ces études varient du jour même à sept

jours. Tout d'abord, l'étude 4 a testé le CAPA en utilisant l'information combinée du parent et de l'adolescent(e). En le comparant au DAWBA⁵², le coefficient d'accord est de 0,49 (K); en le comparant ensuite au DISC, il est de 0,52 (bons). Les études 6, 7 et 11 ont évalué le ChIPS en utilisant l'information reçue de l'adolescent(e). Les deux premières études ont mesuré le ChIPS au DICA avec des coefficients d'accord de 0,57 et 0,50. L'étude 11, qui a comparé le ChIPS au K-SADS, conclut à un coefficient de 0,26. Une quatrième étude [12], utilisant l'information combinée pour comparer le ChIPS au CASI⁵³, affiche des coefficients de 0,40 pour tout type de TDA, de 0,47 pour le type hyperactif, de 0,30 pour le type inattentif et de 0,35 pour le type mixte. Deux études ont analysé le DICS. La première [4], en utilisant l'information combinée du parent et de l'adolescent(e), puis le comparant au DAWBA, parvient à un résultat de 0,57. En le confrontant au CAPA, le résultat est de 0,52. La seconde [36] l'a comparé au CAF⁵⁴. À partir des informations provenant du parent, le résultat est de 0,53, comparativement à 0,48 pour celles obtenues de l'adolescent(e) et de 0,46 lorsque les renseignements sont combinés. Pour terminer, l'étude 48 a comparé le K-SADS au CAS⁵⁵. L'information provenant de l'adolescent(e) génère un coefficient d'accord de 0,36. Si on y ajoute l'information scolaire, le résultat est meilleur, avec 0,64. Même observation si on y ajoute l'information du parent : 0,60. Toutefois, c'est l'information combinée qui génère le meilleur taux (0,65).

La validité concurrente est une qualité psychométrique qui suscite également de l'intérêt, 12 études l'ayant examinée pour quatre instruments. D'une étude à l'autre, les intervalles prévus varient : sur dossier, d'un à 30 jours, jusqu'à deux mois après un traitement et même jusqu'à 135 jours suite à l'obtention du congé de l'hôpital. Cela dit, quatre études [6, 16, 21, 47] ne spécifient pas cet élément méthodologique. Trois études ont analysé le ChIPS. Parmi elles, les études 6 et 7, en utilisant l'information de l'adolescent(e) obtiennent des taux d'accord fort différents, soit 0,36 et 1,00 (évaluation sur dossier). La troisième utilise une

⁵² *Development and Well-Being Assessment* (Goodman et al., 2000)

⁵³ *Child and Adolescent Symptom Inventory* (Gadow et Sprafkin, 2002)

⁵⁴ *Clinical Assessment Form* (Piacentini et al., 1993)

⁵⁵ *Child Assessment Schedule* (Hodges et al., 1982)

information combinée et obtient un taux d'accord de 0,74. Les deux études [16, 21] qui testent le DICA, obtiennent des résultats d'accord semblables. La première génère, avec l'information du parent, un taux d'accord avec l'entrevue clinique de 0,46. Avec l'information fournie par l'adolescent(e), ce taux est de 0,40 et lorsque cette information est combinée, de 0,41. Pour la deuxième, à partir d'informations combinées (parent-adolescent(e)), le taux d'accord au moment du congé de l'hôpital, est de 0,50. Quatre études [22, 39, 41, 42] testent le DISC. Lorsque l'information est recueillie auprès de l'adolescent(e), leurs résultats respectifs sont de 0,22 (au congé), 0,36, 0,27 et 0,09 (à l'admission). Ces taux d'accord sont faibles et légers. Pour les études 39 et 41, les taux sont meilleurs soit de 0,42 et 0,72 lorsque l'information provient du parent et de 0,43 et 0,70 lorsqu'elle est combinée. Finalement, trois études testent le K-SADS, en utilisant l'information combinée du parent et de l'adolescent(e). L'étude 47 obtient un taux d'accord de 0,91 avec l'entrevue diagnostique. L'étude 53, obtient des taux d'accord de 0,80 pour le diagnostic actuel et 0,74 pour le diagnostic à vie. La dernière, l'étude 54, obtient un taux d'accord de 0,94, deux mois après la fin du traitement.

Quatre études ont mis à l'épreuve la validité convergente du diagnostic de TDA/H pour deux instruments. L'intervalle est inconnu pour trois de ces études et il est de moins de 15 jours pour l'étude 20. Les études 14, 15 et 20 ont évalué le DICA. La première utilise l'OCHS-R⁵⁶ et ses résultats sont concluants lorsque l'information émane du parent et du professeur en lien avec le fonctionnement social détérioré, les pauvres performances scolaires et les problèmes disciplinaires (k de 0,36). La deuxième le confronte au CBCL⁵⁷ avec l'information combinée du parent et de l'adolescent(e) et ses résultats sont modérément concluants (r de 0,34 (information provenant des parents) et 0,42 (de l'adolescent(e)) à l'échelle des troubles externalisés). La troisième, toujours avec l'information combinée, mais cette fois avec le PIC⁵⁸: les résultats sont peu concluants avec une valeur K de 0,11. L'étude 49 utilise le

⁵⁶ *Revised Ontario Child Health Study scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

⁵⁷ *Child Behavior Checklist* (Achenbach, 1991a)

⁵⁸ *Personality Inventory for Children* (Lacher et Gdowski, 1979; Wirt et al., 1977)

Conners⁵⁹ pour évaluer la validité convergente du K-SADS : avec de l'information provenant du parent, les résultats sont concluants ($ADHD/POS = 25.8 \pm 13.8$; $ADHD/NEG = 13.1 \pm 15.7$; $t_{63} = 3.43, P < .001$).

Pour terminer, trois études ont vérifié la validité discriminante du ChIPS et du DICA. L'intervalle va de 7 à 30 jours pour l'étude 8 et il est inconnu pour les études 14 et 15. L'étude 8 portant sur le ChIPS, obtient des résultats concluants en comparant le nombre de diagnostics (39,5% versus 64,5%) et le niveau de dysfonctionnement (famille, école, amis : 2,3% versus 32,3%) d'un groupe contrôle à un groupe clinique. Les études 14 et 15 pour le DICA, obtiennent également des résultats concluants en lien avec les problèmes de développement, le nombre de diagnostics et de symptômes. En effet, dans l'étude 14, tel qu'attendu, des relations *non significatives* ont été observées entre le diagnostic de TDA/H et la présence de problèmes d'une autre nature (de discipline ou EDM). Dans l'étude 15, trois versions du DICA-R ont pu mesurer une différence significative entre un groupe de 196 jeunes patients consultant en psychiatrie externe (moyennes = 72,7, 84, 9 ou 81,3) et 130 sujets contrôles provenant d'une école ou d'une clinique pédiatrique (moyennes = 58,2, 57,1 ou 65,3; $p < .0005$).

Pour le diagnostic de TDA/H, la qualité psychométrique la plus utilisée est celle du test-retest et la moins utilisée, celle de la validité discriminante. Deux mesures rapportent des taux parfaits (K de 1,00), soit celle de la fidélité interjuge et de la validité concurrente. Par contre, la fidélité interjuge est aussi la mesure avec les scores les moins robustes : K de -0,02 et -0,01. Toutes les mesures utilisées avec le TDA/H obtiennent des taux d'une très grande variabilité (de taux considérés faibles à presque parfaits et parfaits). Seule exception : la fidélité de la cohérence interne, qui globalement, obtient à nouveau les meilleurs taux.

4.6 État de stress post-traumatique (ESPT)

Référence Tableau 6 : État de stress post-traumatique (ÉSPT) (annexe G)

⁵⁹ *Conners Abbreviated Questionnaire, parent version* (Conners et Barkle, 1985)

Le dernier diagnostic est celui de l'ÉSPT, qui a été mis à l'épreuve par 14 études et ce, pour les instruments suivants : CAPA, ChIPS, DICA, K-SADS et PADDI.

Tableau XXI : Nombre d'études par types de fidélité et validité pour les ESPT

Fidélité			Validité			
Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
1	5	2	3	7	0	1

La validité concurrente a été la qualité psychométrique la plus utilisée pour le tester, avec sept études. Par contre, seulement cinq ont pu produire des valeurs K (impossible lorsque 100% d'accord sur l'absence de diagnostic). Les instruments analysés sont le ChIPS [6], le DICA [16] et le K-SADS [47, 53, 54]. Pour le ChIPS, l'intervalle est inconnu. L'information est recensée auprès de l'adolescent(e) et le coefficient K est de 0,62. Pour le DICA, l'intervalle est également inconnu. L'information provenant de l'adolescent(e) génère un coefficient K de 0,55, celle des parents, 0,66 et l'information combinée, 0,55. Les intervalles, entre la passation de l'instrument et l'entrevue diagnostique, sont de un jour [53] et de deux mois après la fin du traitement [54] pour le K-SADS (inconnu pour l'étude 47). Dans ces trois études, l'information utilisée est combinée (parent et adolescent(e)). Les résultats sont de 0,17 et 0,20 (diagnostic actuel et à vie), 1,00 et 0,88.

La seule étude ayant utilisé la fidélité interjuge pour mettre à l'épreuve le diagnostic d'ÉSPT avec le K-SADS est l'étude 54. L'intervalle est de 14 jours et les valeurs obtenues sont parfaites (1,00) : entre deux intervieweurs et entre la mère et l'adolescent(e).

Cinq études ont pour leur part examiné ce diagnostic à l'aide de la fidélité test-retest pour le CAPA [5], le ChIPS [9], le DICA [17, 19] et le K-SADS [49]. Les intervalles varient entre 5 et 17,9 jours. Exception faite de l'étude 9, toutes ont utilisé deux intervieweurs. Les coefficients K sont de 0,22 [17], 0,63 [19] et 0,64 [5] lorsque l'information est recueillie auprès de l'adolescent(e) et de 0,54 pour l'étude 5 qui inclut l'information du parent. L'étude 49, avec une information combinée du parent et de l'adolescent(e) présente des taux de 0,67

(diagnostic actuel) et de 0,60 (diagnostic à vie). Outre l'étude 17, ces résultats sont qualifiés d'excellents. L'étude 9 ne précise pas l'intervalle et utilise le pourcentage pour traduire l'accord entre une passation avec le parent en face à face et ensuite au téléphone : 100%.

Deux études ont évalué la cohérence interne des échelles diagnostiques de l'ÉSPT, toutes deux pour le PADDI [56, 57]. Les résultats sont excellents (0,90 et 0,87).

La validité concordante a été utilisée par trois études [6, 7, 11], qui ont toutes vérifié le CHIPS. Seulement deux ont pu produire des valeurs K [6, 11]. La première en utilisant le DICA et la deuxième, le K-SADS. Les intervalles sont de moins de sept jours et trois jours. Les taux sont considérés comme excellents pour la première (0,67) et bons pour la seconde (0,42).

Aucune étude n'a examiné la validité convergente du diagnostic d'ÉSPT.

Une étude a vérifié la validité discriminante [5] de ce dernier. À l'aide d'échantillons cliniques comparés à un groupe contrôle, les résultats sont concluants. L'étude a comparé le résultat diagnostic du CAPA d'un groupe clinique à un groupe contrôle à l'aide du pourcentage d'évènements traumatiques importants (77,3% versus 43,1%) et faibles à vie (56,8% versus 39,2%), de souvenirs douloureux (51,4% versus 17,2%), d'hyperexcitation (18,9% versus 1,8%) et d'évitement (21,6% versus 7,4%). Dans tous les cas, le pourcentage de ces indicateurs est plus élevé chez les adolescent(e)s avec un diagnostic d'ÉSPT, y compris le taux diagnostic lui-même (9,1% versus 0,9%).

La qualité psychométrique la plus analysée pour le diagnostic d'ÉSPT est la validité concurrente. Outre la mesure de la validité convergente qui n'a pas été utilisée, la moins répertoriée est celle de la fidélité interjuge avec une seule étude ; elle démontre par ailleurs le taux le plus robuste (K de 1,00), ex æquo avec la mesure de la validité concurrente. Par contre, cette dernière rapporte également le score le plus faible (0,17). Globalement, c'est à nouveau la fidélité de la cohérence interne qui démontre les meilleurs résultats.

Pour cinq des six diagnostics retenus, la qualité psychométrique de la fidélité test-retest est la plus examinée. Seule exception : le diagnostic d'ÉSPT où il s'agit plutôt de la validité concurrente. Le même patron se reproduit pour la qualité psychométrique la moins utilisée.

Il s'agit de la mesure de la validité discriminante pour tous les diagnostics, sauf pour celui d'ÉSPT, où une des mesures n'a pas été utilisée du tout : la validité convergente. Il est plus difficile de dégager des généralités en regard des mesures les plus robustes ou les moins robustes. D'abord, parce que chacune des mesures utilisées l'est dans des proportions fort différentes d'un diagnostic à l'autre : aucun n'est testé avec exactement les mêmes mesures, le même nombre de fois, dans les mêmes conditions. Ensuite, parce que dans les taux obtenus à l'aide d'une même mesure, ces derniers passent d'un extrême à l'autre : de léger (et même quelques fois négatifs) à presque parfaits et à parfaits. La seule constance observable est celle de la fidélité de la cohérence interne qui démontre la meilleure stabilité dans les taux, en grande majorité considérés comme excellents pour tous les diagnostics retenus.

La section suivante reprend les taux obtenus, mais par instrument. Neuf des onze instruments seront présentés, puisque pour deux d'entre eux, aucune étude n'a satisfait aux critères du QUADAS, comme le démontrent les Tableaux XII et XV (pp.67 et 78).

5. Analyse par instrument

Les Tableaux X (p.48), XII (p.67) et XV (p.78) montrent très clairement que les instruments n'ont pas tous été évalués avec la même rigueur. Sur les 11 instruments retenus, aucune étude n'a pu être recensée pour deux d'entre eux : le ISCA et le PAI-A. Ce dernier étant somme toute assez récent, (Morey, 2007), la rareté d'études à son sujet pourrait être expliquée par ce fait. Un article de 2014 (Turner, 2014) rapporte effectivement le peu d'études à son sujet, mais indique qu'il s'agit d'un instrument prometteur, qui permettrait l'obtention d'informations utiles pour le traitement et les comportements à risque. Pour l'ISCA, une partie de l'explication peut se trouver dans le fait qu'il s'agit au départ d'un instrument pour les études longitudinales. Il est peu exploité en clinique, même si certaines études l'ont utilisé (Sherrill et Kovacs, 2000) et que l'auteur prévoit en dériver des versions brèves spécifiquement pour les milieux cliniques.

Les résultats seront repris pour chacun des neuf instruments restants. Ils seront présentés de la façon suivante. D'abord le nombre d'études recensées lors de la phase de la SS sera

comparé au nombre d'études retenues pour l'évaluation finale, une fois l'analyse de la robustesse méthodologique complétée. Un rappel des types de la fidélité et de la validité testés sera également effectué, en identifiant les diagnostics mis à l'épreuve. Les résultats seront ensuite résumés. Pour terminer, une appréciation globale de la validité et de la fidélité des instruments sera portée.

5.1 ADI

Une seule étude a été retenue en regard de l'ADI, une étude, qui selon l'analyse du QUADAS-2 est très solide d'un point de vue méthodologique. Il s'agit de l'étude 1 de Winter (1993). Quatre des sept types de la validité et de la fidélité ont été testés et les résultats démontrent que l'instrument en est un très valide pour l'obtention de diagnostics de TUS. Par contre, il s'agit du seul diagnostic ayant été testé pour cet instrument. Étant donné qu'aucune version parent n'existe, seule l'information de l'adolescent(e) était disponible pour effectuer les évaluations.

5.2 APS

Pour l'APS, une étude avait été retenue au moment de la SS. Par contre, suite à l'analyse de la grille du QUADAS-2, elle fut retirée. Les éléments méthodologiques en lien avec l'échantillonnage, l'instrument à l'étude et la séquence n'étaient pas assez robustes.

5.3 CAPA

Pour le CAPA, trois études ont été recensées et toutes trois ont démontré une robustesse méthodologique plus qu'acceptable selon l'analyse de la grille du QUADAS-2. Trois des sept types de la validité et de la fidélité ont été testés : concordante, discriminante et test-retest. À l'exception d'un coefficient de la validité concordante se situant à 0,39 (léger), les taux sont considérés bons, excellents et presque parfaits, y compris un résultat parfait de 1,00 au test-retest du diagnostic de TUS. La validité discriminante de l'ÉSPT a été clairement démontrée en comparant un groupe clinique à un groupe contrôle. Outre ces deux derniers diagnostics, les six diagnostics ont été testés. Autant l'information du parent que de l'adolescent(e) a été utilisée. Les résultats globaux semblent démontrer que le CAPA est un instrument capable

d'être stable et valide avec les diagnostics retenus au fin du présent projet, mais en notant que seules trois mesures de la validité et de la fidélité ont été analysées.

5.4 ChIPS

Sept études furent d'abord recensées pour le ChIPS. Suite à l'analyse de la grille du QUADAS-2, une étude fut retirée. Des éléments méthodologiques en lien avec l'instrument à l'étude et de référence, ainsi que des questions de correspondance ont fait en sorte qu'elle ne fut pas retenue. Six études ont donc fait partie de l'analyse finale pour le ChIPS. Les sept types de validité et fidélité ont été testés : chacun des six diagnostics a été mis à l'épreuve avec l'un ou l'autre. Les coefficients sont variables : de légers à parfaits, selon le type de validité ou fidélité et du diagnostic. Pour le diagnostic de TUS, les taux sont considérés comme excellents, exception faite de la validité discriminante (non concluante). Pour le TAG, outre un coefficient de 0,29 (validité concordante), ils sont tous considérés comme bons, excellents et presque parfaits, ainsi que concluants pour la validité discriminante. Pour l'ÉDM, les coefficients sont légers et bons. Pour le TC, ils sont bons, excellents et parfaits, ainsi que concluants en regard de la validité discriminante. Il n'y a que la cohérence interne où les résultats sont sous les taux acceptables (K-R de 0,65). Pour le TDA/H, les résultats sont dans les extrêmes : de légers à bons, d'excellents à presque parfaits et parfaits et concluants pour la validité discriminante. Certaines variations sont observables lorsque les types de TDH sont décortiqués. Par exemple, la fidélité interjuge varie de 0,23 (tout type) à 0,60 (type inattentif), avec des résultats intermédiaires pour le type hyperactif avec 0,43 et de 0,27 lorsqu'il s'agit d'un type mixte. Finalement, pour l'ÉSPT les résultats de la validité concordante et concurrente sont bons et excellents, ainsi qu'un 100% d'accord au test-retest (face à face versus entrevue au téléphone). Autant l'information du parent que de l'adolescent(e) a été utilisée. Les résultats globaux font du ChIPS un instrument fiable et valide selon les différents diagnostics.

5.5 DICA

Pour le DICA, neuf études furent retenues et toutes furent incluses dans l'échantillon pour l'analyse finale, les qualités méthodologiques étant robustes ou très acceptables. Cinq des

sept mesures psychométriques de la validité et de la fidélité ont été testées : chacun des six diagnostics a été mis à l'épreuve avec l'une ou l'autre. Les taux sont très variables et vont d'accords faibles à parfaits (un seul coefficient parfait en lien avec la validité interjuge pour le diagnostic de TC). Pour le diagnostic de TUS, seule la fidélité test-retest a été analysée, avec des coefficients presque parfaits variant de 0,80 à 0,99. Pour le TAG, les taux d'accords interjuge sont faibles et ceux du test-retest sont légers, bons et excellents. La validité concurrente génère des résultats légers et bons. La validité convergente est non concluante en lien avec le fonctionnement social détérioré et les performances scolaires, mais concluante en lien avec l'humeur dépressive. Pour l'ÉDM, à nouveau l'accord interjuge est faible et les taux du test-retest sont légers et excellents. La validité concurrente génère des taux faibles et légers et la validité convergente génère des taux semblables au TAG. Pour le TC, l'accord interjuge produit des taux aux extrêmes : d'un côté négatifs et de l'autre presque parfaits et parfaits. Même courbe de résultats pour le test-retest : accords faibles à presque parfaits. La validité concurrente démontre des résultats légers, bons et excellents et la validité convergente est concluante en lien avec le fonctionnement social détérioré, les pauvres performances scolaires et les problèmes disciplinaires. Le diagnostic de TDA/H produit également des taux dans les extrêmes : accords interjuge négatifs à excellents et pour les test-retest, de légers à presque parfaits. La validité concurrente est bonne et la validité convergente génère des résultats concluants en lien avec un problème développemental, le nombre de diagnostic et de symptôme, mais également des résultats modérément concluants face à ceux du CBCL⁶⁰ et faible face à ceux du PIC⁶¹. La validité discriminante est non concluante face aux problèmes développementaux, mais l'est lorsque l'on compare le nombre de diagnostics et de symptômes d'un groupe clinique à un groupe contrôle. Finalement la fidélité test-retest de l'ÉSPT génère des taux légers et excellents et sa validité concurrente des résultats bons et excellents. Pour le DICA, il est donc plus difficile d'en tirer des conclusions claires, les résultats se situant souvent dans les extrêmes. Malgré

⁶⁰ *Child Behavior Checklist* (Achenbach, 1991a)

⁶¹ *Personality Inventory for Children* (Lacher et Gdowski, 1979; Wirt et al., 1977)

tout, quelques taux considérés excellents à parfaits, tel que pour la stabilité dans le temps du diagnostic de TUS (test-retest) ou encore la validité interjuge de celui de TC, sont à souligner.

5.6 DISC

Pour le DISC, 22 études ont d'abord été retenues. Suite à l'analyse à l'aide de la grille du QUADAS-2, seulement 15 font partie de l'analyse finale. Sept études furent éliminées en raison de difficultés méthodologiques : informations manquantes, éléments au niveau de l'échantillon, de l'instrument à l'étude, de la séquence et de la correspondance. Treize études sont considérées très robustes méthodologiquement et deux démontrent une méthodologie acceptable. Les sept types de validité et de fidélité ont été testés : cinq des diagnostics ont été mis à l'épreuve avec l'un ou l'autre de ces types. L'ESPT ne l'a pas été. Pour le TUS, seule la fidélité test-retest a été utilisée. Les taux varient de 0,22 (faibles) à 0,80 (excellents) et sont meilleurs lorsque l'information provient de l'adolescent(e). Pour le TAG, l'accord interjuge passe de valeurs négatives (-0,03) à excellentes (0,77) et sont distribués de façon assez égale entre les deux pôles. Pour la fidélité test-retest les résultats varient entre 0,28 et 0,78, soit de légers à excellents. La cohérence interne, entre 0,70 et 0,80, est quelque peu sous les normes (excellent à partir de 0,80). Pour la validité concurrente, les résultats sont considérés comme faibles (0,23) et bons (0,60, 0,51). Le diagnostic d'ÉDM affiche des résultats légers pour l'accord interjuge et généralement bons et excellents pour le test-retest, exception faite de deux résultats sous la barre du 0,31 (légers). La cohérence interne est excellente (entre 0,85 et 0,88). Pour ce qui est de la validité concordante, les résultats sont en dessous de 0,41, soit légers. Pour la validité concurrente, les résultats varient entre 0,17 et 0,63 (faible à excellent) avec une exception à 0,79. Pour le diagnostic de TC, l'accord interjuge est à nouveau faible et très léger (0,01 et 0,24). Les résultats de la mesure du test-retest varient entre 0,42 et 0,87, distribués assez également entre les deux taux (de bons à presque parfaits). Les résultats de la cohérence interne se retrouvent sous les normes acceptables avec une valeur maximale de 0,60. Les résultats de la validité concordante sont considérés comme légers et bons (entre 0,21 et 0,60). Ceux de la validité concurrente sont dans les extrêmes : entre 0,09 et 0,80 de faibles à excellents. Finalement, pour le TDA/H,

l'accord interjuge affiche à nouveau des résultats faibles. Les coefficients de la fidélité test-retest se situent entre 0,10 (faibles) 0,75 (excellents). La cohérence interne, entre 0,80 et 0,90, est excellente. La validité concordante génère des résultats considérés bons et la validité concurrente de faibles (0,09) à excellents (0,72). Pour ces deux dernières mesures, plusieurs des résultats faibles proviennent de l'information de l'adolescent(e). Les résultats obtenus pour le DISC sont très semblables à ceux du DICA : souvent dans les extrêmes. Outre les résultats de la fidélité de la cohérence interne des diagnostics d'ÉDM et TDA/H, qui sont excellents, il y a une grande variabilité dans les valeurs des autres mesures, ce qui fait en sorte qu'il est difficile de porter un jugement clair sur la validité ou fidélité de l'instrument.

5.7 GAIN

Pour le GAIN, une seule étude avait été recensée et elle a obtenu une note parfaite à la grille du QUADA-2 : méthodologiquement très robuste. Seule la validité convergente a été testée pour le diagnostic de TUS (le seul testé) : les résultats sont concluants lorsque comparé au *Form 90*, un questionnaire auto-révélé sur la consommation. Sur la seule foi de cette seule étude, il est difficile de porter un jugement sur la validité et fidélité du GAIN. Notons toutefois que ce n'est pas la rareté d'études sur cet instrument qui en est la cause. Il s'agit plutôt du fait que les critères de la présente recherche pour l'inclusion des études ne rejoignaient pas les protocoles méthodologiques des études portant sur le GAIN.

5.8 K-SADS

Treize études ont d'abord été recensées pour le K-SADS. Suite à l'analyse de la grille du QUADAS-2, 10 sont utilisées pour les analyses finales. Sur ces 10, neuf démontrent une robustesse méthodologique excellente, la dixième, une méthodologie acceptable. Les trois études non retenues démontraient des failles au niveau de l'instrument à l'étude et de référence, ainsi qu'en lien avec la correspondance, ou encore, manquaient d'informations. Les sept types de la validité et de la fidélité ont été testés : chacun des six diagnostics a été mis à l'épreuve avec l'un ou l'autre. Pour le TUS, seuls l'accord interjuge et la validité concurrente ont été testés. Les accords interjuge ont généré des coefficients faibles à parfaits et la validité concurrente, excellents. Pour le TAG, l'accord interjuge, le test-retest, la

cohérence interne et la validité concurrente affichent des taux entre 0,76 et 0,85 (exception faite d'un 0,24 à un test-retest), ce qui représente des niveaux excellents et presque parfaits. La validité convergente est concluante en lien avec le CBCL⁶² et le SCARED⁶³. Pour l'ÉDM, les résultats sont semblables : entre 0,54 et 1,00 pour l'accord interjuge, le test-retest, la cohérence interne et la validité concurrente (exception faite d'un 0,30 pour un accord interjuge). Ces résultats sont bons, excellents, presque parfaits et parfaits. La validité convergente est concluante en lien avec le CBCL, mais la validité discriminante ne l'est pas en lien avec un diagnostic de TAG ou de TC. Pour le TC, encore une fois les résultats sont aussi semblables : l'accord interjuge, le test-retest, la cohérence interne et les validités concordante et concurrente obtiennent des taux entre 0,46 et 0,92 (exception faite d'un 0,36 avec un accord interjuge), donc des résultats bons, excellents et presque parfaits. Le TDA/H obtient aussi des résultats bons, excellents et presque parfaits : l'accord interjuge, le test-retest, la cohérence interne et les validités concordante et concurrente démontrent des coefficients entre 0,55 et 0,91 (exception faite d'un 0,36 en lien avec la validité concordante). La validité convergente est concluante en lien avec le Conners⁶⁴. Finalement, l'ESPT obtient des taux de 0,60 et 0,67 pour la fidélité test-retest et des résultats aux extrêmes pour la validité concurrente : 0,17, 0,20, 0,88 et 1,00. De façon globale, les résultats du K-SADS démontrent une validité et fidélité souvent excellente et presque parfaite, ce qui en fait un instrument plus que recommandable.

5.9 PADDI

Deux études furent recensées pour le PADDI. Suite à l'analyse de la grille du QUADAS-2, les deux études démontrèrent une robustesse méthodologique excellente. Seule la qualité psychométrique de la fidélité de la cohérence interne fut analysée pour quatre diagnostics. Les résultats des diagnostics de TUS, EDM et ESPT sont excellents, soit entre 0,79 et 0,90. Tandis que pour le TC, ils sont quelque peu sous les normes acceptables avec 0,62 et 0,87. Il

⁶² *Child Behavior Checklist* (Achenbach et Edelbrock, 1983; Achenbach, 1991a)

⁶³ *Screen for Child Anxiety Related Emotional Disorders* (Birmaher et al., 1997)

⁶⁴ *Conners Abbreviated Questionnaire* (Conners et Barkle, 1985)

est également difficile de se prononcer pour le PADDI étant donné le peu d'études et le peu de mesures de validité et fidélité recensées.

Les données contenues dans le présent chapitre sur les résultats sont nombreuses et variées. D'abord les particularités des 11 instruments répertoriés ont été présentées. Ensuite, ce sont les études les ayant mis à l'épreuve qui l'ont été. Les types de validité et de fidélité recensés sont également détaillés. Les deux dernières sections présentaient les résultats selon les diagnostics retenus et les instruments d'évaluation. Dans le chapitre suivant sur la conclusion, il sera discuté des résultats de la présente synthèse systématique, de ses limites, des éléments à retenir pour le choix d'un instrument d'évaluation et des avenues possibles pour tenter de pousser plus avant les connaissances.

Conclusion

Par définition, les périodes de l'enfance et de l'adolescence en sont de développement (physique, psychologique, cognitif, neuronale, etc.). Il s'ensuit qu'on ne peut utiliser les mêmes critères, ni les mêmes définitions pour identifier les pathologies qu'à l'âge adulte. Steinberg (2002) affirme d'ailleurs qu'il faut analyser les manifestations morbides dans leur contexte développemental et non seulement comme des phénomènes cliniques touchant des individus plus âgés que les enfants, mais plus jeunes que les adultes. C'est à partir de ce cadre conceptuel que le présent projet a été élaboré.

Dans la recension des écrits, un survol des taux de prévalence et de comorbidité a bien montré l'importance d'une évaluation adéquate qui permette de maximiser l'efficacité des interventions auprès de ces jeunes, de les orienter vers des traitements efficaces, d'éviter une détérioration de leur état et de leur fonctionnement.

Il apparaît donc assez clairement que les adolescent(e)s ont des besoins différents des adultes en fonction de leur développement. La consommation de SPA et différents troubles mentaux peuvent être présents, chez les jeunes comme chez les adultes, mais la façon de les identifier et d'évaluer leur sévérité ne doit pas être conçue de la même façon.

Le Centre canadien de lutte contre les toxicomanies (CCLT) (Comité directeur du projet collaboratif en toxicomanie et en santé mentale, 2014), a d'ailleurs fait état de l'importance d'impliquer tous les fournisseurs de services du réseau, des organismes communautaires aux services publics, afin d'offrir de l'information professionnelle, des consultations approfondies et de l'aiguillage en vue d'offrir des soins supplémentaires en lien avec la SM et la consommation de SPA. Cela est d'autant plus vrai pour les enfants et les adolescent(e)s, chez qui les problèmes de SM s'accompagnent souvent d'un risque d'aggravation de la consommation de SPA (Adair, 2009).

Une des recommandations d'une revue de littérature sur les meilleures pratiques cliniques chez les adolescent(e)s avec comorbidité (Bender et al., 2006) concerne l'évaluation. Elle préconise un processus multiple et continu comprenant des données recueillies auprès de

plusieurs acteurs (adolescent(e)s, cliniciens impliqués, parents, enseignants) afin de permettre que le traitement soit adapté aux besoins changeants de l'adolescent(e).

La validité et la fidélité sont des concepts au cœur de l'examen de tout instrument d'évaluation. C'est à l'aide de ces concepts, qui se déclinent en plusieurs sous-types, que les études portant sur les instruments d'évaluation de la comorbidité qui ont été recensés ont été évaluées.

1. Discussion des résultats

En l'absence de consensus en matière de pratiques professionnelles, la SS s'avère tout indiquée puisqu'elle permet d'identifier ce qui a déjà été publié sur un sujet donné et ensuite de se prononcer sur l'efficacité d'un traitement, ou comme dans le cas présent, les qualités psychométriques d'instruments d'évaluation.

Cette recherche avait pour but d'améliorer les connaissances en lien avec les instruments d'évaluation des troubles psychiatriques comorbides présentés par les adolescent(e)s. Voici les trois objectifs principaux visés au départ :

- Répertorier les instruments d'évaluation multidimensionnelle en mesure de produire des diagnostics de l'axe I (DSM-IV-TR) incluant obligatoirement les TUS, spécifiquement conçus pour les besoins particuliers des adolescent(e)s;
- Identifier les études ayant mis ces instruments à l'épreuve sous l'angle de leur validité et fidélité diagnostiques;
- Évaluer la rigueur méthodologique de ces mêmes études pour répondre aux critères de transparence et d'efficacité des SS

Le premier objectif a donné lieu à l'identification de 11 instruments, spécifiquement conçus ou adaptés pour les besoins des adolescent(e)s présentant des troubles comorbides à l'Axe I, incluant les TUS. Par contre, seulement neuf de ceux-ci se sont vus étudier du point de vue de la validité ou de la fidélité des diagnostics retenus pour le présent projet. De ces neuf outils, quatre l'ont été à l'aide de plus de six études (entre six et 15), études ayant passé le

test d'une robustesse méthodologique au minimum acceptable. Il s'agit des instruments suivants : ChIPS, DICA, DISC et K-SADS. Les cinq restants, soient l'ADI, le CAPA, le GAIN et le PADDI, ne l'ont été qu'avec peu d'études (entre trois et une chacun).

La deuxième phase de la SS, pour sa part, a permis d'identifier 57 études ayant mis ces instruments à l'épreuve, répondant ainsi au deuxième objectif.

L'atteinte du troisième objectif l'a été à l'aide de la grille d'analyse du QUADAS-2. À l'aide de cette grille, 47 études ont été retenues pour former l'échantillon final. Pour une SS, un total de 47 études peut sembler un chiffre assez élevé. Par contre, lorsqu'on le divise par le nombre d'instruments retenus, ces chiffres se répartissent entre 0 et 15 études par instrument.

L'analyse par diagnostic a aussi permis d'observer de très grandes disparités entre les six retenus (TUS, TAG, EDM, TC, TDA/H et ESPT). Certains étant très étudiés, comme l'EDM, le TC et le TDA/H, avec 31, 37 et 36 des 47 études de l'échantillon. D'autres un peu moins, avec 16 études pour les TUS, 26 pour le TAG et 14 pour l'ESPT.

Malgré une disparité un peu moindre, les types de la validité et de la fidélité ont aussi été étudiés de façon inégale. La mesure la plus testée est sans contre dit celle du test-retest avec 20 études. Viennent ensuite les mesures de la fidélité interjuge avec 12 et de la validité concurrente avec 13 études. Les mesures de cohérence interne (fidélité), de validité concordante, convergente et discriminante sont moins utilisées avec sept, six, dix et six études chacune.

Le fait que la fidélité des instruments soit la plus testée peut être expliqué par le fait que le concept de fidélité renvoie à la constance des scores obtenus (Bernier et Pietrulewicz, 1997). Il semble donc normal que l'on veuille d'abord vérifier si les scores sont constants, avant de vérifier leur validité : mesure-t-il réellement le concept ciblé. Certains auteurs résument bien cela : « un instrument non fidèle ne saurait être valide » (Laveault et Grégoire, 2002 : 110) ou encore « aucun instrument ne peut être valide et infidèle, puisque son infidélité signifie qu'il ne peut rien mesurer convenablement » (Berthelot, 2000 : 184).

Finalement, lorsque les données sont analysées de façon globale, les seuls instruments qui ont été testés pour tous les diagnostics retenus, à l'aide d'une des sept mesures psychométriques, sont le CAPA, le ChIPS, le DICA et le K-SADS. Par ailleurs, les sept mesures ont toutes été utilisées seulement avec le ChIPS et le K-SADS, avec l'un ou l'autre des six diagnostics. (Voir le Tableaux XIV et XV, p.77 et p.78.)

Il est difficile de fournir une explication claire de ces constats : lors de la phase d'identification des études ayant mis à l'épreuve les instruments, le K-SADS avait généré 1765 occurrences et 96 études avaient été retenues lors du premier tri sur le titre ou le résumé. Par contre, pour le ChIPS, seulement 118 occurrences avaient été recensées au départ, pour 21 études retenues lors du premier tri, comme le démontre le Tableau X, en page 49. Bien évidemment, si le K-SADS peut être considéré comme un instrument pouvant servir d'étalon (Rathman, 2008) (tout comme la version adulte, le SADS, (Blacker et Endicott, 2008), il est normal que le corpus d'études à son sujet soit plus volumineux. Est-ce à dire qu'un plus grand nombre d'études entraîne une meilleure validité et une meilleure fidélité ? Non, puisque le ChIPS se démarque également avec un corpus qui semble moindre. Plusieurs éléments entrent en ligne de compte, dont les éléments et choix méthodologiques des auteurs de ces études et de celle du présent mémoire. À cet égard, les limites des études rapportées par leurs auteurs (présentées plus bas) peuvent, en partie, expliquer la difficulté de généralisation à partir des résultats présentement obtenus.

Lorsque les taux globaux sont pris en considération, se sont à nouveau le ChIPS et le K-SADS qui se démarquent. Selon les résultats de la présente SS, ces deux instruments ont été les plus évalués (en lien avec les choix méthodologiques du présent mémoire et non en lien avec les chiffres du Tableau X), à l'aide du plus grand nombre de mesures psychométriques et ce pour les six diagnostics retenus et avec les meilleures moyennes des taux inférieurs et supérieurs des indices de la validité et de la fidélité. (Voir le Tableaux XIV et XV, p.77 et p.78 et le Tableau XXII en page 117.)

Selon l'information présentée ci-haut, il est difficile de statuer clairement sur la validité et la fidélité globales des neuf instruments analysés. Les taux sont souvent dans les extrêmes, variant parfois de négatifs à presque parfaits et même parfaits pour une même mesure psychométrique. De plus, étant donné la grande variabilité dans les diagnostics testés et la mesure utilisée, combinaison différente avec chaque instrument, aucune généralisation ne peut être effectuée. Par contre, à petite échelle, certains éléments se dégagent.

Le Tableau XXII permet de présenter les résultats sous un autre angle.

Tableau XXII : Valeurs inférieures et supérieures des indices de validité et de fidélité par instrument

Fidélité				Validité				Moyenne inf. et sup.
Instrument	Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante	
K-SADS (10)	0,07-1,00	(0,24) ¹ 0,54-1,00	0,68-0,89 ²	0,36-0,72	0,17-1,00	0,30-0,61 ³ Concluants x5 ⁴	Non concluant x1 ⁴	0,56-0,87
ChIPS (6)	(0,23) ¹ 0,43-0,77	75-1,00 ⁵	0,65 ⁶	0,26-1,00	0,30-1,00	–	% concluants x4 % non concluants x1 ⁴	0,44-0,88
CAPA (3)	–	0,54-1,00	–	0,39-0,52	–	–	% concluants x1 ⁴	0,47-0,76
DICA (9)	-0,05-1,00	0,17-0,99	–	–	0,13-0,79	0,05-0,45	%concluant x1 ⁴ Non concluants x2 ⁴	0,08-0,81
DISC (15)	-0,05-0,77	0,15-0,91	0,56-0,90 ²	0,21-0,60	0,06-0,80	–	–	0,18-0,80
PADDI (2)	–	–	0,63-0,92 ²	–	–	–	–	0,63-0,92
ADI (1)	0,53-0,86	0,80-0,83	–	–	0,71-0,82	(0,28) ¹ 0,58-0,75 ⁷	–	0,66-0,82
GAIN (1)	–	–	–	–	–	0,75-0,84 ³	–	0,75-0,84
APS (0)	–	–	–	–	–	–	–	–
ISCA (0)	–	–	–	–	–	–	–	–
PAI-A (0)	–	–	–	–	–	–	–	–
Total (47)								

Note 1 : une même étude peut avoir utilisé plus d'une mesure.

Note 2 : sauf indication, il s'agit de valeur K.

(X) : nombre d'études recensées, retenues à la suite de l'analyse du QUADAS-2.

Valeurs négatives ou parfaites

1- Le chiffre entre (...) signifie que l'écart de la première valeur inférieure est d'au moins 20 points avec la valeur suivante.

2- α

3- r

4- nombre d'étude

5- pourcentage

6- K-R

7- *correlational analysis* : indique le degré de corrélation entre variables

Dans le groupe d'instruments moins étudiés, l'ADI se révèle être un bon choix en lien avec le diagnostic de TUS, mais il s'agit du seul testé : les moyennes inférieures et supérieures étant 0,66 et 0,92 (valeur K), ce qui est considéré excellent et presque parfait. Il en est de même pour le GAIN qui démontre une excellente validité discriminante avec des indices se situant entre 0,75 et 0,84 pour le diagnostic de TUS (le seul testé).

Malgré d'excellents taux de cohérence interne obtenus avec le PADDI (seule mesure utilisée), les analyses n'ont été effectuées qu'avec quatre des six diagnostics. Il est donc difficile de se prononcer sur sa validité et fidélité de façon globale, mais les taux sont tout de même excellents, spécialement avec les diagnostics de TUS, EDM et ESPT (α entre 0,79 et 0,90). La même constatation peut être faite avec le CAPA. Il démontre une excellente performance des six diagnostics analysés, mais seulement à l'aide d'une ou deux des trois mesures psychométriques suivantes : validités concordante et discriminante et fidélité test-retest. La moyenne des indices inférieurs et supérieurs est de 0,47 et 0,76, tel que le démontre le Tableau XXII.

Dans un deuxième groupe d'instruments plus étudiés, soit le DICA et le DISC, les taux sont très variables : de valeurs K négatives à parfaites (-0,05 à 1,00). Il est donc ardu de porter un jugement sur leur validité et fidélité générale. Le DICA se distingue par ses résultats démontrant une stabilité dans le temps du diagnostic de TUS (test-retest; valeurs K entre 0,80 et 0,99) et un coefficient parfait en lien avec la validité interjuge pour le diagnostic de TC. Pour sa part, le DISC démontre d'excellents résultats au niveau de la fidélité de la cohérence interne des diagnostics d'EDM et TDA/h (valeurs α entre 0,80 et plus de 0,90).

Deux instruments se sont démarqués par leurs résultats globaux excellents à parfaits au niveau de leur validité et fidélité. Tous deux ont testé les six diagnostics à l'aide de l'une ou l'autre des sept mesures. Il s'agit du ChIPS et du K-SADS. Comme le démontre le Tableau XXII, plusieurs indices présentent des valeurs parfaites. Pour le K-SADS, les valeurs moyennes inférieures et supérieures sont de 0,56 et 0,87 (entre bon et presque parfait). Les résultats sont semblables pour le ChIPS (0,44 et 0,88).

La section qui suit présente un résumé des limites des études recensées, telles que les rapportent leurs auteurs. Ces limites permettent de nuancer certains des résultats.

2. Les limites des études

Lorsque les limites des 47 études sont compilées, telles que les rapportent les auteurs de ces dernières, on note plusieurs similitudes.

Winters (1993) mentionne des limites en lien avec les individus qui composent les échantillons : trop peu nombreux ou pas assez diversifiés (jeunes blancs seulement, consommant alcool et cannabis seulement). Il rapporte également des éléments propres à l'évaluation des TUS, tels que le phénomène de tolérance aux SPA et les changements développementaux rapides qui brouillent les indicateurs de la tolérance chez les jeunes.

Angold (2012) quant à lui, souligne que le protocole d'échantillonnage (cliniques pédiatriques) utilisé dans son étude a généré des taux de prévalence qui n'étaient pas représentatifs de la population en générale.

L'utilisation d'échantillons diversifiés et plus substantiels est une préoccupation rapportée par plusieurs auteurs en regard du petit nombre d'individus ou de la spécificité de l'échantillon de leur étude (Brasil et Bordin, 2010; Bravo et al., 2001; Costello et al., 1998; Dennis et al., 2004; Edelbrock et al., 1999; Flisher et al., 2012; Friman et al., 2000; Fristad, Glickman, et al., 1998; Ghanizadeh et al., 2006; Granero Perez et al., 1998; Hope et al., 1999; Kaufman et al., 1997; Paing et al., 2010; Ribera et al., 1996; Schwab-Stone et al., 1993; Swenson et al., 2007; Witwer et al., 2012). Selon plusieurs, des échantillons plus larges permettraient des analyses plus poussées en lien avec l'impact du genre, de l'âge, de l'ethnie et du lieu (par exemple en traitement interne). Un plus grand nombre d'individus rendrait également possibles des analyses en lien avec les différentes populations cliniques démontrant certaines caractéristiques précises, incluant la variété de la population générale. La diversification des échantillons permet une réelle analyse de la validité et de la fidélité des instruments et aide à les ajuster pour les rendre plus performants avec la population en général ou un groupe particulier. Le petit nombre d'individus composant un échantillon fait

également en sorte que certains diagnostics ne peuvent être analysés dû à un trop faible taux de prévalence (Bravo et al., 2001; Flisher et al., 2012).

Bravo (2012) rappelle que l'utilisation d'individus provenant d'échantillons cliniques tend à produire des résultats plus stables des diagnostics, probablement dû à la gravité de leurs symptômes et leur « expérience » des entrevues diagnostiques. Ils précisent également qu'un instrument destiné à la population générale doit être testé avec cette dernière et non avec des populations cliniques.

Toujours en lien avec la diversité des échantillons, certains auteurs (Bravo et al., 2001; Jensen et al., 1995; Jewell et al., 2004; Kaufman et al., 1997; Lauth et al., 2010; Rothen et al., 2009; Rubio-Stipec et al., 1994; Shahrivar et al., 2010) déplorent une trop grande diversité (ou un échantillon trop ciblé), d'une étude à une autre, puisque cela rend difficile de comparer ou généraliser leurs résultats (*comparability of results*). D'autres auteurs (Edelbrock et al., 1999; Ho et al., 2005) vont aussi dans le même sens, mais en regard des diagnostics : certaines études regroupent des diagnostics, par exemple une seule catégorie pour tous les troubles anxieux, ce qui rend la comparaison difficile avec les études dont les diagnostics sont distincts.

D'autres encore identifient la nécessité d'alterner l'ordre de passation des évaluations (*counterbalanced administration*), c'est-à-dire ne pas toujours débiter par l'adolescent, le parent ou le même intervieweur pour éviter « l'effet de l'ordre » (*order effect*) (Dennis et al., 2004; Paing et al., 2010; Swenson et al., 2007). Ribera (1996) précise que l'utilisation d'un même intervieweur, malgré une alternance (devis de sa propre étude), est aussi une limite et un biais important à prendre en considération. Il précise également que l'utilisation de pédopsychiatres (versus des non-cliniciens) statut connu d'un des groupes de répondants, a pu avoir un impact sur leurs réponses et ainsi les modifier.

La provenance de l'information est aussi une préoccupation (Edelbrock et al., 1999; Hope et al., 1999). Swenson (2007) précise cette idée en expliquant que l'utilisation exclusive de l'information provenant de l'adolescent apporte un biais et que l'apport du parent et

d'autres informateurs vient augmenter la validité de certains diagnostics, surtout en lien avec des comportements moins rapportés par les adolescents, tels que les troubles externalisés (Loeber, Burke, Lahey, Winters et Zera, 2000).

Swenson (2007) et Shahrivar (2010) mentionnent le fait que le niveau de formation des interviewers peut également avoir un impact sur les résultats, surtout en lien avec l'identification d'un diagnostic de SM : une formation de niveau baccalauréat, de maîtrise ou postdoctorale comporte des différences de connaissances importantes qui peuvent influencer sur le verdict final.

Une autre limite rapportée par Boyle (1993) et Lauth (2010), est l'attrition d'une partie de leur échantillon en cours d'étude. Il s'agit là d'une limite importante qui s'est traduite par des biais potentiels pour plusieurs études, comme l'a révélé l'analyse méthodologique à l'aide de la grille du QUADAS-2.

Friman (2000) mentionne qu'une des limites de leur étude est en lien avec l'instrument utilisé pour mettre à l'épreuve le DISC et le diagnostic de TC, soit le DIR⁶⁵ (outil d'observation de certains comportements). Il mentionne que sa validité et sa fidélité étaient en 2000, moment de la publication de l'étude, discutables en tant qu'instrument de mesure psychologique. Par contre, le DIR offrait des avantages non négligeables en tant que mesure de validation externe du diagnostic de TC du DISC, jamais égalé par d'autres études. Jewell (2004) rapporte également cette préoccupation, ajoutant que le DIR étant un outil d'observation, certains comportements n'y étaient manifestement pas inscrits s'ils n'étaient pas observés par le personnel en place.

Toujours en lien les moyens utilisés pour mettre à l'épreuve un diagnostic ou un instrument, Rothen (2009) rapporte l'utilisation d'instruments différents pour effectuer la collecte d'information auprès du parent (*Family History– Research Diagnostic Criteria*⁶⁶) et de

⁶⁵ *Daily Incident Report* (Larzelere, 1996)

⁶⁶ Évaluation semi-structurée (Andreasen N.C., Endicott J., Spitzer R.L. et G., 1977)

l'adolescent (K-SADS), ce qui pourrait avoir eu comme conséquence un taux d'accord plus faible.

Par ailleurs, Jewell (2004) mentionne que le délai, mais surtout le contexte de l'évaluation avant la période d'observation (centre interne de traitement pour adolescents avec TC), pourrait avoir faussé les données si l'adolescent n'a pas dévoilé certains comportements croyant que cela pouvait lui nuire. Jewell mentionne également qu'il faut prendre en considération la période de traitement dans l'analyse des résultats, puisque des changements, résultant du traitement même, ont pu survenir. Shanee (1997) confirme que l'utilisation d'un échantillon clinique interne pourrait avoir faussé les résultats du test-retest, dû aux changements cliniques rapides dans ce contexte de traitement.

L'absence d'un instrument ou d'évaluation diagnostique étalon (*gold standard*) a été soulevée par l'équipe de Kaufman (1997) qui la considère comme une limite de la portée des études.

Ce survol succinct des limites des différentes études retenues pour l'analyse finale, telles que rapportées par leurs auteurs, démontre combien il peut être ardu de comparer des résultats de recherche pour en tirer une conclusion claire.

3. Limites et forces de la synthèse systématique

Plusieurs limites sont identifiables dans le cadre de la présente SS.

Le biais de publication est une limite commune aux SS. Il s'agit d'une possibilité de non-publication des études qui démontrent des résultats peu probants ou négatifs. Une autre limite inhérente aux SS est la « littérature grise » (*gray literature*), c'est-à-dire possiblement non publiée dans les bases de données électroniques (Petticrew et Roberts, 2006). Il est difficile de se prononcer sur l'impact de ces limites sur le présent mémoire, puisqu'à cause de la méthodologie choisie (voir le chapitre sur la méthodologie pour les bases de données utilisées), ces études sont inconnues. Par contre, il est clair que certaines études auraient pu être retenues, mais qu'à cause de la langue (autre que l'anglais ou le français, N=11) ou de l'impossibilité de les obtenir via les services des bibliothèques de l'Université de Montréal

(N=10), elles ont été rejetées. Par ailleurs, advenant qu'elles répondent aux critères d'inclusion, il est peu probable que ces études aient apporté un éclairage bien différent de la conclusion du présent projet, comme le démontrent les limites supplémentaires qui suivront.

Dans le même ordre d'idée, la recherche des études s'étant échelonnée du 8 au 30 novembre 2012, on peut se demander si les données utilisées sont toujours d'actualité au moment du dépôt du présent mémoire, soit décembre 2015. Un bref retour vers les bases de données démontre que peu d'études auraient pu être intégrées au corpus utilisé. En effet, en reprenant le nom de chaque instrument, dans chacune des plateformes utilisées (Ovid, Web of science et Proquest) pour les années 2013 à 2016 et en faisant un rapide survol des titres, 13 études auraient mérité une analyse plus poussée pour en déterminer l'admissibilité. Il est fort probable qu'une fois les critères d'inclusion et d'exclusion appliqués, peu aient été retenues. (Voir l'annexe H pour les détails.)

Une limite non négligeable est celle de la diversité des échelles contenues dans les instruments d'évaluation. Chacun de ces instruments, bien que devant être en mesure de produire des diagnostics spécifiques, est construit différemment. Conséquemment, la multitude de devis de recherche des études les ayant testés également, sans compter les choix méthodologiques fort différents d'un auteur à un autre. Il est donc délicat de comparer les résultats entre eux. Les auteurs des études ventilent⁶⁷ également leurs données différemment. Par exemple, en regroupant plusieurs diagnostics dans une même grande catégorie : tous les troubles de l'humeur ou anxieux dans un seul groupe ou encore les TC et TOP ensemble. De plus, certaines données n'ont pas pu être utilisées parce non ventilées par groupe d'âge. Elles étaient alors écartées pour ne conserver que celles ventilées par diagnostics et catégories d'âge (entre 13 et 18 ans) retenus.

Outre la grande diversité des instruments et des devis de recherche des études qui limite les comparaisons, il y a aussi le petit nombre d'études retenues par instrument. Pour certains, comme le ISCA et la PAI-A, aucune étude n'a été recensée. Il en va de même pour l'APS où

⁶⁷ Répartition des données selon certaines catégories, par exemple, par âge ou par genre.

une seule étude l'avait été, mais n'a pas été retenue une fois l'analyse de la robustesse méthodologique effectuée. Pour six des instruments identifiés, il y a moins de trois études retenues les ayant mis à l'épreuve (incluant le ISCA et le PAI-A).

Une dernière limite de ce projet est le fait qu'il fut accompli par l'étudiante seule. Il est recommandé d'être plus d'un chercheur pour identifier les études qui seront incluses dans l'analyse (Higgins et Green, 2009), ce qui fut, de façon générale, impossible, outre un certain nombre d'avis obtenus du directeur du présent projet.

Une des forces de l'étude est le fait qu'il semble s'agir d'une première compilation d'instruments d'évaluation multidimensionnelle étant en mesure de générer des diagnostics de l'Axe I du DSM-IV-TR, incluant un TUS, instruments spécifiquement destinés aux adolescent(e)s. Malgré les limites, cela permet un premier tour d'horizon lorsque vient le temps pour une organisation de choisir un instrument de ce type.

De plus, la SS identifie clairement les lacunes en ce qui concerne les instruments pouvant être utilisés au Québec, surtout en langue française. Elle démontre en outre les failles méthodologiques des études recensées et peut ainsi servir de guide pour de futures études de même type.

4. Conclusion

Si on prend en considération ce que nous apprend la présente SS, voici les éléments qu'une organisation devrait prendre en considération au moment du choix d'un instrument pour évaluer les troubles concomitants avec une clientèle adolescente, chez qui des TUS et des TM ont préalablement été détectés.

Tout d'abord, avant même de choisir un instrument d'évaluation, il faut identifier le temps, les ressources financières et le type de professionnels dont on dispose (incluant la formation spécifique à l'instrument choisi), la clientèle visée, ses besoins d'évaluation, la manière dont seront utilisées les données recueillies, mais également, la mission de cette organisation. Les organisations qui bénéficieraient davantage des résultats de la présente recherche sont celles travaillant auprès de jeunes aux prises avec des TUS, chez qui on retrouve également

les diagnostics retenus pour le présent projet. Selon la recension des écrits du Chapitre 1, ces derniers sont reconnus les plus prévalents chez les adolescent(e)s. Voici un rappel de ces diagnostics :

- Trouble lié aux substances (TUS)
- Trouble d'anxiété généralisée (TAG)
- Épisode dépressif majeur (ÉDM)
- Troubles des conduites (TC)
- Trouble du déficit de l'attention /hyperactivité (TDA/H)
- État de stress post-traumatique (ÉSPT)

La recension des écrits au Chapitre 1 a clairement démontré que les TUS et tout autre TM sont étroitement liés, qu'on ne peut plus ignorer la concomitance de ces difficultés et qu'il faut répondre aux deux. Leur identification est donc une étape primordiale pour l'élaboration d'un plan de traitement adapté pour ces adolescent(e)s.

Partant de la clientèle visée, soit les adolescent(e)s nécessitant une évaluation approfondie des TM et TUS, l'organisation devra se questionner sur leurs besoins spécifiques. Un instrument conçu pour eux devra être privilégié. Il devra donc avoir été pensé ou adapté en fonction des étapes de développement propres à cette période de vie. Les questions et concepts devront être clairs et compréhensibles pour eux. Il faudra également que les professionnels désignés pour évaluer ces troubles chez les jeunes, soient au fait des trajectoires de consommations de ces derniers, autant d'un point de vue « normal » que « pathologique ». Être en mesure de pousser plus à fond l'exploration lorsque le tableau clinique comporte des symptômes pouvant être répertoriés dans plusieurs catégories sera un atout majeur pour une meilleure précision diagnostique.

Pour éviter de longs délais de traduction, on pourrait considérer une version déjà disponible en français puisqu'il y en a. Il est alors plus réaliste d'entreprendre une adaptation ethnoculturelle rapidement. Cette adaptation semble une condition intégrale du projet d'un

choix d'instrument. En s'assurant d'effectuer les modifications en lien avec la réalité québécoise, on augmentera la validité des résultats, puisqu'une telle adaptation permet une meilleure compréhension des concepts évalués de la part des participants. Cet investissement permettrait d'obtenir un portrait clinique plus exact, d'où découle un plan de traitement plus efficace et plus efficient.

Plusieurs données, autres que celles directement reliées à la présence ou absence des symptômes d'un TM, devront aussi être collectées. Un portrait clair de l'environnement, des antécédents de traitement, du réseau social, du fonctionnement scolaire, doit être obtenu.

Outre l'information obtenue auprès de l'adolescent(e), les renseignements recueillis auprès de tiers significatifs, comme les parents, les enseignants, les autres fournisseurs de soins et services, doivent faire partie intégrante de la démarche menant à un diagnostic et ultimement à un plan de traitement.

Du point de vue de l'exactitude, l'instrument choisi devra avoir démontré que les résultats obtenus reflèteront la réalité de cet(te) adolescent(e) : il devra conclure à un ou des troubles réellement présents, vécus par le jeune. C'est donc le côté de la validité qui entre en jeu. Ces résultats devront être produits de façon constante dans le temps, peu importe qui procède à la passation. La fidélité des résultats est une nécessité indéniable. Ces questions n'en sont pas que l'on se pose d'emblée.

Un début de réponse peut être apporté à ces organisations à partir des détails contenus dans la présente SS sur les 11 instruments retenus, ainsi que sur les études les ayant mis à l'épreuve. Les résultats et grandes conclusions dégagés, pourront également servir de base de discussion pour un choix éventuel.

Malgré les limites de la présente recherche, les informations qu'elle fournit sur plusieurs instruments d'évaluation multidimensionnelle, conçus pour les adolescent(e)s, sont intéressantes. Une des particularités recherchées de ces instruments était leur capacité d'évaluer plusieurs troubles de l'Axe I, mais incluant obligatoirement les TUS. Tous les instruments recensés sont en mesure d'effectuer cette évaluation. Certains d'entre eux ont

été conçus d'abord dans cette optique, comme l'ADI et le GAIN. Malgré le peu de données recueillies sur ces instruments dans le cadre du présent mémoire, ils démontrent d'excellents résultats en lien avec le diagnostic de TUS.

Selon les résultats de la SS, les études ayant mis à l'épreuve les différents instruments en démontrent les forces et les faiblesses. Ainsi, si l'on recherche un instrument démontrant des résultats stables dans le temps, le DICA répond à ces attentes en lien avec le diagnostic de TUS. De plus, la majorité des instruments démontrent une excellente cohérence interne des éléments utilisés pour mener à un diagnostic.

Une mention spéciale doit être faite pour le CAPA. Malgré que seulement trois mesures psychométriques aient été utilisées (validités concordante et discriminante et fidélité test-retest), les six diagnostics ont été évalués. Les résultats démontrent une excellente performance.

Comme mentionné précédemment, deux instruments se démarquent, tant par le nombre des diagnostics testés et les types de validité et fidélité utilisés pour les évaluer, que par les résultats. Il s'agit du ChIPS et du K-SADS. Certaines caractéristiques qui les différencient sont à prendre en considération lorsque vient le temps de faire un choix ou d'amorcer une réflexion en ce sens.

Rappelons les caractéristiques de ces instruments.

Tout d'abord le ChIPS est un instrument basé sur le participant, donc structuré. Ainsi, malgré le besoin d'y être formé, des non-cliniciens peuvent s'acquitter de la passation. Il est en mesure d'identifier 20 diagnostics et est disponible en version adolescent(e) et parent. Le temps de passation rapporté est assez court (autour de 40 minutes). Il n'est disponible qu'en anglais et nécessite donc une traduction et une adaptation ethnoculturelle, avant d'être validé au Québec auprès de la clientèle ciblée. Selon les résultats obtenus au niveau de la fidélité, il se révèle un instrument très stable dans le temps (test-retest) et assez solide en ce qui concerne la passation par différents intervenants (accord interjuge). Du point de vue de sa validité, il semble être en mesure de distinguer les jeunes avec et sans la condition visée

(discriminante). Lorsque comparés aux résultats d'une évaluation par un professionnel ou des instruments semblables, les résultats varient, mais obtiennent tout de même des résultats parfaits (validité concordante et concurrente).

Le K-SADS offre une alternative au format structuré du CHIPS : de format semi-structuré, il est basé sur l'intervieweur, ce qui nécessite la passation par des cliniciens formés, mais qui leur laisse plus de latitude dans l'exploration de certains items. Générant plus de 30 diagnostics, il est disponible en français. L'organisation le choisissant pourra passer directement à la phase d'adaptation ethnoculturelle auprès de la clientèle adolescente. Son temps de passation auprès d'une clientèle clinique nécessite plus de temps, soit de 60 à 85 minutes. Il ne comporte par contre pas de module pour collecter les données auprès de tiers. Il se révèle un instrument assez stable dans le temps (test-retest), tout comme la passation par divers intervenants (interjuge) mais avec des valeurs inférieures plus faibles que le CHIPS. Sa valeur discriminante se révèle aussi plus faible que ce dernier, mais sa validité concordante est plus solide. Lorsque les résultats sont comparés à ceux des évaluations par des professionnels ou à ceux d'autres types d'outils, il se démarque avec des résultats plus élevés que le CHIPS.

Comme il a été mentionné plus haut, le choix d'un instrument d'évaluation doit prendre en considération plusieurs éléments, dont le personnel dont on dispose, le type de services et de soins offerts, le temps disponible, la clientèle cible, etc. Par ailleurs, pour le Québec, les choix impliquent également soit une traduction ou une validation auprès d'une clientèle en majorité francophone. Ces deux instruments offrent donc des qualités différentes, pouvant répondre à des impératifs différents.

Outre, une adaptation ethnoculturelle, il serait intéressant de pousser plus avant les recherches en lien avec des groupes d'adolescent(e)s spécifiques au Québec, tels que les jeunes recevant des services de réadaptation en dépendance en interne ou en externe; ou encore les jeunes des Centres jeunesse ou suivis en pédopsychiatrie.

Les instruments recensés pour le présent projet répondent aux principes de la PPD, puisqu'ils ont été conçus ou adaptés pour les jeunes. Il faudrait se questionner sur les instruments de toutes sortes utilisés présentement dans le réseau de la santé et des services sociaux du Québec pour évaluer les besoins des jeunes en termes de troubles mentaux. Le survol de la recension des écrits en début de mémoire, démontre clairement que la présence de troubles mentaux, incluant un TUS, souvent de façon concomitante, touche particulièrement les jeunes. Avons-nous toujours des instruments adaptés à leurs besoins évolutifs et à leur niveau de compréhension en mesure de bien identifier ces portraits cliniques ? Effectuons-nous une cueillette de données auprès des tiers significatifs, tel que le préconisent Bender et ses collaborateurs (2006) ? Sommes-nous toujours en mesure de bien expliquer aux adolescent(e)s comment les aspects évalués se présentent chez un jeune de leur âge ? Une bonne utilisation des principes de base de la PPD le préconise, mais rien n'est moins certain.

Il a aussi été démontré que rien ne sert d'utiliser des *evidence based treatment* (EBT), si au préalable les instruments d'évaluation et tout le processus qui en découle, ne répondent pas aux mêmes normes (*EBA-evidence-based assessment*). Pour ce faire, des études pour bien comprendre comment cela se fait au Québec aideraient à mieux adapter les actions d'évaluation et de traitement : processus d'évaluation, d'analyse de la collecte de données, de mise en place d'un plan d'intervention et d'application du traitement.

Les adolescents et adolescentes de toutes cultures et de tous horizons méritent des services et des soins à la hauteur de leurs besoins propres.

Bibliographie

Note : les références identifiées à l'aide d'un * représentent les études composant l'échantillon final d'analyse (47 études).

- *Abrantes, A. M., Hoffmann, N. G., Anton, R. P. et Estroff, T. W. (2004). Identifying co-occurring disorders in juvenile justice populations. *Youth Violence and Juvenile Justice*, 2(4), 329-341. doi: <http://dx.doi.org/10.1177/1541204004267781>
- Académie canadienne de psychiatrie de l'enfant et de l'adolescent. (2011). Politiques et directives [http://www.cacap-acpea.org/fr/cacap/Politiques_et_directives_p810.html]. Repéré le 9 mars 2011
- Achenbach, T. (1993). *Empirically based taxonomy : How to use syndromes and profile types derived from the CHCL/4-18, TRF, ans YSR*. Burlington, VT: University of Vermont, Department of psychiatry.
- Achenbach, T. et Edelbrock, C. S. (1984). Psychopathology of children. *Annual Review of Psychology*, 35, 227-256.
- Achenbach, T. et Edelbrock, G. (1983). *Manual for the Child and Behavior Checklist and revised Child Behavior Profil*. Burlington: University of Vermont: Department of Psychiatry.
- Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.
- Achenbach, T. M. (1991b). *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Adair, C. E. (2009). *Concurrent substance use and mental disorders in adolescents: A review of the literature on current science and practice*. Alberta Centre for Child, Family and Community Research.
- Ambrosini, P. J., Metz, C., Prabucki, K. et Lee, J.-c. (1989). Videotape reliability of the third revised edition of the K-SADS. *Journal of the American Academy of Child & Adolescent Psychiatry*, 28(5), 723-728. doi: <http://dx.doi.org/10.1097/00004583-198909000-00013>
- American Academy of Child and Adolescent Psychiatry. (2011). Guidelines. Repéré le 11 mars 2011 : <http://www.jaacap.com/search/quick>
- American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders : DSM-IV*. (4th ed..^e éd.). Washington, DC: Washington, DC : American Psychiatric Association.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders, text revision (DSM-IV-TR)*. American Psychiatric Association.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. (Fifth edition..^e éd.). Washington, DC : American Psychiatric Publishing.

- American Psychological Association Presidential Task Force on Evidence-based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, 61(4), 271-285. doi: 10.1037/0003-066X.61.4.271
- American Society of Addiction Medicine (ASAM). (2001). Public Policy Statement: Definition of Addiction ; <http://www.asam.org/for-the-public/definition-of-addiction>. Repéré le 31 mars 2015
- Anastasi, A. (1986). Evolving concepts of test validation .7. *Annual Review of Psychology*, 37, 1-15.
- Andreasen N.C., Endicott J., Spitzer R.L. et G., W. (1977). The family history method using diagnostic criteria. Reliability and validity. *Archives of General Psychiatry*, 34, 1229-1253.
- Angold, A. et Costello, E. (1995). A test-retest reliability study of child-reported psychiatric symptoms and diagnoses using the Child and Adolescent Psychiatric Assessment (CAPA-C). *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences*, 25(4), 755-762. doi: <http://dx.doi.org/10.1017/S0033291700034991>
- Angold, A., Erkanli, A., Copeland, W., Goodman, R., Fisher, P. W. et Costello, E. J. (2012). Psychiatric Diagnostic Interviews for Children and Adolescents: A Comparative Study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(5), 506-517. doi: 10.1016/j.jaac.2012.02.020
- Angold, A. et Fisher, P. W. (1999). Interviewer-based interviews (*Diagnostic assessment in child and adolescent psychopathology* (p. 34-64). New York, NY: Guilford Press; US.
- Arial, P. (2009). *La réadaptation de jeunes toxicomanes qui présentent un problème de santé mentale*. Université du Québec à Trois-Rivières.
- Aronen, E. T., Noam, G. G. et Weinstein, S. R. (1993). Structured diagnostic interviews and clinicians' discharge diagnoses in hospitalized adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 32(3), 674-681.
- Association canadienne de psychiatrie. (2011). Guides de pratique clinique. Repéré le 9 mars 2011
- Association des centres de réadaptation du Québec (ACRDQ) et Chestnut Health Systems (GAIN Coordinating Center). (2012). *Document interne de formation*. Document inédit.
- Aulakh, H. (2010). *Biological, psychological and gambling variables associated with problem gambling: A functional magnetic resonance imaging study*. (Université McGill, Montréal (Québec)).
- Bartko, J. J. et Carpenter, W. T. (1976). On the methods and theory of reliability. *Journal of Nervous and Mental Disease*, 163, 307-317.
- Beck, J. S., Beck, A. T. et Jolly, J. (2001). *Manual for the Beck Youth Inventories of Emotional and Social Impairment*. San Antonio, TX: Psychological Corporation.
- Bender, K., Springer, D. et Kim, J. (2006). Treatment effectiveness with dually diagnosed adolescents : A systematic review. *Brief Treatment and Crises Intervention*, 6(3), 177-205.

- Bernier, J.-J. et Pietrulenwicz, B. (1997). *La psychométrie : traité de mesure appliquée*. Montréal: G. Morin.
- Berthelot, J., M., (2000). *Sociologie: Épistémologie d'une discipline*. Bruxelles: Éditions De Boeck Université.
- Bertrand, K., Brunelle, N., Ménard, J.-M., Richer, I., Lemieux, L. et Ledoux, C. (2009). Étude évaluative auprès d'adolescents en traitement de la toxicomanie (partie 1) : suivi trois mois et six mois. Dans Rapport final remis au Ministère de la santé et des services sociaux (dir.). Québec.
- Bertrand, K., Ménard, J.-M. et Plourde, C. (2004). Synthèse de l'évaluation du programme jeunesse, recension des écrits et recommandation pour sa révision: Domrémy Mauricie/Centre du Québec, Centre de réadaptation en toxicomanie/jeu excessif.
- Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J. et Neer, S. M. (1997). The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(4), 545-553.
- Blacker, D. et Endicott, J. (2008). Psychometric Properties : Concepts of Reliability and Validity. Dans A. J. R. Jr, M. B. First & D. Blacker (dir.), *Handbook of Psychiatric Measures, Second Edition* (p. 7-13). London, England: American Psychiatric Publishing, Inc.
- Bolton, J. M., Robinson, J. et Sareen, J. (2009). Self-medication of mood disorders with alcohol and drugs in the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Affective Disorders*, 115(3), 367-375. doi: 10.1016/j.jad.2008.10.003
- Boyle, M. H., Offord, D. R., Racine, Y., Fleming, J. E., Szatmari, P. et Sanford, M. (1993). Evaluation of the revised Ontario Child Health Study scales. *Journal of Child Psychology and Psychiatry*, 34(2), 189-213.
- Boyle, M. H., Offord, D. R., Racine, Y., Sanford, M. et et al. (1993). Evaluation of the Diagnostic Interview for Children and Adolescents for use in general population samples. *Journal of Abnormal Child Psychology: An official publication of the International Society for Research in Child and Adolescent Psychopathology*, 21(6), 663-681. doi: <http://dx.doi.org/10.1007/BF00916449>
- Boyle, M. H., Offord, D. R., Racine, Y. A., Szatmari, P., Sanford, M. et Fleming, J. E. (1997). Adequacy of interviews vs checklists for classifying childhood psychiatric disorder based on parent reports. *Archives of General Psychiatry*, 54(9), 793-799.
- Brasil, H. H. et Bordin, I. A. (2010). Convergent validity of K-SADS-PL by comparison with CBCL in a Portuguese speaking outpatient population. *BMC Psychiatry*, 10, 83. doi: <http://dx.doi.org/10.1186/1471-244X-10-83>
- Bravo, M., Ribera, J., Rubio-Stipec, M., Canino, G., Shrout, P., Ramirez, R., . . . Taboas, A. M. (2001). Test-retest reliability of the Spanish version of the Diagnostic Interview Schedule for Children (DISC--IV). *Journal of Abnormal Child Psychology*, 29(5), 433-444. doi: <http://dx.doi.org/10.1023/A:1010499520090>

- Breton, J. J., Bergeron, L., Valla, J. P., Berthiaume, C. et St-Georges, M. (1998). Diagnostic Interview Schedule for Children (DISC-2.25) in Québec: Reliability findings in light of the MECA study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 37(11), 1167-1174.
- Brochu, S. (2007). Nos réponses face à l'abus de drogues chez les jeunes (*Toxicomanie au Canada : Plein feux sur les jeunes* (p. 22-29). Ottawa (Ontario): Centre canadien de lutte contre l'alcoolisme et les toxicomanies.
- Brown, T. E. (1996). Brown attention deficit disorder scales. San Antonio, TX: The Psychological Corporation.
- Brown, T. G. et Wood, W.-J. (2002). Abus de substances: tous les traitements se valent-ils? Québec, Gouvernement du Québec: Comité permanent de lutte à la toxicomanie (CPLT).
- Carroll, B. J., Feinberg, M., Smouse, P. E., Rawson, S. G. et Greden, J. F. (1981). The Carroll Rating Scale for Depression, I: Development, reliability and validation. *Br J Psychiatry*, 138, 194-200.
- Center for Substance Abuse Treatment. (1999). Screening and assessing adolescents for substance use disorders. *Treatment Improvement Protocol (TIP no 31)*. Rockville (MD): Substance Abuse and Mental Health Services Administration (SAMHSA).
- Center for substance abuse treatment. (2005). Substance abuse treatment for persons with co-occurring disorders *Treatment Improvement Protocol (TIP no 42)*. Rockville, MD: Substance Abuse and Mental Health Services Administration (SAMHSA).
- Chambers, W. J. et et al. (1985). The assessment of affective disorders in children and adolescents by semistructured interview: Test-retest reliability of the Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present Episode Version. *Archives of General Psychiatry*, 42(7), 696-702. doi: <http://dx.doi.org/10.1001/archpsyc.1985.01790300064008>
- Chan, Y. F., Dennis, M. L. et Funk, R. R. (2008). Prevalence and comorbidity of major internalizing and externalizing problems among adolescents and adults presenting to substance abuse treatment. *Journal of Substance Abuse Treatment*, 34(1), 14-24. doi: 10.1016/j.jsat.2006.12.031
- Charles, G. et Alexander, C. (2007). Lacune dans nos approches en matière de consommation et d'abus de substances chez les jeunes (*Toxicomanie au Canada : Plein feux sur les jeunes* (p. 38-45). Ottawa (Ontario): Centre canadien de lutte contre l'alcoolisme et les toxicomanies.
- Chassin, L., Beltran, I., Lee, M., Haller, M. et Villalta, I. (2010). Vulnerability to substance use disorders in childhood and adolescence (*Vulnerability to psychopathology: Risk across the lifespan* (2nd^e éd., p. 113-140). New York, NY: Guilford Press; US.
- Chauvet, M., Kamgang, E., Ngamini Ngui, A., Fleury, M. J. et Simoneau, H. (2015). Les troubles liés à l'utilisation de substances psychoactives : Prévalence, utilisation des services et bonnes pratiques. Montréal, Qc: Centre de réadaptation en dépendance de Montréal – Institut universitaire (CRDM-IU).

- Chung, T. et Martin, C. S. (2005). What were they thinking? Adolescents' interpretations of DSM-IV alcohol dependence symptom queries and implications for diagnostic validity. *Drug and Alcohol Dependence*, 80(2), 191-200. doi: 10.1016/j.drugalcdep.2005.03.023
- Cicchetti, D. (1984). The emergence of developmental psychopathology. *Child Dev*, 55(1), 1-7.
- Cicchetti, D. (1987). Developmental psychopathology in infancy: Illustration from the study of maltreated youngsters. *J Consult Clin Psychol*, 55(6), 837-845.
- Cicchetti, D. et Garmezy, N. (1993). Prospects and promises in the study of resilience. *Development and psychopathology*, 5(04), 497-502.
- Cicchetti, D. et Toth, S. L. (2009). The past achievements and future promises of developmental psychopathology: The coming of age of a discipline. *J Child Psychol Psychiatry*, 50(1-2), 16-25. doi: 10.1111/j.1469-7610.2008.01979.x
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284-290.
- Cohen, J. (1968). Weighted kappa : Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*, 70, 213-220.
- Comité directeur du projet collaboratif en toxicomanie et en santé mentale. (2014). Collaboration pour les soins en toxicomanie et en santé mentale : Meilleurs conseils. Ottawa, On.: Centre canadien de lutte contre les toxicomanies.
- Comité permanent de lutte à la toxicomanie (CPLT). (2001). Avis sur les services de réadaptation pour les mineurs qui font un usage inapproprié de substances psychoactives: Gouvernement du Québec. Ministère de la santé et des services sociaux.
- Commission de la santé mentale du Canada. (2009). Vers le rétablissement et le bien-être : Cadre pour une stratégie en matière de santé mentale au Canada, 129 pp.
- Commission de la santé mentale du Canada. (2013). La nécessité d'investir dans la santé mentale au Canada. Calgary, Alberta.
- Conners, C. et Barkle, y. R. (1985). Rating scales and checklists for child psychopharmacology. *Psychopharmacol Bull*, 21, 809-843.
- Conners, C. K. (1989a). Conners' Parent Rating Scale. North Tonawanda, NY: Multi-Health Systems.
- Conners, C. K. (1989b). Conners' Teacher Rating Scale. North Tonawanda, NY: Multi-Health Systems.
- Conners, C. K., Wells, K. C., Parker, J. D., Sitarenios, G., Diamond, J. M. et Powell, J. W. (1997). A new self-report scale for assessment of adolescent psychopathology: Factor structure, reliability, validity, and diagnostic sensitivity. *J Abnorm Child Psychol*, 25, 487-497.
- Costello A.J., Edelbrock C.S., Dulcan M.D., Kalas R. et Klaric S.H. (1984). *Report of the NIMH Diagnostic Interview Schedule for Children (DISC)*. Washington, DC: National Institute of Mental Health
- Costello, E., Angold, A., March, J. et Fairbank, J. (1998). Life events and post-traumatic stress: The development of a new measure for children and adolescents. *Psychological*

- Medicine: A Journal of Research in Psychiatry and the Allied Sciences*, 28(6), 1275-1288. doi: <http://dx.doi.org/10.1017/S0033291798007569>
- Costello, E. J., Erkanli, A., Federman, E. et Angold, A. (1999). Development of psychiatric comorbidity with substance abuse in adolescents: Effects of timing and sex. *Journal of Clinical Child Psychology*, 28(3), 298-311.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Crowley, T. J., Mikulich, S. K., Ehlers, K. M., Whitmore, E. A. et Macdonald, M. J. (2001). Validity of structured clinical evaluations in adolescents with conduct and substance problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(3), 265-273. doi: <http://dx.doi.org/10.1097/00004583-200103000-00005>
- Currie, J. (2001). *Meilleures pratiques - Traitement et réadaptation des jeunes ayant des problèmes attribuables à la consommation d'alcool et d'autres drogues*. Ottawa: Division de la stratégie antidrogue - Santé Canada.
- de la Osa, N., Ezpeleta, L., Domenech, J., Navarro, J. et Losilla, J. (1997). Convergent and discriminant validity of the Structured Diagnostic Interview for children and adolescents (DICA-R). *Psychology in Spain*, 1(1), 37-44.
- Dennis, M. L., Funk, R., Godley, S. H., Godley, M. D. et Waldron, H. (2004). Cross-validation of the alcohol and cannabis use measures in the Global Appraisal of Individual Needs (GAIN) and Timeline Followback (TLFB; Form 90) among adolescents in substance abuse treatment. *Addiction*, 99(Suppl2), 120-128. doi: <http://dx.doi.org/10.1111/j.1360-0443.2004.00859.x>
- Dennis, M. L., Rourke, K. M., Lennox, R., Campbell, R. S. et Caddell, J. M. (1995). *Global Appraisal of Individual Needs: Background and Psychometric Properties Manual (GAIN-B)*. Research Triangle Park, NC: Research Triangle Institute.
- Dennis, M. L., Titus, J. C., White, M. K., Unsicker, J. I. et Hodgkins, D. (2003). *Global appraisal of individual needs: Administration guide for the GAIN and related measures (version 5)*. Bloomington, IL: Chestnut Health Systems.
- Desrodières, P. et Jacques, C. (2009). *Les services en jeu pathologique dans les centres de réadaptation en dépendance. Guide de bonnes pratiques et offre de services de base*. Association des centres de réadaptation en dépendance du Québec.
- Diamond, G., Panichelli-Mindel, S. M., Shera, D., Dennis, M., Tims, F. et Ungemack, J. (2006). Psychiatric syndromes in adolescents with marijuana abuse and dependency in outpatient treatment. *Journal of Child & Adolescent Substance Abuse*, 15(4), 37-54. doi: 10.1300/J029v15n04_02
- Dillon, F. R., Turner, C. W., Robbins, M. S. et Szapocznik, J. (2005). Concordance among biological, interview, and self-report measures of drug use among African American and Hispanic adolescents referred for drug abuse treatment. *Psychology of Addictive Behaviors*, 19(4), 404-413.
- Drabick, D. A. G. et Kendall, P. C. (2010). Developmental psychopathology and the diagnosis of mental health problems among youth. *Clinical Psychology: Science and Practice*, 17(4), 272-280. doi: 10.1111/j.1468-2850.2010.01219.x

- Dufour, M., Nadeau, L. et Gagnon, S. R. (2014). Tableau clinique des personnes cyberdépendantes demandant des services dans les centres publics de réadaptation en dépendance au Québec : étude exploratoire. *Santé mentale au Québec*, 39(2), 149-168.
- Edelbrock, C., Cnric, K. et Bohnert, A. (1999). Interviewing as communication: An alternative way of administering the Diagnostic Interview Schedule for Children. *Journal of Abnormal Child Psychology*, 27(6), 447-453. doi: <http://dx.doi.org/10.1023/A:1021979925865>
- Endicott, J. et Spitzer, R. L. (1978). A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Arch. Gen. Psychiatry* 35, 837-844.
- Endicott J. et Spitzer R.L. (1978). A diagnostic interview: The schedule for affective disorders and shizophrenia. *Arch Gen Psychiatry*, 35, 837-844.
- Estroff, T. et Hoffmann, N. (2001). *PADDI: Practical Adolescent Dual Diagnosis Interview*.
- Ezpeleta, L., de la Osa, N., Domenech, J. M., Navarro, J. B., Losilla, J. M. et Judez, J. (1997). Diagnostic agreement between clinicians and the Diagnostic Interview for Children and Adolescents--DICA-R--in an outpatient sample. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 38(4), 431-440.
- Farrington, D. P. et Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *Annals of the American Academy of Political and Social Science*, 578(ArticleType: research-article / Issue Title: What Works in Preventing Crime? Systematic Reviews of Experimental and Quasi-Experimental Research / Full publication date: Nov., 2001 / Copyright © 2001 American Academy of Political and Social Science), 35-49.
- Fazel, S., Doll, H. et Långström, N. (2008). Mental Disorders Among Adolescents in Juvenile Detention and Correctional Facilities: A Systematic Review and Metaregression Analysis of 25 Surveys. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(9), 1010-1019. doi: <http://dx.doi.org/10.1097/CHI.ObO13e31817eecf3>
- Federman, E. B., Costello, E. J., Angold, A., Farmer, E. M. Z. et Erkanli, A. (1997). Development of substance use and psychiatric comorbidity in an epidemiologic study of white and American Indian young adolescents The Great Smoky Mountains Study. *Drug and Alcohol Dependence*, 44(2-3), 69-78.
- First, M. B., Spitzer, R. L., Gibbon M. et Williams, J. B. W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P)*. New York: Biometrics Research, New York State Psychiatric Institute.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Flisher, A., Sorsdahl, K. et Lund, C. (2012). Test-retest reliability of the Xhosa version of the Diagnostic Interview Schedule for Children. *Child: Care, Health and Development*, 38(2), 261-265. doi: <http://dx.doi.org/10.1111/j.1365-2214.2010.01195.x>
- Friedman A. S. et Utada A. (1989). A method for diagnosing and planning the treatment of adolescent drug abusers (the Adolescent Drug Abuse Diagnosis [ADAD] instrument). *Journal of Drug Education*, 19, 285-312.

- Friman, P. C., Handwerk, M. L., Smith, G. L., Larzelere, R. E., Lucas, C. P. et Shaffer, D. M. (2000). External validity of conduct and oppositional defiant disorders determined by the NIMH Diagnostic Interview Schedule for children. *Journal of Abnormal Child Psychology*, 28(3), 277-286.
- Fristad, M. A., Cummins, J., Verducci, J. S., Teare, M., Weller, E. et Weller, R. A. (1998). Study IV: Concurrent validity of the DSM-IV revised Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology*, 8(4), 227-236. doi: <http://dx.doi.org/10.1089/cap.1998.8.227>
- Fristad, M. A., Glickman, A. R., Verducci, J. S., Teare, M., Weller, E. B. et Weller, R. A. (1998). Study V: Children's Interview for Psychiatric Syndromes (ChIPS): Psychometrics in two community samples. *Journal of Child and Adolescent Psychopharmacology*, 8(4), 237-245. doi: <http://dx.doi.org/10.1089/cap.1998.8.237>
- Gadow, K. D. et Sprafkin, J. (2002). *Child symptom inventory 4 : Screening and norms manual*. Stony Brook, NY: Checkmate Plus.
- Gagnon, S. R., Hickey, N., Hugron, M. et Racine, P. (2006). *Étude d'enquête sur les habitudes de jeu de hasard et d'argent des adolescents lanaudois*. Joliette: Centre Le Tremplin, CSSSNL, publication interne.
- Geller, B., Williams, M., Zimmerman, B. et Frazier, J. (1996). WASH-U-KSADS (Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia). *St. Louis (Missouri): Washington University*.
- Ghanizadeh, A., Mohammadi, M. R. et Yazdanshenas, A. (2006). Psychometric properties of the Farsi translation of the Kiddie Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version. *BMC Psychiatry*, 6, 10.
- Goldstein, N. E., Arnold, D. H., Weil, J., Mesiarik, C. M., Peuschold, D., Grisso, T. et Osman, D. (2003). Comorbid symptom patterns in female juvenile offenders. *International Journal of Law and Psychiatry*, 26(5), 565-582. doi: 10.1016/s0160-2527(03)00087-6
- Goodman, R., Ford, T., Richards, H., Gatward, R. et Meltzer, H. (2000). The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry*, 41, 645-656.
- Goyette, M. et Nadeau, L. (2008). Utilisation pathologique d'Internet : une intégration des connaissances [Pathological internet use : Integration of knowledge]. *Alcoologie et addictologie*, 30(3), 275-283.
- Granero Perez, R., Ezpeleta Ascaso, L., Domenech Massons, J. M. et de la Osa Chaparro, N. (1998). Characteristics of the subject and interview influencing the test-retest reliability of the Diagnostic Interview for Children and Adolescents-Revised. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 39(7), 963-972.
- Griffiths, M. et Sutherland, I. (1998). Adolescent gambling and drug use. *Journal of community & applied social psychology*, 8(6), 423-427.
- Griffiths, M., Wood, R. T. A., Gupta, R. et Derevensky, J. L. (2000). Risk factors in adolescence : The case of gambling, videogame playing, and the Internet. *Journal of gambling studies*, 16(2-3), 199-225.

- Griffiths, M. D., Davies, M. N. O. et Chappell, D. (2004). Online computer gaming: A comparison of adolescent and adult gamers. *Journal of adolescence*, 27(1), 87-96.
- Grisso, T. (2004). *Double jeopardy: Adolescent offenders with mental disorders* (Chicago: University of Chicago Press).
- Grisso, T. (2008). Adolescent offenders with mental disorders. *Future Child*, 18(2), 143-164.
- Grisso, T., Vincent, G. et Seagrave, D. (2005). *Mental health screening and assessment in juvenile justice*. New York: Guilford Press.
- Grisso, T. et Vincent, G. M. (2005). The context for mental health screening and assessment. Dans T. Grisso, G. M. Vincent & D. Seagrave (dir.), *Mental health screening and assessment in juvenile justice* (p. 44-70). New York, NY, US: Guilford Press.
- Harrison, P. A., Fulkerson, J. A. et Beebe, T. J. (1998). DSM-IV substance use disorder criteria for adolescents: A critical examination based on a statewide school survey. *Am J Psychiatry*, 155(4), 486-492. doi: doi:10.1176/ajp.155.4.486
- Harthman, C. A., Gelhorn, H., Crowley, T. J., Sakai, J. T., Stalling, M., Young, S. E. et al. (2008). Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(2), 165-173.
- Heilman, R. O. (1973). *Early recognition of alcoholism and other drug dependence*. Center City, MN: Hazelden Foundation.
- Herjanic, B. et Reich, W. (1982). Development of a structured psychiatric interview for children - agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychology*, 10(3), 307-324. doi: 10.1007/bf00912324
- Higgins, J. et Green, S., (editors), (2009). *Cochrane Handbook of systematic review of intervention (version 5.0.2, updated sept. 2009)*. The Cochrane Collaboration.
- Ho, T.-p., Leung, P. W.-l., Lee, C.-c., Tang, C.-p., Hung, S.-f., Kwong, S.-l., . . . Shaffer, D. (2005). Test-retest reliability of the Chinese version of the Diagnostic Interview Schedule for Children-Version 4 (DISC-IV). *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 46(10), 1135-1138.
- Hodges, H., Stern, L., Cytryn, L. et McKnew, D. (1982). The Child Assessment Schedule (CAS) diagnostic interview : A report on reliability and validity. *This Journal*, 21, 468-473.
- Hodges, K., McKnew, D., Burbach, D. J. et Roebuck, L. (1987). Diagnostic concordance between the Child Assessment Schedule (CAS) and the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS) in an outpatient sample using lay interviewers. *Journal of the American Academy of Child & Adolescent Psychiatry*, 26(5), 654-661. doi: <http://dx.doi.org/10.1097/00004583-198709000-00008>
- Hoffmann, N. G., Bride, B. E., MacMaster, S. A., Abrantes, A. M. et Estroff, T. W. (2004). Identifying co-occurring disorders in adolescent population. *Journal of Addictive Diseases*, 23(4), 41-53.
- Hope, T. L., Adams, C., Reynolds, L., Powers, D., Perez, R. A. et Kelley, M. L. (1999). Parent vs. self-report: Contributions toward diagnosis of adolescent psychopathology. *Journal of*

- Psychopathology and Behavioral Assessment*, 21(4), 349-363. doi: 10.1023/a:1022124900328
- Hunsley, J. et Mash, E. J. (2008). Developing criteria for Evidence-Based Assessment : An introduction to assessments that work (*A guide to assessment that work* (p. 3-14). New York: Oxford University Press.
- Ingram, R. E. et Price, J. M. (2010). *Vulnerability to psychopathology risk across the lifespan*. (2nd^e éd.). New York: Guilford Press.
- Institute of Medicine (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academies Press.
- Jensen, P., Roper, M., Fisher, P., Piacentini, J. et et al. (1995). Test-retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1): Parent, child, and combined algorithms. *Archives of General Psychiatry*, 52(1), 61-71. doi: <http://dx.doi.org/10.1001/archpsyc.1995.03950130061007>
- Jewell, J., Handwerk, M., Almquist, J. et Lucas, C. (2004). Comparing the validity of clinician-generated diagnosis of conduct disorder to the Diagnostic Interview Schedule for Children. *Journal of Clinical Child and Adolescent Psychology*, 33(3), 536-546. doi: http://dx.doi.org/10.1207/s15374424jccp3303_11
- Kairouz, S., Boyer, S., Nadeau, L., Perreault, M. et Fiset, L. (2008). Troubles mentaux, toxicomanie et autres problèmes liés à la santé mentale chez les adultes québécois *Enquêtes sur la santé dans les collectivités canadiennes* (Vol. cycle 1.2). Québec: Institut sur la statistique du Québec.
- Kairouz, S., Nadeau, L., LoSiou, G., Streiner, D. L., Cairney, J. et Lesage, A. (2005). Area variations in the prevalence of substance use and gambling behaviours and problems in Quebec : A multilevel analysis. *Canadian journal of psychiatry*, 50(10), 591-598.
- Kaminer, Y. h. et Bukstein, O. G. (2008). *Adolescent substance abuse : Psychiatric comorbidity and high-risk behaviors*. New York: Routledge.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U. M. A., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(7), 980-988. doi: <http://dx.doi.org/10.1097/00004583-199707000-00021>
- Kazdin, A. E. (2005). Evidence-based assessment for children and adolescents: Issues in measurement development and clinical application. *Journal of Clinical Child & Adolescent Psychology*, 34(3), 548-558. doi: 10.1207/s15374424jccp3403_10
- Kessler, R. C., Angermeyer, M., Anthony, J. C., de Graaf, R., Demyttenaere, K., Gasquet, I., . . . W.H.O. World Mental Health Survey Conso. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*, 6(3), 168-176.
- Khan, K. S., Riet, G. t., Glanville, J., Sowden, A. J. et Kleijnen, J. (2001). Undertaking systematic reviews of research on effectiveness : CRD's guide for those carrying out or commissioning reviews (NHS, Trans.) *CRD Report Number 4* (2^e éd.): University of York.

- King, C. A., Katz, S. H., Ghaziuddin, N., Brand, E., Hill, E. et McGovern, L. (1997). Diagnosis and assessment of depression and suicidality using the NIMH Diagnostic Interview Schedule for Children (DISC-2.3). *Journal of Abnormal Child Psychology*, 25(3), 173-181.
- Kolaitis, G., Korpa, T., Kolvin, I. et Tsiantis, J. (2003). Schedule for Affective Disorders and Schizophrenia for school-age children-present episode (K-SADS-P): A pilot inter-rater reliability study for Greek children and adolescents. *European Psychiatry*, 18(7), 374-375. doi: <http://dx.doi.org/10.1016/j.eurpsy.2003.02.002>
- Kovacs M. (1985). The Interview Schedule for Children (ISC). *Psychopharmacol Bull*, 21, 991-994.
- Kovacs M. (1992). *Children's Depression Inventory (CDI) manual*. New York, NY: Multi-Health Systems.
- Lacher, D. et Gdowski, C. L. (1979). *Actuarial assessment of child and adolescent personality : An interpretive guide for th Personality Inventory for Child Profile*. Los Angeles, CA: Western Psychological Services.
- Ladouceur, R. (1996). The prevalence of pathological gambling in Canada. *Journal of Gambling Studies*, 12(2), 129-142. doi: 10.1007/BF01539170
- Landis, J. R. et Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174. doi: 10.2307/2529310
- Landry M. et Brochu S. (2012). Le concept d'addiction : L'ASAM et la tradition biopsychosociale du Québec. *RISQ-INFO*, 20(1).
- Langenbucher, J. W., Labouvie, E., Martin, C. S., Sanjuan, P. M., Bavly, L., Kirisci, L. et al. (2004). An application of item response theory analysis to alcohol, cannabis and cocaine criteria in DSM-IV. *Journal of Abnormal Child Psychology*, 113 (1), 72-80.
- Larzelere, R. E. (1996). Inter-coder reliabilities and construct groupings for important codes on the Daily Incident Report. (Residential Research Technical Report 96-1). Boys Town.
- Lauth, B., Arnkelsson, G. B., Magnusson, P., Skarpheinnsson, G. A., Ferrari, P. et Petursson, H. (2010). Validity of K-SADS-PL (Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version) depression diagnoses in an adolescent clinical population. *Nordic Journal of Psychiatry*, 64(6), 409-420. doi: <http://dx.doi.org/10.3109/08039481003777484>
- Laveault, D. et Grégoire, J. (2002). *Introduction aux théories des tests: En psychologie et en sciences de l'éducation*. (2^e éd.). Bruxelles: Éditions De Boeck Université.
- Laventure, M., Déry, M., Pauzé, R. et Toupin, J. (2008). Caractéristiques associées à la cooccurrence du trouble des conduites et de problèmes de consommation de psychotropes chez les garçons et les filles. [Characteristics associated with concomitant conduct disorders and substance abuse among males and females.]. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 40(1), 11-20. doi: 10.1037/0008-400x.40.1.11
- Legendre, R. (1993). *Dictionnaire actuel de l'éducation (2e Éd.)*. Montréal: Guérin, éditeur.

- Lemos, I. T., Faisca, L. M. et Valadas, S. T. (2011). Assessment of psychopathological problems in the school context: The psychometric properties of a portuguese version of the adolescent psychopathology scale-short form. *Journal of Psychoeducational Assessment*, 29(1), 63-74. doi: 10.1177/0734282910362250
- Leslie, K. (2008). Youth substance use and abuse: challenges and strategies for identification and intervention. *Canadian Medical Association Journal*, 178(2), 145-148. doi: 10.1503/cmaj.071410
- Lewczyk, C. M., Garland, A. F., Hurlburt, M. S., Gearity, J. et Hough, R. L. (2003). Comparing DISC-IV and clinician diagnoses among youths receiving public mental health services. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42(3), 349-356.
- Lewis, D. O., Pincus, J. H., Shanok, S. S. et Glaser, G. H. (1982). Psychomotor epilepsy and violence in a group of incarcerated adolescent boys. *Am J Psychiatry*, 139, 882-887.
- Lingler, J., Bedoya, D. et Findling, R. L. (2007). *The Kiddie-SADS Present and Lifetime Version Plus (K-SADS-PL-W)*. [Unpublished measure]. Document inédit.
- Loeber, R., Burke, J. D., Lahey, B. B., Winters, A. B. A. et Zera, M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years, part I. *J Am Acad Child Adolesc Psychiatry*, 39, 1468-1484.
- Marmorstein, N. R. (2010). Longitudinal associations between depressive symptoms and alcohol problems: The influence of comorbid delinquent behavior. *Addictive Behaviors*, 35(6), 564-571. doi: 10.1016/j.addbeh.2010.01.004
- Martin, C. S., Chung, T., Kirisci, L. et Langenbucher, J. W. (2006). Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents : Implications for DSM-V. *Journal of Abnormal Child Psychology*, 115(4), 807-814.
- McGrath, A. M. (2000). The validity of the ADHD section of the Diagnostic Interview Schedule for Children. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 60(11-B), 5782.
- Mee-Lee, D., Shulman, G. D., Fishman, M. J., Gastfriend, D. R. et Miller, M. M., eds (2013). *The ASAM Criteria : Treatment Criteria for Addictive, Substance-Related, and Co-Occurring Condition*. Carson City, NV: The Change Companies.
- Mee-Lee, D., Shulman, G. D., Gastfriend, D. R. et Griffith, J. H. (2001). *ASAM patient placement criteria for the treatment of substance-related disorders, Second edition-revised (ASAM PPC-2R)*. Chevy Chase, MD : American Society of Addiction Medicine.
- Merikangas, K. R., Avenevoli, S., Costello, E. J., Doreen, K. et Kessler, R. C. (2009). National comorbidity survey replication adolescent supplement (NCS-A): I. Background and measures. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(4), 367-379. doi: 10.1097/CHI.0b013e31819996f1
- Milich, R., Roberts, M. A., Loney, J. et Caputo, J. (1980). Differentiating practice effects and statistical regression on the Conners Hyperkinesis Index. *J Abnorm Child Psychol* 8, 549-552.
- Miller, W. R. (1996). *Form 90: Structured assessment interview for drinking and related behavior*. Washington, DC: Department of Health and Human Services.

- Miller, W. R. et Boca, F. K. D. (1994). Measurement of drinking behavior using the Form 90 family of instruments. *Journal of Studies on Alcohol, Supplement(s12)*, 112-118. doi: doi:10.15288/jsas.1994.s12.112
- Moreno, A. M., Basurto, F. Z., Becerra, I. G. et San Pedro, E. M. (2006). Discriminant and criterion validity of the Spanish version of the Children's Interview for Psychiatric Syndromes-Parents' Version (P-ChIPS). *European Journal of Psychological Assessment*, 22(2), 109-115. doi: http://dx.doi.org/10.1027/1015-5759.22.2.109
- Morey, L. (1991). *Personality Assessment Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. (2007). *Personality Assessment Inventory – Adolescent professional manual*. Lutz, FL: Psychological Assessment Resources.
- Morrall, A., McCaffrey, D., Ridgeway, G., Mukherji, A. et Beighley, C. (2006). *The relative effectiveness of 10 adolescent substance abuse treatment programs in the United States*. Santa Monica, CA: Rand Corporation Technical Report (TS-346-CSAT).
- National Institute on Drug Abuse (NIDA). (2010). Comorbidity : Addiction and mental illnesses. *Research Report Series*.
- Nottelmann, E. (2001). National Institute of Mental Health research roundtable on prepubertal bipolar disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(8), 871-878.
- Obel, C., Heiervang, E., Rodriguez, A., Heyerdah, I. S., Smedje, H., Sourander, A. et et al. (2004). The Strengths and Difficulties Questionnaire in the Nordic countries. *Eur Child Adolesc Psychiatry*, 13(suppl 2), 1132-1139.
- Organisation mondiale de la Santé. (1946). Préambule à la Constitution de l'Organisation mondiale de la Santé, tel qu'adopté par la Conférence internationale sur la Santé, New York, 19-22 juin 1946, signé le 22 juillet 1946 par les représentants de 61 Etats. (Actes officiels de l'Organisation mondiale de la Santé, n°. 2, p. 100) et entré en vigueur le 7 avril 1948.
- Paing, W. W., Weller, R. A., Dixon, T. A. et Weller, E. B. (2010). Face-to-face versus telephone administration of the parent's version of the Children's Interview for Psychiatric Syndromes (P-ChIPS). *Current Psychiatry Reports*, 12(2), 122-126.
- Pearson, C., Janz, T. et Ali., J. (2013). « Troubles mentaux et troubles liés à l'utilisation d'une substance au Canada » Coup d'oeil sur la santé. *Septembre, produit no 82-624-X au catalogue de Statistique Canada*.
- Pearson, K. (1896). Mathematical contribution to the theory of evolution. III : Regression, heredity and panmixia. *Philosophical Transactions, A*, 187, 252-318.
- Perepletchikova, F., Krystal, J. H. et Kaufman, J. (2008). Practitioner review: Adolescent alcohol use disorders: Assessment and treatment issues. *J Child Psychol Psychiatry*, 49(11), 1131-1154. doi: 10.1111/j.1469-7610.2008.01934.x
- Petticrew, M. et Roberts, H. (2006). *Systematic reviews in the social sciences; A practical guide*. Meldon : Blackwell publishing.
- Piacentini, J., Shaffer, D., Fisher, P. W., Schwab-Stone, M. et et al. (1993). The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): III. Concurrent criterion

- validity. *Journal of the American Academy of Child & Adolescent Psychiatry*, 32(3), 658-665. doi: <http://dx.doi.org/10.1097/00004583-199305000-00025>
- Pica , L. A., Traoré, I., Camirand, H., Laprise, P., Bernèche, F., Berthelot, M., . . . et autres. (2013). L'Enquête québécoise sur la santé des jeunes du secondaire 2010-2011. Le visage des jeunes d'aujourd'hui : leur santé mentale et leur adaptation sociale, Tome 2 (p. 141). Québec: Institut de la statistique du Québec.
- Poznanski, E. O., Grossman, J. A., Buchsbaum, Y., Banegas, M., Freeman, L. et Gibbons, R. (1984). Preliminary studies of the reliability and validity of the Children's Depression Rating Scale. *Journal of the American Academy of Child and Adolescent Psychiatry*, 23(191-197).
- Rahdert, E. R., National Institute on Drug Abuse, Pacific Institute for Research Evaluation et Westover Consultants (1991). *The Adolescent assessment/referral system manual*. Rockville, MD: U.S. Dept. of Health and Human Services, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute on Drug Abuse.
- Rathman, F. S. (2008). *A comparison of diagnostic interviews for children*. (The Ohio State University, Ohio, US).
- Regier, D. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L. et Goodwin, F. K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA*, 264, 2511 - 2518.
- Reich, W. (2000). Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(1), 59-66.
- Reich, W., Cottler, L., McCallum, K., Corwin, D. et VanEerdewegh, M. (1995). Computerized interviews as a method of assessing psychopathology in children. *Comprehensive psychiatry*, 36(1), 40-45.
- Reynolds, W. M. (1987). *Reynolds Adolescent Depression Scale: Professional manual* Odessa, FL: Psychological Assessment Resources.
- Reynolds, W. M. (1998). *Adolescent psychopathology scale : Administration and interpretation manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Ribera, J. C., Canino, G., Rubio-Stipec, M., Bravo, M., Bauermeister, J. J., Alegria, M., . . . Shrout, P. E. (1996). The Diagnostic Interview Schedule for Children (DISC-2.1) in Spanish: Reliability in a Hispanic population. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 37(2), 195-204.
- RISQ (2007). *Grille de dépistage de consommation problématique d'alcool et de drogues chez les adolescents et adolescentes. Version 3.2*. Montréal: Recherche et intervention sur les substances psychoactives-Québec.
- Roberts, R. E., Solovitz, B. L., Chen, Y. W. et Casat, C. (1996). Retest stability of DSM-III-R diagnoses among adolescents using the Diagnostic Interview Schedule for Children (DISC-2.1C). *Journal of Abnormal Child Psychology*, 24(3), 349-362.
- Robins L. N., Cottler, L. B. et Babor, T. (1986). *The WHO/ADAMHA CIDI-SAM Interview, Revised 1987, 1988, 1989*. St.Louis.

- Robins, L. N., Helzer, J. E., Ratcliff, K. S. et Seyfried, W. (1982). Validity of the Diagnostic Interview Schedule, version II : DSM-III diagnoses. *Psychological Medicine*, 12(4), 855-870.
- Rothen, S., Vandeleur, C. L., Lustenberger, Y., Jeanpretre, N., Ayer, E., Gamma, F., . . . Preisig, M. (2009). Parent-child agreement and prevalence estimates of diagnoses in childhood: Direct interview versus family history method. *International Journal of Methods in Psychiatric Research*, 18(2), 96-109. doi: <http://dx.doi.org/10.1002/mpr.281>
- Rubio-Stipec, M., Canino, G. J., Shrout, P., Dulcan, M., Freeman, D. et Bravo, M. (1994). Psychometric properties of parents and children as informants in child psychiatry epidemiology with the Spanish Diagnostic Interview Schedule for Children (DISC.2). *Journal of Abnormal Child Psychology*, 22(6), 703-720.
- Rush, B., Urbanoski, K., Bassani, D., Castel, S., Wild, T. C., Strike, C., . . . Somers, J. (2008). Prevalence of co-occurring substance use and other mental disorders in the Canadian population. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, 53(12), 800-809.
- Saban, A. et Flisher, A. J. (2010). The association between psychopathology and substance use in young people: A review of the literature. *Journal of Psychoactive Drugs*, 42(1), 37-47.
- Samet, S., Waxman, R., Hatzenbuehler, M. et Hasin, D. S. (2007). Assessing addiction : Concepts and instruments. *Addiction Science and Clinical Practice*, 4(1), 19-31.
- Santé Canada. (2012). Enquête de surveillance canadienne de la consommation d'alcool et de drogues ; http://www.hc-sc.gc.ca/hc-ps/drugs-drogues/stat/_2012/summary-sommaire-fra.php.
- Schwab-Stone, M., Fisher, P. W., Piacentini, J., Shaffer, D. et et al. (1993). The Diagnostic Interview Schedule for Children-Revised version (DISC-R): II. Test-retest reliability. *Journal of the American Academy of Child & Adolescent Psychiatry*, 32(3), 651-657. doi: <http://dx.doi.org/10.1097/00004583-199305000-00024>
- Schwab-Stone, M. E., Shaffer, D., Dulcan, M. K., Jensen, P. S. et et al. (1996). Criterion validity of the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3). *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(7), 878-888. doi: <http://dx.doi.org/10.1097/00004583-199607000-00013>
- Shaffer, D., Fisher, P., Dulcan, M. K., Davies, M., Piacentini, J., Schwab-Stone, M. E., . . . Regier, D. A. (1996). The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): description, acceptability, prevalence rates, and performance in the MECA Study. Methods for the Epidemiology of Child and Adolescent Mental Disorders Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(7), 865-877.
- Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K. et Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses.

- Journal of the American Academy of Child and Adolescent Psychiatry*, 39(1), 28-38. doi: 10.1097/00004583-200001000-00014
- Shaffer, D., Schwab-Stone, M., Fisher, P., Cohen, P., Piacentini, J., Davies, M., . . . Regier, D. (1993). The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): I. Preparation, field testing, interrater reliability, and acceptability. *Journal of the American Academy of Child & Adolescent Psychiatry*, 32(3), 643-650.
- Shahrivar, Z., Kousha, M., Moallemi, S., Tehrani-Doost, M. et Alaghband-Rad, J. (2010). The reliability and validity of Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Life-time Version-Persian Version. *Child and Adolescent Mental Health*, 15(2), 97-102. doi: <http://dx.doi.org/10.1111/j.1475-3588.2008.00518.x>
- Shanee, N., Apter, A. et Weizman, A. (1997). Psychometric properties of the K-SADS-PL in an Israeli adolescent clinical population. *Israel Journal of Psychiatry and Related Sciences*, 34(3), 179-186.
- Sherrill, J. T. et Kovacs, M. (2000). Interview schedule for children and adolescents (ISCA). *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(1), 67-75.
- Sobell, L. C. et Sobell, M. B. (1996). *Timeline Followback user's guide: A calendar method for assessing alcohol and drug use*. Toronto, Ontario, Canada: Addiction Research Foundation.
- Steinberg, L. (2002). Clinical adolescent psychology: What it is, and what it needs to be. *Journal of Consulting and Clinical Psychology*, 70(1), 124-128. doi: 10.1037//0022-006x.70.1.124
- Substance Abuse and Mental Health Services Administration et Center for Behavioral Health Statistics Quality. (2014). Results from the 2013 national survey on drug use and health: Detailed tables. Rockville, MD: United States Department of Health and Human Services.
- Sung, M. J., Erkanli, A., Angold, A. et Costello, E. J. (2004). Effects of age at first substance use and psychiatric comorbidity on the development of substance use disorders. *Drug and Alcohol Dependence*, 75(3), 287-299. doi: 10.1016/j.drugalcdep.2004.03.013
- Swenson, L. P., Esposito-Smythers, C., Hunt, J. I., Hollander, B. L., Dyl, J., Rizzo, C. J., . . . Spirito, A. (2007). Validation of the Children's Interview for Psychiatric Syndromes (ChIPS) with psychiatrically hospitalized adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(11), 1482-1490. doi: <http://dx.doi.org/10.1097/chi.0b013e31814cef0e>
- Sylvester, C. E., Hyde, T. S. et Reichler, R. J. (1987). The Diagnostic Interview for Children and personality Inventory for Children in Studies of Children at Risk for Anxiety Disorders or Depression. *Journal of the American Academy of Child & Adolescent Psychiatry*, 26(5), 668-675. doi: <http://dx.doi.org/10.1097/00004583-198709000-00010>
- The Campbell Collaboration. (2011). The campbell collaboration library of systematic reviews. Repéré le 10 mars 2001 : <http://www.campbellcollaboration.org/library.php>
- The Campbell Collaboration. (2015). <http://www.campbellcollaboration.org/>. Repéré le 26 avril 2015










- The Change Companies. (2014). <https://www.changecompanies.net/products/product.php?id=PT-V>. Repéré le 7 avril 2014
- The Cochrane Collaboration. (2015). <http://www.cochrane.org/what-is-cochrane-evidence>. Repéré le 26 avril 2015
- The Cochrane Library. (2011a). Cochrane Database of systematic reviews (CDSR). Repéré le 10 mars 2011 : http://onlinelibrary.wiley.com/o/cochrane/cochrane_clsystrev_articles_fs.html
- The Cochrane Library. (2011b). Database of abstracts of reviews on effect (DARE). Repéré le 10 mars 2011 : http://onlinelibrary.wiley.com/o/cochrane/cochrane_cldare_articles_fs.html
- The Cochrane Library. (2011c). Health Technology Assessment (HTA). Repéré le 10 mars 2011 : http://onlinelibrary.wiley.com/o/cochrane/cochrane_clhta_articles_fs.html
- Traoré, I., Pica, L. A., Camirand, H., Cazale, L., Berthelot, M. et Plante, N. (2014). Enquête québécoise sur le tabac, l'alcool, la drogue et le jeu chez les élèves du secondaire, 2013. Évolution des comportements au cours des 15 dernières années. (p. 208). Québec: Institut de la statistique du Québec.
- Tremblay, J., Allaire, G., Dufresne, J., Lecavalier, M. et Neveu, Y. (2004). *Les Centres de réadaptation pour personnes alcooliques et autres toxicomanes. Chefs de file des services en toxicomanie et jeu pathologique*. Montréal: Fédération québécoise des centres de réadaptation pour personnes alcooliques et toxicomanes.
- Tupker, E. (2004). *First contact : A brief treatment for youth substance users with mental health problems*. Toronto: Centre for Addiction and Mental Health.
- Turner, E. A. (2014). Use of the PAI-A in the diagnosis and treatment of psychiatric disorders: A clinical application. *Ann Psychiatry Ment Health* 2(1), 1007.
- Vaughn, M. G. et Howard, M. O. (2004). Adolescent substance abuse treatment: A synthesis of controlled evaluations. *Research on Social Work Practice*, 14(5), 325-335. doi: 10.1177/1049731504265834
- Vincent, G. M., Grisso, T., Terry, A. et Banks, S. (2008). Sex and Race differences in mental health symptoms in juvenile justice: The MAYSI-2 national meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(3), 282-290. doi: 10.1097/chi.0b013e318160d516
- Vitaro, F., Rorive, M., Zoccolillo, M., Éliça, R. et Tremblay, R.-E. (2001). Consommation de substances psychoactives, troubles du comportement et sentiments dépressifs à l'adolescence. *Santé mentale au Québec*, 26(2), 106-131.
- Wechsler, D. (1995). *Échelle de Wechsler pour la période préscolaire et primaire. Forme révisée*. Paris: Les Éditions du Centre de psychologie appliquée.
- Weinstein, S. R., Stone, K., Noam, G. G., Grimes, K. et Schwab-Stone, M. (1989). Comparison of DISC with clinicians' DSM-III diagnoses in psychiatric inpatients. *Journal of the American Academy of Child & Adolescent Psychiatry*, 28(1), 53-60.
- Weissman M.M., Wickramaratne P., Warner V. et al., e. (1987). Assessing psychiatric disorders in children: Discrepancies between mothers' and childrens' reports.

- Archives of General Psychiatry*, 44(8), 747-753. doi: 10.1001/archpsyc.1987.01800200075011
- Weller, E. B., Weller, R. A., Fristad, M. A. et Rooney, M. T. (1999). *Administration manual Children's Interview for Psychiatric Syndromes (ChIPS)*. Washington, DC American Psychiatric Press.
- Weller, E. B., Weller, R. A., Fristad, M. A., Rooney, M. T. et Schecter, J. (2000). Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(1), 76-84. doi: <http://dx.doi.org/10.1097/00004583-200001000-00019>
- Welner, Z., Reich, W., Herjanic, B., Jung, K. G. et Amado, H. (1987). Reliability, validity, and parent-child agreement studies of the Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child & Adolescent Psychiatry*, 26(5), 649-653.
- Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., . . . Quadas- Group. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, 155(8), 529-536. doi: 10.7326/0003-4819-155-8-201110180-00009
- Winters, K. et Henly, G. (1986). The chemical dependency adolescent assessment project: Progress update. *Digest of Alcoholism Theory and Application*, 5, 56-64.
- Winters, K. C. et Henly, G. A. (1989). *The Personal Experience Inventory test and manual*. Los Angeles: Western Psychological Services.
- Winters, K. C. et Henly, G. A. (1993). *Adolescent Diagnostic Interview Schedule and Manual*. Los Angeles: Western Psychological Services.
- Winters, K. C., Latimer, W. et Stinchfield, R. D. (1999). The DSM-IV criteria for adolescent alcohol and cannabis use disorders. *Journal of Studies on Alcohol*, 60(3), 337-344.
- Winters, K. C., Martin, C. S. et Chung, T. (2011). Substance use disorders in DSM-V when applied to adolescents. *Addiction*, 106(5), 882-884. doi: <http://dx.doi.org/10.1111/j.1360-0443.2010.03334.x>
- Winters, K. C., Stinchfield, R. D. et Bulkstein, O. G. (2008). Assessing adolescent substance use and abuse. Dans Y. H. Kaminer & O. G. Bukstein (dir.), *Adolescent substance abuse : Psychiatric comorbidity and high-risk behaviors* (p. 53-86). New York: Routledge.
- Winters, K. C., Stinchfield, R. D., Fulkerson, J. et Henly, G. A. (1993). Measuring alcohol and cannabis use disorders in an adolescent clinical sample. *Psychology of Addictive Behaviors*, 7(3), 185-196. doi: <http://dx.doi.org/10.1037/0893-164X.7.3.185>
- Wirt, R. D., Lacher, D., Klinedinst, J. K. et Seat, P. D. (1977). *Multidimensional description of child personality : A manual for the Personality Inventory for Children*. Los Angeles, CA: Western Psychological Services.
- Witwer, A. N., Lecavalier, L. et Norris, M. (2012). Reliability and Validity of the "Children's Interview for Psychiatric Syndromes-Parent Version" in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 42(9), 1949-1958.

- Wood, R. T. A., Griffiths, M. D., Derevensky, J. L. et Gupta, R. (2002). Adolescent accounts of the UK National Lottery and scratchcards: An analysis using Q-sorts. *Journal of gambling studies*, 18(2), 161-183.
- Wu, P., Goodwin, R. D., Fuller, C., Liu, X. H., Comer, J. S., Cohen, P. et Hoven, C. W. (2010). The relationship between anxiety disorders and substance use among adolescents in the community: Specificity and gender differences. *Journal of Youth and Adolescence*, 39(2), 177-188. doi: 10.1007/s10964-008-9385-5








Annexes

Annexe A : Légende des Tableaux diagnostiques 1 à 6 (annexes B à G)

-  : échantillon clinique externe
- !! : échantillon clinique externe intensif
-  : échantillon clinique interne
-  : communauté générale
-  : communauté scolaire
-  : traitement santé mentale
-  : judiciaire
-  : traitement SPA
- # : traitement TC
- ‡ : Dx autisme
- : centre pour jeunes
-  : intervalle
-  : intervieweur
- Cpt délinq : comportement délinquant
- ÉC : évaluation clinique
- Dép : dépendance
- Dx : diagnostic
- Sx : symptômes
- Tx : traitement
- OAD : *over anxious disorder*
- ∅ : aucun
- : K non disponible lorsque 100% d'accord sur absence de diagnostic/aucune donnée/résultat/moins de cinq cas/0
- Rangée ombragée : étude non retenue selon l'analyse méthodologique de la grille QUADAS-2
- Q-2 : numéro de l'étude du Tableau XII : Analyse des études retenues à l'aide de la grille QUADAS-2

Annexe B : Tableau 1, Trouble lié à l'utilisation de substance psychoactive (TUS)

Voir la légende des Tableaux à l'Annexe A.

Troubles liés à l'utilisation de substance psychoactive (TUS)								
Q-2	Instrument étude échantillon	Fidélité			Validité			
		Inter juge	Test-retest	Cohérence interne	Concordance	Concurrente	Convergente	Discriminante
1	ADI Winter 1993 	12-19 ans (n=72)  (sur enregistrement)  2 Info ado K Abus alcool 0,86 Dépendance alcool 0,53 Abus cannabis 0,76 Dépendance cannabis 0,68	12-19 ans (n=49)  7 jrs.  ? Info ado K Dépendance alcool 0,83 Dépendance cannabis 0,80			12-19 ans (n=50)  3 jrs Info ado ADI vs EC K Abus alcool 0,71 Dépendance alcool 0,82 Abus cannabis 0,80 Dépendance Cannabis 0,79	12-19 ans (n=276)  7jrs Info ado Correlational (eta) analysis ⁶⁸ Alcool DUF ⁶⁹ (n=276) 12ms 0,68 6ms 0,58 Cotation du clinicien ⁷⁰ (n=115) 0,60 Heilman's warning ⁷¹ (n=115) 0,75 PICS ⁷² (n=124) 0,65 Cannabis DUF (n=276)	













⁶⁸ *correlational analysis* : indique le degré de corrélation entre variables

















⁶⁹ *Drug Use Frequency*















⁷⁰ Évaluation de la sévérité de la problématique selon le clinicien évaluateur.

⁷¹ *Heilman's warning signs Checklist* (Heilman, 1973)

⁷² *Personal Involvement with Chemicals Scale* dans PEI (Winters et Henly, 1989)














							12ms 0,68 6ms 0,28 Cotation du clinicien (n=115) 0,71 Heilman's warning (n=115) 0,73 PICS (n=124) 0,70	
2	APS Lemos 2011  Version portugaise		12-19 ans (n=68)  12 sem.  ? Info ado Abus de SPA r 0,42	12-19 ans α Abus de SPA Portugal (n=656) ÉU (n=1827) 0,70 0,85				
3	CAPA Angold 1995 		10-18 ans (n=77)  moyen 5 jrs  2 Info ado K Abus et dépendance de SPA 1,00					
6	ChIPS Fristad 1998a 				13-18 ans (n= 19)  moins 7 jrs Info ado ChIPS vs DICA Abus de SPA K % 0,627 84	6-18 ans (n=47)  ? Info ado ChIPS vs ÉC Abus de SPA K % 0,676 94		
7	ChIPS Fristad 1998b  dont n=18 endeuillés				6-18 ans (n=40)  même jr Info ado ChIPS vs DICA K %	6-18 ans (n=40)  (sur dossier) Info ado ChIPS vs ÉC K %		

					Abus alcool – 1,00 Abus drogue – 1,00	Abus alcool – 1,00 Abus drogue – 1,00		
8	ChIPS Moreno 2006  + vs  version espagnole					6-18 ans ( n=31)  30 jrs Info parent-ado ChIPS vs EC K % – 1,00		6-18 ans  (n=43)  7jrs vs  + (n=31)  30jrs Non concluant vs nbre de Dx (4,6%/0%)
9	ChIPS Paing 2010  +		Âge moyen 12.2 n=12  ?  ? Info parent Tél. vs face à face Abus de SPA (marijuana) % 100					
10	ChIPS Rathman 2008  +				11-14 ans (n=10)  7jrs Info parent-ado ChIPS vs K-SADS Abus de SPA K % – 1,00 Info ado – 1,00	11-14 ans (n=10)  (sur dossier) Info parent-ado ChIPS vs EC Abus de SPA K % – 1,00		
11	ChIPS Swenson 2007  +				13-18 ans (N=93)  3 jrs Info ado ChIPS vs K-SADS K %			

					0,66	87		
16	DICA Granero 1998  Version espagnole		13-17 ans (n=52)  moyen 11 jrs  2 Info ado K Abus/utilisation alcool 0,801 Colle prisée 0,994 Marijuana 0,956 Drogues de rue 0,974					
23	DISC Bravo 2001  Version espagnole		11-17 ans (n=85)  moyen 12jrs  4 Toute SPA K Info parent Dernière an 0,22 À vie 0,46 Info ado (n=83) Dernière an 0,80 À vie 0,80					
25	DISC Crowley 2001  Certains  Vs 						14-18 ans (n=variable de 149 à 172)  ? Info parent-ado  Concluant Vs <i>SumDep</i> ⁷³ (CIDI ⁷⁴) Vs échelle <i>Days of use</i> (DISC)	14-18 ans Concluant vs prévalence Abus de SPA  n=87 60,9%  n=85 7,1% Dép SPA  78,2%









⁷³ *Sum of all reported SD symptoms* (somme de tous les Sx de dépendance)

⁷⁴ CIDI-SAM : *Composite International Diagnostic Interview-Substance Abuse Module* (Robins L. N., Cottler et Babor, 1986)









								r_s 0,67	 4,7% Dép. vs <i>SumDep</i> !!  13,6  0,6
26	DISC Dillon 2005  (80%),  (16%),  (4%) Minorité ethnique	12-17 ans  18 ms  ? Score composé Dx abus/dép Info parent vs ado r Marijuana Temps 0(n= 181) 0,32 18 ms (n=102) 0,23 Cocaïne Temps 0 (n= 181) 0,21 18 ms (n=102) 0,13						12-17 ans Info ado Éval temps 0 (n= 190) 6 ms (n=141) 18 ms (n=111) Conculant vs ADAD ⁷⁵ vs TLFB ⁷⁶	
38	DISC Roberts 1996    ● Minorité ethnique		12-17 ans (n=101)  7 à 14 jrs  2 Info ado K Y Abus-dép SPA 0,574 0,662 Abus alcool 0,532 0,669 Abus drogue 0,477 0,592 15-17 (n= ?) Abus-dép						

⁷⁵ *Adolescent Drug Abuse Diagnosis Scale* (Friedman A. S. et Utada A., 1989)

⁷⁶ *Timeline Followback*, méthode calendrier auto rapportée (Sobell et Sobell, 1996)

			0,73	-				
43	GAIN Dennis 2004 							12-18 ans (n=93) 89%  7jrs/98%  14 jrs Info ado Concluant vs Form 90 ⁷⁷ $r(p < 0.01)$ Alcool Nbre de jrs de conso 0,84 Nbre max de conso 0,75 Cannabis Nbre de jrs de conso 0,84 Nbre max de joints 0,75
	ISCA							
10	K-SADS Rathman 2008 							11-14 ans (n=10)  (sur dossier) Info parent-ado K-SADS vs EC Abus de SPA K % - 100
54	K-SADS Shanee 1997  Version hébreux	6-19 ans (n=57)  14 jrs Abus de drogue K ‡ 2 1,00 Mère-ado						6-19 ans (n=57)  après 2 mois de Tx Info parent-ado K-SADS Vs EC Abus de drogue

⁷⁷ Sorte de journal de bord de la consommation (Miller et Boca, 1994)

		0,47				K 0,73		
53	K-SADS Shahrivar 2010  Version perse					9-18 ans (n=102)  ? Info parent-ado K-SADS vs ÉC (BED ⁷⁸) K Dx actuel 0,34 Dx vie 0,23		
55	K-SADS Weissman 1987  mais avec parent avec Dx EDM vs sans Dx	6-23 ans (n=175) ≡ même jr  1 Mère vs ado Garçons (n=81) Parent Dx EDM Parent  Dx EDM	K 0,07 0,12 -					
56	PADDI Abrantes 2004 			13-18 ans (n=284) Dépendance (18 items) Abus (9 items)	α 0,889 0,861			
57	PADDI Hoffman 2004  certains 			12-18 ans (n=279) Dépendance (18 items) Abus (9 items)	α 0,870			


















⁷⁸ Best estimate diagnostic (meilleur estimé du diagnostic)








				0,790				
PAI-A								

Annexe C : Tableau 2, Trouble anxiété généralisé (TAG)

Voir la légende des Tableaux à l'Annexe A.

Trouble anxiété généralisé (TAG)								
Q-2	Instrument études échantillon	Fidélité			Validité			
		Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
	ADI							
2	APS Lemos 2011 Version portugaise		12-19 ans (n=68) 📅 12 sem. 👤 ? Info ado r 0,61	12-19 ans Portugal (n=656) ÉU (n=1827) α 0,83 0,88				
3	CAPA Angold 1995		10-18 ans (n=74) Info ado 📅 moyen 5 jrs 👤 2 K 0,79					
6	ChIPS Fristad 1998a				13-18 ans (n=19) 📅 moins 7 jrs Info ado ChIPS vs DICA K % 0,532 74	6-18 ans (n=47) 📅 ? Info ado ChIPS vs EC K % 0,460 85		
7	ChIPS Fristad 1998b 👤 dont n=18 endeuillés				6-18 ans (n=40) 📅 même jr Info ado ChIPS vs DICA K %	6-18 ans (n=40) 📅 (sur dossier) Info ado ChIPS vs EC K %		










					0,487	95	0,740	97.5	
8	ChIPS Moreno 2006  vs  Version espagnole						6-18 ans  (n=31)  30 jrs Info parent ChIPS vs EC K % 0,475 93		6-18 ans  (n=43)  7 jrs  (n=31)  30 jrs Info parent Concluant vs nbre de Dx (0%/9,7%) vs niveau de détérioration (2,3%/32,3%)
9	ChIPS Paing 2010  +		Âge moyen 12.2 (n=12)  ? ? ? Info parent Tél. vs face à face % 91.7						
10	ChIPS Rathman 2008  +						11-14 ans (n=10)  7 jrs Info parent-ado ChIPS vs K-SADS K % 0,74 90 Info ado 0,62 90		11-14 ans (n=10)  (sur dossier) Info parent-ado ChIPS vs EC K % 0,74 90
11	ChIPS Swenson 2007  +						12-18 ans (n=93)  3 jrs Info ado ChIPS vs K-SADS K % 0,29 73		
12	ChIPS Witwer 2012	6-17 ans (n=43)  (sur enregistrement)		6-17 ans (n=61) K-R 0,65			6-17 ans (n=61)  même jr Info parent-ado		















		<p>♂ 2</p> <p>Info ado</p> <p>K %</p> <p>0,86 95,4</p>			ChIPS vs CASI ⁷⁹			
					<p>K %</p> <p>0,50 78,7</p>			
13	<p>DICA</p> <p>Boyle, 1993</p> <p> mais à forte probabilité de Dx</p> <p>Version OAD</p>	<p>12-16 ans (n=137)</p> <p> 10 à 20 jrs</p> <p>♂ 4 (dont 2 spécialistes⁸⁰)</p> <p>Accord parent-ado</p> <p>K %</p> <p>Temps 1 0,06 77</p> <p>Temps 2 0,06 86</p>	<p>12-16 ans (n=137)</p> <p> 10 à 20 jrs</p> <p>♂ 4</p> <p>K %</p> <p>Info parent 0,69 96</p> <p>Info ado 0,54 88</p>					
14	<p>DICA</p> <p>Boyle, 1997</p> <p> mais à forte probabilité de Dx</p> <p>Version OAD</p>		<p>6-16 ans (n=210)</p> <p> moyen 17 jrs</p> <p>♂ ?</p> <p>Info parent</p> <p>K r</p> <p>0,57 0,61</p>				<p>6-16 ans (n=223)</p> <p> ?</p> <p>DICA vs OCHS-R⁸¹</p> <p>K</p> <p>0,31</p> <p>Info parent (n=213)</p> <p>Non concluant vs fonctionnement social détérioré et pauvres performances scolaires</p> <p>Info profs (n=184)</p> <p>Non concluant vs fonctionnement social détérioré</p> <p>Info parent-ado</p>	














⁷⁹ *Child and Adolescent(e) Symptom Inventory* (Gadow et Sprafkin, 2002)

⁸⁰ Pédopsychiatre et psychologue pour enfant.












⁸¹ *Revised Ontario Child Health Study scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

							(n=214) Concluant vs humeur dépressive	
16	DICA Ezpeleta 1997  + Version espagnole					13-17 ans  ? DICA vs EC K % Info ado (n=65) 0,21 74 Info parent (n=66) 0,47 89 Combinée (n=64) 0,26 73		
17	DICA Granero 1998  + Version espagnole		13-17 ans (n=52)  moyen 11 jrs ↓ 2 Info ado K 0,324					
19	DICA Reich 1995  + Teste version Informatique Version OAD		13-18 ans (n=?)  approximatif 5 jrs ↓ 2 Info ado K 0,40					
23	DISC Bravo 2001  +  + Version espagnole		11-17 ans (n=85)  moyen 12jrs ↓ 4 K Info parent Dernière an 0,39 Info ado (n=83) -					
24	DISC		6-14 ans (n= 260)	12-14 ans				

	Breton 1998  TAG + OAD version française		 7 et 21 jrs  4 K Info parents 0,57 Info ado 12-14 ans (n=145) 0,53	Info parents (n=148) 0,73 Info ado (n=145) Entre 0,70 et 0,80	á			
27	DISC Edelbrock 1999 		6-15 ans (n=24)  moyen 7,5 jrs  1 Info parent K 0,78					
37	DISC Ribera 1996  vs  Version espagnole	9-17 ans  14 jrs  4 K  (n=124) Accord non-cli. -0,03 Info parent - Info ado 0,09  (n=28) Accord non-cli. 0,77 Info parent 0,64 Info ado 0,64  (n=124) Accord non-cli./ pédopsychiatre 0,42 Info parent 0,74 Info ado -  (n=28) Accord non-cli./ pédopsychiatre 0,20						

		Info parent 0,38 Info ado 0,66					
38	DISC Roberts 1996  Minorité ethnique		12-17 ans (n=101)  7 à 14 jrs ↓ 2 Info ado K Y TAG 0,375 0,622 OAD 0,434 0,580				
39	DISC Rubio-Stipeč 1994  vs  Version espagnole	9-17 ans  moins de 15 jrs. ↓ 2 Info parent vs ado K  (n=248) 0,09  (n=28) -0,05					
41	DISC Schwab-Stone 1996  avec Dx vs  sans Dx Version OAD		9-18 ans (n=247)  1 à 15 jrs ↓ 4 K Info parent 0,60 Info ado 0,28		9-18 ans (n=247)  1 à 15jrs DISC vs ÉC K Info parent 0,60 Info ado 0,23 Combinée 0,51		
	GAIN						
	ISCA						
44	K-SADS Ambrosini 1989  Version OAD	6-17 ans (n=25)  (enregistrement vidéo) ↓ 2 Info parent-ado	6-17 ans (n=25) α SPE ⁸² Évaluateur 1 0,80 Évaluateur 2 0,81 SLW ⁸³				

⁸² Score Sx present illness 12 months (présence de Sx au cours des 12 derniers mois)

		K 0,85		Évaluateur 1 0,78 Évaluateur 2 0,76				
45	K-SADS Brasil 2010  + Version brésilienne						6-14 ans (n=78)  15 jrs Info parent-ado r 0,30 Légèrement concluant vs CBCL/4-18 internalisés ⁸⁴	
46	K-SADS Chambers 1985  +		6-17 ans (n=52)  moins de 3 jrs ↓ 4 Info combinées parent/ado/ évaluateur K 0,24					
47	K-SADS Ghanizadeh 2006  + plus  version Perse					4-19 ans (n=109)  ? Info parent-ado K-SADS vs ÉC K 0,781		
49	K-SADS Kaufman 1997  + vs 		7-17 ans (n=20)  moyen 17,9 jrs ↓ 2 Info parent-ado K Dx actuel 0,78 Dx vie 0,78				7-17 ans (n=66)  ? Info parent-ado Concluant vs SCARED-P ⁸⁵ ANX/POS = 25.5 ± 14.6;	

⁸³ Score Sx present week (presence de Sx au cours de la présente semaine)

⁸⁴ Child Behavior Checklist (Achenbach, 1991a)

⁸⁵ Screen for Child Anxiety Related Emotional Disorders Parent and child versions (Birmaher et al., 1997)

							ANX/NEG = 11.4 ± 11.4; $t_{63} = 4.32$, $P < .0001$ SCARED-C ANX/POS = 26.1 ± 14.7; ANX/NEG = 12.6 ± 9.3; $t_{56.4} = 4.47$, $P < .0001$ vs CBCL ⁸⁶ internalisé ANX/POS = 66.8 ± 11.5; ANX/NEG = 58.1 ± 13.7; $t_{63} = 2.76$, $P < .01$	
10	K-SADS Rathman 2008 ☞+					11-14 ans (n=10) 📅 (sur dossier) Info parent-ado K-SADS vs ÉC K % 1,00 80		
52	K-SADS Rothen 2009 ☞☞☞ mais avec parent avec Dx SM Version française	7-17 ans 📅? ♂ ? Y Ado vs mère (n=235) -0,01 Ado vs père (n=189) 0,00 Mère vs père						























⁸⁶ Child Behavior Checklist (Achenbach et Edelbrock, 1983)













		(n=128) 0,35 baseline vs suivi Info ado (n=60) 0,00 Info parent (n=89) 0,31						
53	K-SADS Shahrivar 2010 ☞+ Version perse					9-18 ans (n=102) 🗓️ même jr Info parent-ado K-SADS vs ÉC K Dx actuel 0,28 Dx vie 0,28		
55	K-SADS Weissman 1987 🗓️🗓️🗓️ mais avec parent avec Dx ÉDM vs sans Dx	6-23 ans (n=175) 🗓️? 👤 1 K Info mère vs ado 0,25 13-18 ans (n=70) Parent ÉDM 0,17 Parent ≠ ÉDM 0,24						
	PADDI							
	PAI-A							

Annexe D : Tableau 3, Épisode dépression majeure (ÉDM)









Voir la légende des Tableaux à l'Annexe A.

Épisode dépression majeure (ÉDM)								
		Fidélité			Validité			
Q-2	Instrument études	Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante

	échantillon							
	ADI							
2	APS Lemos 2011  Version portugaise		12-19 ans (n=68)  12 sem.  ? Info ado r 0,72	12-19 ans Portugal (n=656) ÉU (n=1827)	α 0,89 0,91			
3	CAPA Angold 1995  		10-18 ans (n=77)  moyen 5jrs  2 Info ado K 0,90					
6	ChIPS Fristad 1998a  				13-18 ans (n=19)  moins de 7 jrs Info ado ChIPS vs DICA K % 0,379 74	6-18 ans (n=47)  ? Info ado ChIPS vs EC K % 0,394 89		
7	ChIPS Fristad 1998b  dont n=18 endeuillés				6-18 ans (n=40)  même jr Info ado ChIPS vs DICA K % - 100	6-18 ans (n=40)  (sur dossier) Info ado ChIPS vs EC K % - 100		
8	ChIPS Moreno 2006  vs  Version espagnole					6-18 ans  (n=31)  30 jrs Info parent ChIPS vs EC K % - 93	6-18 ans  (n=43)  7 jrs  (n=31)  30 jrs Info parent Concluant vs nbre de Dx (2,3%/6,4%) vs niveau de détérioration	
















								(famille, école, amis) (2,3%/32,3%)
9	ChIPS Paing 2010  +		Âge moyen 12.2 (n=12)  ? ↓ ? Info parent Tél. vs face à face % 75					
10	ChIPS Rathman 2008  +				11-14 ans (n=10)  7 jrs Info parent-ado ChIPS vs K-SADS K % 1,00 100	11-14 ans (n=10)  (sur dossier) Info parent-ado ChIPS vs ÉC K % 1,00 100		
11	ChIPS Swenson 2007  +				12-18 ans (n=93)  3 jrs Info ado ChIPS vs K-SADS K % 0,49 75			
12	ChIPS Witwer 2012  +‡	6-17 ans (n=43)  (sur enregistrement) ↓ 2 Info ado K % 0,44 86.1		6-17 ans (n=61) K-R 0,83	6-17 ans (n=61)  même jr Info parent-ado ChIPS vs CASI ⁸⁷ K % 0,44 83.6			
13	DICA Boyle 1993	12-16 ans (n=137)  10 à 20 jrs ↓ 4 (dont 2 spécialistes ⁸⁸)	12-16 ans (n=137)  10 à 20 jrs ↓ 2 K %					

⁸⁷ Child and Adolescent(e) Symptom Inventory (Gadow et Sprafkin, 2002)

	 mais à forte probabilité de Dx	Info parent vs ado K % Temps 1 -0,05 91 Temps 2 0,13 95	Info parent 0,31 96 Info ado 0,45 95					
14	DICA Boyle 1997  mais à forte probabilité de Dx		6-16 ans (n=210)  moyen 17 jrs  ? Info parent K 0,70				6-16 ans (n=223)  ? DICA vs OCHS-R ⁸⁹ K 0,05 Info parent (n=213) Non concluant vs fonctionnement social détérioré et pauvres performances scolaires Info profs (n=184) Non concluant vs fonctionnement social détérioré Info parent-ado (n=214) Concluant vs humeur dépressive	6-16 ans (n=214)  ? Non concluant vs retard développemental
16	DICA Ezpeleta 1997 					13-17 ans  ? DICA vs EC K %		
























⁸⁸ Pédopsychiatre et psychologue pour enfant.

⁸⁹ *Revised Ontario Child Health Study scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

	Version espagnole					Info ado (n=67) 0,31 72 Info parent (n=65) 0,13 69 Combinée (n=66) 0,25 66		
19	DICA Reich 1995    Teste version informatique		13-18 ans (n=?)  approximatif 5 jrs  2 Info ado K 0,45					
20	DICA Sylvester 1987  parent avec Dx (anxiété ou dépression) vs  parent sans Dx	7-17 ans (n=74)  même jr  2 Info parent avec Dx vs parent  Dx du même ado K % 0,39 81 Info parent vs ado 0,11 63				7-17 ans (n=82)  moins de 15 jrs Info parent-ado DISC vs PIC ⁹⁰ K % 0,2413 70		
22	DISC Aronen 1993  					12-15 ans (n=163)  moyen 135,56 jrs ⁹¹ Info ado DISC Vs EC au congé K 0,186		
23	DISC Bravo 2001		11-17 ans (n=85)  moyen 12jrs					

⁹⁰ *Personality Inventory for Children* (Lacher et Gdowski, 1979; Wirt et al., 1977)

⁹¹ Durée moyenne de l'hospitalisation.

	 Version espagnole		 4 K Info parent Dernière an 0,31 Info ado Dernière an 0,15					
25	DISC Crowley 2001 !!  #  Certains  vs 	Âge moyen 16,1  ?  ? Info ado-parent K !!  (n=69)-0,017  (n=84) aucun					Âge moyen 16,1  ? Info parent-ado Concluant vs Carroll ⁹² (nbre de Sx) !!o : 15,7 n : 8,8	Âge moyen 16,1  ? Info parent-ado Concluant vs prévalence/nbre de Sx/Carroll !!  (n=87) 9,3%/1,3/11  (n=85) 3,5/ 0,6/ 4,9 vs % prévalence info parent !!  (n=69) 23,2  (n=84) 0,0
28	DISC Flisher 2012  plus 		Âge moyen 15 ans  approximatif 15 jrs  1 K Info parent 0,662 Info ado 0,661 Combinée 0,660					
30	DISC Ho 2005  Version		9-18 ans (n=79)  moyen 22.7 jrs  4 K Info parent 0,82					

⁹² Carroll Depression Self-Rating scale (Carroll, Feinberg, Smouse, Rawson et Greden, 1981)

	chinoise		Info ado 0,58 Combinée 0,61					
33	DISC King 1997 👤					12-18 ans (n=266) 📅? % de Dx Info parents 46,8 Info ado 36 Combinée 66,1 DISC vs ÉC 58,9	12-18 (n=266) 📅? Info parent-ado Concluant vs RADS ⁹³ vs CDRS-R ⁹⁴	
36	DISC Piacentini 1993 👤+				11-17 ans 📅 même jr DISC Vs CAF ⁹⁵ K Info parent (n=73) 0,36 Info ado (n=72) 0,39 Combinée (n=71) 0,41			
38	DISC Roberts 1996 👤👤+ 👤• Minorité ethnique		12-17 ans (n=101) 📅 7 à 14 jrs 👤 2 Info ado K Y 0,501 0,614					
39	DISC Rubio-Stipec 1994 👤👤+ Vs 👤 Version	9-17 ans 📅 moins de 15 jrs 👤 2 Info parent vs ado K				9-17 ans (👤 n=124/👤 n=46) 📅 (sur dossier) DISC vs ÉC(BED ⁹⁶)		

⁹³ Reynolds Adolescent Depression Scale (Reynolds, 1987)






⁹⁴ Children's Depression Rating Scale-Revised (Poznanski et al., 1984)

⁹⁵ Clinical Assessment Form (Piacentini et al., 1993)

⁹⁶ Best estimate diagnostic (meilleure estimation du Dx)









	espagnole	n= 248 0,18 n=28 0,25				K Info parent 0,33 Info ado 0,54 Combinée 0,43		
40	DISC Schwab-Stone 1993 		11-17 ans (n=39) 7 à 21 jrs 4 K Info parent 0,72 Info ado 0,77 Combinée 0,71	11-17 ans (n=39) 7 à 21 jrs α Info parent 0,88 Info ado 0,85				
41	DISC Schwab-Stone 1996 avec Dx vs sans Dx		9-18 ans (n=247) 1 à 15 jrs 4 K Info parent 0,55 Info ado 0,37 Combinée 0,48			9-18 ans (n=247) 1 à 15jrs DISC vs ÉC K Info parent 0,60 Info ado 0,79 Combinée 0,63		
42	DISC Weinstein 1989 					12-15 ans (n=163) max. 15 jrs Info ado DISC vs ÉC à l'admission K 0,17		
	GAIN							
	ISCA							
44	K-SADS Ambrosini 1989 	6-17 ans (n=25) (enregistrement vidéo) 2 Info parent-ado		6-17 ans (n=25) α 9 items SPE ⁹⁷ Évaluateur 1 0,79 Évaluateur 2 0,87 SLW ⁹⁸				

⁹⁷ Score Sx present illness 12 months (présence de Sx au cours des 12 derniers mois)

		K 0,80		Évaluateur 1 0,78 Évaluateur 2 0,83 12 items SPE Évaluateur 1 0,79 Évaluateur 2 0,87 SLW Évaluateur 1 0,78 Évaluateur 2 0,83 17 items SPE Évaluateur 1 0,81 Évaluateur 2 0,89 SLW Évaluateur 1 0,80 Évaluateur 2 0,84				
45	K-SADS Brasil 2010  Version brésilienne						6-14 ans (n=78)  15 jrs Info parent-ado r 0,30 Concluant (léger) vs CBCL/4-18 internalisés ⁹⁹	
46	K-SADS Chambers 1985 		6-17 ans (n=52)  moins de 3 jrs  4 Info combinée parent/ado/ évaluateur K 0,54	6-17 ans (n=52) <i>Depression summary scales</i> α 17 items 0,72 12 items 0,72 9 items 0,68 2items 0,84				

⁹⁸ Score Sx present week (présence de Sx au cours de la présente semaine)

⁹⁹ Child Behavior Checklist (Achenbach, 1991a)

49	K-SADS Kaufman 1997  vs 		7-17 ans (n=20)  moyen 17,9 jrs  2 Info parent-ado K Dx actuel 0,90 Dx vie 1,00				7-17 ans (n=27)  ? Info parent-ado Concluant vs CBCL ¹⁰⁰ internalisé <i>DEP/POS</i> = 68.7 ± 10.0 ; <i>DEP/NEG</i> = 58.4 \pm 13.6; $t_{62,4} = 3.54$, $P < .001$	
51	K-SADS Lauth 2010  Version islandaise						11-18 ans (n=86)  ? Concluant vs info ado YSR <i>internalizing</i> ¹⁰¹ <i>YSR anxious depressed</i> <i>SDQ-S emotional symptoms</i> ¹⁰² <i>CASS:L :emotional problems</i> ¹⁰³ <i>BYI depression</i> ¹⁰⁴ <i>CDI total</i> ¹⁰⁵ vs info parent CBCL <i>anxious-</i>	11-18 (n=86)  ? Non concluant vs Dx TAG et TC

¹⁰⁰ *Child Behavior Checklist* (Achenbach et Edelbrock, 1983)

¹⁰¹ *Youth Self-Report* (Achenbach, 1991b)

¹⁰² *Strengths and Difficulties Questionnaire* (Obel et al., 2004)

¹⁰³ *Conners/Wells Adolescent(e) Self-report Scales Long Version* (Conners et al., 1997)

¹⁰⁴ *Beck Youth Inventories of Emotional and Social Impairment* (Beck et al., 2001)

¹⁰⁵ *Children's Depression Inventory* (Kovacs M., 1992)









							<i>depressed</i>	
10	K-SADS Rathman 2008 ☞+						11-14 ans (n=10) 📅 (sur dossier) Info parent-ado K-SADS vs ÉC K % 1,00 80	
52	K-SADS Rothen 2009 ☞☞☞ mais avec parent avec Dx Version française	7-17 ans 📅? ☞ ? Y Ado vs mère (n=235) 0,35 Ado vs père (n=189) 0,39 Mère vs père (n=128) 0,57 Baseline vs suivi Info ado (n=60) 0,68 Info parent (n=89) 0,23						
53	K-SADS Shahrivar 2010 ☞+ Version perse						9-18 ans (n=102) 📅 même jr Info parent-ado K-SADS vs ÉC (BED ¹⁰⁶) K Dx actuel 0,74	
54	K-SADS Shanee 1997 ☞☞+ Version	6-19 ans (n=57) 📅14 jrs K ☞ 2 0,84 Mère-ado 0,47					6-19 ans (n=57) 📅 après 2 ms de Tx Info parent-ado K-SADS vs ÉC	

¹⁰⁶ Best estimate diagnostic (meilleur estimation du Dx)
























	hébreux					K 0,89		
55	K-SADS Weissman 1987 ☞☞☞ mais avec parent avec Dx vs sans Dx	6-23 ans (n=175) ☞ même jr ☞ 1 K Mère vs ado 0,30 13-18 ans (n=70) Parent EDM 0,38 Parent ≠ EDM 0,49						
56	PADDI Abrantes 2004 ☞☞☞			13-18 ans (n=284) α (8 items) 0,838				
57	PADDI Hoffman 2004 ☞☞☞ plus certains ☞☞☞			12-18 ans (n=279) α (9 items) 0,923				
	PAI-A							

Annexe E : Tableau 4, Trouble des conduites (TC)

Voir la légende des Tableaux à l'Annexe A.

		Trouble des conduites (TC)						
		Fidélité			Validité			
Q-2	instrument étude échantillon	Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
	ADI							
2	APS Lemos 2011  Version portugaise		12-19 ans (n=68)  12 sem.  ? Info ado r 0,42	12-18 ans Portugal (n=656) α 0,73 ÉU (n=1827) 0,80				
3	CAPA Angold 1995 		10-18 ans (n=70)  moyen 5 jrs  2 Info ado K 0,55					
4	CAPA Angold 2012 				9-16 ans  7 jrs Info parent-ado CAPA vs DAWBA ¹⁰⁷ (n=215) 0,39 CAPA vs DISC (n=208) 0,41			
6	ChIPS				13-18 ans (n=19)	6-18 ans (n=47)		

¹⁰⁷ *Development and Well-Being Assessment* (Goodman et al., 2000)










	Fristad 1998a  				 moins de 7 jrs Info ado ChIPS vs DICA K % 0,791 74	 ? Info ado ChIPS vs EC K % 0,421 70		
7	ChIPS Fristad 1998b  dont n=18 endeuillés				6-18 ans (n=40)  même jr Info ado ChIPS vs DICA K % 1,00 100	6-18 ans (n=40)  (sur dossier) Info ado ChIPS vs EC K % 1,00 100		
8	ChIPS Moreno 2006  vs  Version espagnole					6-18 ans  (n=31)  30 jrs Info parent ChIPS vs EC K % 1,00 100		6-18 ans  (n=43)  7 jrs  (n=31)  30 jrs Info parent Concluant vs nbre de Dx (6,9%/16% vs niveau de détérioration (famille-école- amis) (2,3%/32.3%)
9	ChIPS Paing 2010  		Âge moyen 12.2 (DS 3.9) n=12  ?  ? Info parent Tél. vs face à face % 83,3					
10	ChIPS Rathman 2008  				11-14 ans (n=10)  7 jrs Info parent-ado ChIPS vs K-SADS K % - 90	11-14 ans (n=10)  (sur dossier) Info parent-ado ChIPS vs EC K % - 90		


















11	ChIPS Swenson 2007 ☞+				12-18 ans (n=93) 📅 3 jrs Info ado ChIPS vs K-SADS K % 0,44 73			
12	ChIPS Witwer 2012 👨👩+	6-17 ans (n=43) 📅 (sur enregistrement) 👤 2 Info ado K % 0,77 88,4		6-17 ans (n=61) K-R 0,65	6-17 ans (n=61) 📅 même jr Info parent-ado ChIPS vs CASI ¹⁰⁸ K % 0,57 78,7			
13	DICA Boyle 1993 ✍️ mais à forte probabilité de Dx	12-16 ans (n=137) 📅 10 à 20 jrs 👤 4 (dont des spécialistes ¹⁰⁹) Accord parent- ado K % Temps 1 -0,02 95 Temps 2 -0,02 95	12-16 ans (n=137) 📅 10 à 20 jrs 👤 2 K % Info parent 0,87 99 Info ado 0,92 99					
14	DICA Boyle 1997 ✍️ mais à forte probabilité de Dx		6-16 ans (n=210) 📅 moyen 17 jrs 👤 ? Info parent K r 0,21 0,67				6-16 ans (n=223) 📅 ? DICA vs OCHS-R ¹¹⁰ K 0,45 Info parent (n=213) Concluant vs fonctionnement	

¹⁰⁸ *Child and Adolescent Symptom Inventory* (Gadow et Sprafkin, 2002)

¹⁰⁹ Pédiopsychiatre et psychologue pour enfant

¹¹⁰ *Revised Ontario Child Health Study scales* (Boyle, Offord, Racine, Fleming, et al., 1993)

							social détérioré et pauvres performances scolaires Info profs (n=184) Concluant vs fonctionnement social détérioré Info parent-ado (n=214) Concluant vs problèmes disciplinaires	
16	DICA Ezpeleta 1997  + Version espagnole					13-17 ans (n=66)  ? DICA vs EC K % Info ado 0,79 98 Info parent 0,57 – Combinée 0,30 80		
17	DICA Granero 1998  + Version espagnole		13-17 ans (n=52)  moyen 11 jrs ↓ 2 Info ado K 0,589					
19	DICA Reich 1995  + Teste version informatique		13-18 ans (n=?)  approximatif 5 jrs ↓ 2 Info ado K 0,17					
21	DICA Welner	7-17 ans  (n=27)				7-17 ans  (n=27)  ?		

	1987  	 1 à 7 jrs  2 Info ado K % 1,00 100  (n=84)  même jr Info parent vs ado K 0,80				Info parent-ado DICA vs EC lors du congé K % 0,43 78		
4	DISC Angold 2012  				9-16 ans  7 jrs Info parent-ado K Vs DAWBA ¹¹¹ (n=213) 0,26 vs CAPA (n=208) 0,41			
22	DISC Aronen 1993  					12-15 ans (n=163)  moyen 135,56 jrs ¹¹² Info ado Disc vs EC au congé K 0,06		
23	DISC Bravo 2001   		11-17 ans (n=85)  moyen 12jrs  4 K Info parent Dernière an 0,42 À vie 0,48					

¹¹¹ *Development and Well-Being Assessment* (Goodman et al., 2000)

¹¹² Durée moyenne de l'hospitalisation

	Version espagnole		Info ado (n=83) Dernière an 0,62 À vie 0,66					
24	DISC Breton 1998 version française		12-14 ans (n= 145) 7 et 21 jrs 4 Info parent K Info ado 0,49	12-14 ans (n=148) à Info parent 0,62 Info ado (n=145) Un peu plus de 0,60				
25	DISC Crowley 2001 Certains vs						Âge moyen 16,1 ? (n = variable de 149 à 172) Info parent-ado Concluant vs CBCL ¹¹³ cpt délinq vs YSR ¹¹⁴ cpt délinq vs CD ¹¹⁵ Sx agressivité (DISC) vs PABR ¹¹⁶	Âge moyen 16,1 ? Info parent-ado Concluant vs Prévalence/ nbre Sx moyen !! (n= 87) 80,2 %/5,3 (n=85) 3,5%/0,8 vs YSR <i>Delinquent Behavior T score</i> !! 67,4 (95 th percentile) vs CBCL <i>Delinquent Info parent</i> !! :73,8 (98 th percentile) : normal
29	DISC, Friman							9-17 ans 12 mois ¹¹⁷

¹¹³ *Child Behavior Checklist* (Achenbach et Edelbrock, 1983)

¹¹⁴ *Youth Self-Report* (Achenbach, 1991b)

¹¹⁵ *Conduct disorder* : trouble de la conduite

¹¹⁶ *Peak Aggressive Behavior rating (life)* (version modifiée de Lewis, Pincus, Shanok et Glaser, 1982)





















	2000 ↔● avec Dx TC vs ↔● sans Dx TC (groupe contrôle) Version informatique							Dx n=146 ↔Dx n=223 Concluant vs DIR ¹¹⁸
31	DISC Jensen 1995 ↔↔+ ↔+ vs ↔		9-17 ans ↔↔📅moyen 15jrs. ↔📅moyen 21 jrs 👤 4 K Info parent ↔↔ (n=88) 0,70 ↔ (n=273) 0,66 Info ado ↔↔ (n=90) 0,86 ↔ (n=270) 0,60 Info combinée ↔↔ (n=91) 0,71 ↔ (n= 264) 0,64					
32	DISC Jewell 2004 ↔#						9-18 ans (n=534) 📅moyen 164 jrs ¹¹⁹ Info ado Concluant vs le DIR ¹²⁰	
36	DISC Piacentini				11-17 ans 📅 même jr			

¹¹⁷ Observation tout au long du Tx de un an.

¹¹⁸ *Daily Incident Report* (Larzelere, 1996)

¹¹⁹ Évaluation en communauté avant l'arrivée en Tx

¹²⁰ *Daily Incident Report* (Larzelere, 1996)











	1993  +				DISC Vs CAF ¹²¹ k Info parent (n=73) 0,60 Info ado (n=72) 0,21 Combinée (n=71) 0,46			
38	DISC Roberts 1996  +  • Minorité ethnique		12-17 ans (n=101)  moyen 7 à 14 jrs  2 Info ado K Y 0,564 0,632					
39	DISC Rubio-Stipeč 1994  + vs  Version espagnole	9-17 ans  moins de 15 jrs  2 Info parent vs ado K  (n=248) 0,01  (n=28) 0,24				9-17 ans (n=170 :  124 et  46)  (sur dossier) DISC vs BED ¹²² K Info parent 0,15 Info ado 0,15 Combinée 0,34		
40	DISC Schwab-Stone 1993 		11-17 ans (n=39)  7 à 21 jrs  4 Info parent K 0,87 Info ado 0,55 Combinée 0,50	11-17 ans (n=39)  7 à 21 jrs Info parent α 0,56 Info ado 0,59				
41	DISC Schwab-Stone		9-18 ans (n=247)  1 à 15 jrs			9-18 ans (n=247)  1 à 15jrs		

¹²¹ *Clinical Assessment Form* (Piacentini et al., 1993)

¹²² *Best estimate diagnostic* (meilleur estimation du Dx)

	1996 👤👤👤 avec Dx Vs 👤👤 sans Dx		👤 4 K Info parent 0,56 Info ado 0,64 Combinée 0,66			DISC vs EC K Info parent 0,74 Info ado 0,77 Combinée 0,80		
42	DISC Weinstein 1989 👤👤+					12-15 ans (n=163) 👤 max. 15 jrs Info ado DISC vs EC à l'admission K 0,09		
	GAIN							
	ISCA							
45	K-SADS Brasil 2010 👤👤+ Version brésilienne						6-14 ans (n=78) 👤 15 jrs Info parent-ado r 0,61 Concluant vs CBCL/4-18 externalisés ¹²³	
46	K-SADS Chambers 1985 👤👤+		6-17 ans (n=52) 👤 moins de 3 jrs 👤 4 Info combinée parent/ado/ évaluateur K 0,63	6-17 ans (n=52) <i>Conduct disorder scale</i> α 0,86				
47	K-SADS Ghanizadeh 2006 👤👤+	4-19 ans (n=14) 👤? 👤 2 Info parent-ado				4-19 ans (n=109) 👤? Info parent-ado K-SADS vs EC		




¹²³ Child Behavior Checklist (Achenbach, 1991a)

	plus  version perse	K 0,81				K 0,923		
48	K-SADS Hodges 1987  +				7-16 ans K-SADS vs CAS ¹²⁴  approximatif 32 hres K Info parent (n=30) 0,60 Info ado (n=29) 0,50 Info ado ou parent 0,63 Combinée 0,59			
49	K-SADS Kaufman 1997  + plus 		7-17 ans (n=20)  moyen 17,9 jrs  4 Info parent-ado K Dx actuel – Dx vie 0,83				7-17 ans (n=66)  ? Info parent-ado Concluant vs CBCL ¹²⁵ externalisé BEH/POS = 68.7 ± 8.6 BEH/NEG = 53.6 ± 10.7 $t_{63} = 6.29,$ $P < .0001$	
50	K-SADS Kolaitis 2003 ? + Version grecque	7-15 ans (n=26)  ?  2 Info parent-ado K 0,90						

¹²⁴ *Child Assessment Schedule* (Hodges et al., 1982)











¹²⁵ *Child Behavior Checklist* (Achenbach et Edelbrock, 1983)

10	K-SADS Rathman 2008 ↔+					11-14 ans (n=10) 📅 (sur dossier) Info parent-ado K-SADS vs ÉC K % - 100		
52	K-SADS Rothen 2009 👤👤 mais avec parent avec Dx Version française	7-17 ans 📅? 👤 2 Y Ado vs mère (n=235) 0,51 Baseline vs suivi Info ado (n=60) 0,68						
53	K-SADS Shahrivar 2010 ↔+ Version perse					9-18 ans (n=102) 📅 même jr Info parent-ado K-SADS vs ÉC K Dx actuel 0,46 Dx vie 0,52		
54	K-SADS Shanee 1997 👤↔+ Version hébreux	6-19 ans (n=57) 📅 14 jrs K 👤 2 0,92 Mère-ado 0,65				6-19 ans (n=57) 📅 après 2 ms de Tx Info parent-ado K-SADS vs ÉC K 0,86		
55	K-SADS Weissman 1987, 👤👤 mais avec parent avec Dx vs sans Dx	6-23 ans (n=175) 📅 même jr 👤 1 K Info mère vs ado 0,36 13-18 ans (n=70)						


















		Parent EDM 0,13 Parent ≠ EDM 0,49						
56	PADDI Abrantes 2004 			13-18 ans (n=284) α 0,624 (12 items)				
57	PADDI Hoffman 2004  certains 			12-18 ans (n=279) α 0,711 (11items)				
	PAI-A							

Annexe F : Tableau 5, Trouble déficitaires de l'attention avec ou sans hyperactivité (TDA/H)

Voir la légende des Tableaux à l'Annexe A.

Trouble déficitaire de l'attention avec ou sans hyperactivité (TDA/H)								
Q-2	Instrument Études échantillon	Fidélité			Validité			
		Interjeu	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
	ADI							
	APS							
4	CAPA Angold 2012 				9-16 ans  7jrs Info parent-ado K CAPA vs DAWBA ¹²⁶ (n=215) 0,49 CAPA vs DISC (n=208) 0,52			
6	ChIPS Fristad 1998a  				13-18 ans (n=19)  moins de 7 jrs Info ado ChIPS vs DICA K % 0,568 84	6-18 ans (n=47)  ? Info ado ChIPS vs EC K % 0,363 74		
7	ChIPS Fristad 1998b  dont n=18 endeuillés				6-18 ans (n=40)  même jr Info ado ChIPS vs DICA K % 0,494 97.5	6-18 ans (n=40)  (sur dossier) Info ado ChIPS vs EC K % 1,00 100		
8	ChIPS					6-18 ans 		6-18 ans

¹²⁶ *Development and Well-Being Assessment* (Goodman et al., 2000)








	Moreno 2006  vs  Version espagnole					(n=31)  30 jrs Info parent-ado ChIPS vs EC K % 0,74 87		 (n=43) ≡ 7jrs  (n=31) ≡ 30 jrs Concluant vs nbre de Dx (39,5% vs 64,5%) vs niveau de détérioration (2,3% vs 32,3%)
9	ChIPS Paing 2010  +		Âge moyen 12.2 n=12  ?  ? Info parent Tél. vs face à face % 100					
10	ChIPS Rathman 2008  +				11-14 ans (n=10)  7 jrs Info parent-ado ChIPS vs K-SADS K % 0,35 70	11-14 ans (n=10)  (sur dossier) Info parent-ado ChIPS vs EC K % 0,52 80		
11	ChIPS Swenson 2007  +				12-18 ans (n=93)  3 jrs Info ado ChIPS vs K-SADS K % 0,26 63			
12	ChIPS Witwer 2012  + ‡	6-17 ans (n=43)  (sur enregistrement)  2 Info ado		6-17 ans (n=43) K-R TDA-H 0,80 TDA-I 0,84 TDA-C 0,89	6-17 ans (n=61)  même jr Info parent-ado ChIPS vs CASI ¹²⁷ K % tout TDAH			

¹²⁷ Child and Adolescent(e) Symptom Inventory(Gadow et Sprafkin, 2002)













		<p>K %</p> <p>Tout TDAH 0,23 66,6</p> <p>TDA-H 0,43 96,7</p> <p>TDA-I 0,60 88,4</p> <p>TDA-combiné 0,27 62,8</p>			<p>0,40 80,3</p> <p>TDA-H 0,47 91,8</p> <p>TDA-I 0,30 75,4</p> <p>TDA-combiné 0,35 65,5</p>			
13	<p>DICA Boyle 1993 ✍ mais à forte probabilité de Dx</p>	<p>12-16 ans (n=137) 📅 10 à 20 jrs 👤 4 (dont des spécialistes¹²⁸) Info parent vs ado Temps 1 K % -0,02 93 Temps 2 -0,01 94</p> <p>6-16 ans (n=60) K 👤 2 Info-parent 0,75 Info-ado 0,34 Pédo vs non- clinicien Info-parent 0,71 Info-ado 0,78</p>	<p>12-16 ans (n=137) 📅 10 à 20 jrs 👤 2 K % Info parent 0,86 99 Info ado 0,24 99</p>					
14	<p>DICA Boyle 1997 ✍ mais à</p>		<p>6-16 ans (n=210) 📅 moyen 17 jrs 👤 2 Info parent K r</p>				<p>6-16 ans (n=223) 📅 ? DICA vs OCHS-R¹²⁹ K 0,36</p>	<p>6-16 ans (n=210) 📅 ? Non concluant vs problèmes développemen-</p>

¹²⁸ Pédopsychiatre et psychologue pour enfant














¹²⁹ Revised Ontario Child Health Study scales (Boyle, Offord, Racine, Fleming, et al., 1993)

	forte probabilité de Dx		0,59 0,73				Info parent (n=213) Concluant vs fonctionnement social détérioré et pauvres performances scolaires Info profs (n=184) Concluant vs fonctionnement social détérioré Info parent-ado (n=214) Concluant vs problèmes disciplinaires	taux
15	DICA Dela Osa 1997  Vs  Version espagnole						6-17 ans (n=147)  ? Info parent 0,34 Info ado 0,42 Modérément concluant vs CBCL ¹³⁰	6-17 ans  ? Info parent-ado combinée Concluant vs  (n=130) -Dx/-Sx % des Dx bien identifiés 65,3/57,1/58,2  (n=196) +Dx/+Sx % Dx bien identifiés 81,3/84,9/72,7 p < .0005
16	DICA Ezpeleta					13-17 ans (n=65)  ?		

¹³⁰ Child Behavior Checklist(Achenbach, 1991b)
























	1997  + version espagnole					DICS vs EC K % Info ado 0,40 92 Info parent-ado 0,46 91 Combinée 0,41 89		
17	DICA Granero 1998  + Version espagnole		13-17 ans (n=52)  moyen 11 jrs ↓ 2 Info ado K 0,510					
18	DICA Hope 1999  + plus 	11-17 ans (n=24)  même jr ↓ 2 Info parent vs ado (n=121) K % 0,30 39.4						
19	DICA Reich 1995  + Teste version informatique		13-18 ans (n=?)  approximatif 5jrs ↓ 2 Info ado K 0,29					
20	DICA Sylvester 1987  parent avec Dx (trouble anxieux ou	7-17 ans (n=74)  même jr ↓ 2 (spécialistes) K % Info parent avec Dx vs  Dx 0,61 84					7-17 ans (n=82)  moins de 15 jrs Info parent-ado DICA vs PIC ¹³¹ K % 0,1106 76	

































¹³¹ *Personality Inventory for Children* (Lacher et Gdowski, 1979; Wirt et al., 1977)












	dépression) vs  parent sans Dx	Info parent vs ado 0,42 76					
21	DICA Welner 1987   	7-17 ans ○ (n=27)  1 à 7 jrs ↓ 2 Info ado K % 1,00 100  (n=84)  même jour Info parent vs ado K 0,66				7-17 ans (n=27)  ? Info parent-ado DICA vs EC lors du congé K % 0,50 78	
4	DISC Angold 2012 				9-16 ans  7jrs Info parent-ado K DISC vs DAWBA ¹³² (n=213) 0,57 DISC vs CAPA (n=208) 0,52		
22	DISC Aronen 1993  					12-15 ans (n=163)  moyen 135,56 jrs ¹³³ Info ado DISC vs EC au congé K 0,215	
23	DISC		11-17 ans (n=85)				

¹³² *Development and Well-Being Assessment* (Goodman et al., 2000)

¹³³ Durée moyenne de l'hospitalisation

	Bravo 2001   Version espagnole		 moyen 12jrs  4 K Info parent Dernière an 0,37				
24	DISC Breton 1998  Version française		6-14 ans n= 260  entre 7 et 21 jrs  4 K Info parents 0,60 Info ado 12-14 ans (n=145) -	12-14 ans α Info parents (n=148) 0,90+ Info ado (n=145) Entre 0,80 et 0,90			
25	DISC Crowley 2001  Certains  vs 	Âge moyen 16,1  ?  ? Info parent vs ado K  (n=69)-0,079  (n=84) -0,016					Âge moyen 16,1  ? Info parent-ado % prévalence Non concluant vs info ado  (n=69) 4,1  (n=84) 2,4 Concluant vs info parent  (n=69) 23,3  (n=84) 1,2
27	DISC Edelbrock 1999 		6-15 ans (n=24)  moyen 7,5 jrs  1 Info parent K 0,91				
28	DISC Flisher 2012		Âge moyen 15 (n=105)  approximatif 15 jrs				

	 + plus 		 1 K Info parent 0,559 Info ado 0,227					
30	DISC, Ho 2005  + Version chinoise		9-18 ans (n=79)  moyen 22.7 jrs  4 K Info parent 0,81 Info ado 0,25 Combinée 0,75					
31	DISC, Jensen 1995   +   vs 		9-17 ans    moyen 15jrs.   moyen 21 jrs  4 K Info parent   (n=88) 0,69  (n=273) 0,57 Info ado   (n=90) 0,59  (n=270) 0,43 Combinée   (n=91) 0,68  (n= 264) 0,62					
34	DISC, Lewczyk 2003  (? ) +					6-18 ans (n=240)  moyen 168 jrs DICS vs ÉC avant étude K 0,22		
35	DISC McGrath 2000					12-17 ans (n=58)  ? DISC vs ÉC avant	12-17 ans (n=58)  7 jrs Corrélation Dx	12-17 ans (n=58)  7 jrs Concluant vs info tuteur

	Ado en famille d'accueil					étude	Concluant vs Brown ¹³⁵ ado	Non concluant vs info ado
						Info tuteur ¹³⁴ 0,43 Info ado 0,31	CPRS ¹³⁶ tuteur CTRS ¹³⁷ tuteur Corrélation Sx Concluant vs Brown ado CPRS tuteur CTRS tuteur	
36	DISC Piacentini 1993  +				11-17 ans  même jr DISC Vs CAF ¹³⁸	Info parent 0,53 Info ado 0,48 Combinée 0,46		
38	DISC Roberts 1996  + Minorité ethnique		12-17 ans (n=101)  moyen 7 à 14 jrs ↓ 2 Info ado K 0,470 Y 0,599					
39	DISC Rubio-Stipec, 1994  + vs  Version	9-17 ans  moins de 15 jrs ↓ 2 Info parent vs ado K  n= 248 0,07				9-17 ans (n=170 =  124 et  46)  (sur dossier) DISC vs BED ¹³⁹ K		

¹³⁴ Étant donné que les adolescent(e)s sont placés, il s'agit d'adulte significatif de la famille d'accueil.



















¹³⁵ *Brown Attention Deficit Disorder Scale for Adolescent* (Brown, 1996)

¹³⁶ *Conners' Parent Rating Scale* (Conners, 1989a)

¹³⁷ *Conners' Teacher Rating Scale* (Conners, 1989b)








¹³⁸ *Clinical Assessment Form* (Piacentini et al., 1993)

¹³⁹ Best estimate diagnostic (meilleure estimation du Dx)













	espagnole	  n=28 0,21				Info parent 0,42 Info ado 0,36 Combinée 0,43		
40	DISC Schwab-Stone 1993  		11-17 ans (n=39)  7 à 21 jrs  4 K Info parent 0,55 Info ado – Combinée 0,56	11-17 ans (n=39) à info parent 0,87 info ado 0,83				
41	DISC Schwab-Stone 1996  avec Dx Vs  sans Dx		9-18 ans (n=247)  1 à 15 jrs  4 K Info parent 0,60 Info ado 0,10 Combinée 0,48			9-18 ans (n=247)  1 à 15 jrs DISC vs ÉC Info parent 0,72 Info ado 0,27 Combinée 0,70		
42	DISC Weinstein 1989  					12-15 ans (n=163)  max. 15 jrs Info ado DISC vs ÉC à l'admission K 0,09		
	GAIN							
	ISCA							
44	K-SADS Ambrosini 1989  	6-17 ans (n=25)  (enregistrement vidéo)  2 Info parent-ado K		6-17 ans (n=25) à Inattention SPE ¹⁴⁰ Évaluateur 1 0,82 Évaluateur 2 0,75 SLW ¹⁴¹				

¹⁴⁰ Score Sx present illness 12 months (présence de Sx au cours des 12 derniers mois)

¹⁴¹ Score Sx present week (présence de Sx au cours de la présente semaine)

		0,88		Évaluateur 1 0,86 Évaluateur 2 0,84 Impulsivité SPE Évaluateur 1 0,82 Évaluateur 2 0,79 SLW Évaluateur 1 0,86 Évaluateur 2 0,84 Hyperactivité SPE Évaluateur 1 0,82 Évaluateur 2 0,83 SLW Évaluateur 1 0,83 Évaluateur 2 0,82				
47	K-SADS Ghanizadeh 2006  Plus  Version perse	4-19 ans (n=14)  ? † 2 Info parent-ado K 0,69	4-19 ans (n=11)  28 jrs † 2 Info ado K 0,81			4-19 ans (n=109)  ? K-SADS vs EC Info parent-ado K 0,905		
48	K-SADS Hodges 1987 				7-16 ans (parent : n=30 ; ado n=29)  approximatif 32 hres K-SADS vs CAS ¹⁴² K Info ado 0,36 + info scolaire 0,64 Info parent 0,60 + info scolaire 0,72 Ado ou parent 0,39 + info scolaire 0,70 Combinée 0,65			

¹⁴² *Child Assessment Schedule* (Hodges et al., 1982)











49	K-SADS Kaufman 1997  plus 		7-17 ans (n=20)  moyen 17,9 jrs  2 Info parent-ado K Dx actuel 0,63 Dx vie 0,55				7-17 ans (n=66)  ? Info parent Concluant vs Conners ¹⁴³ ADHD/POS = 25.8 ± 13.8; ADHD/NEG = 13.1 ± 15.7; $t_{63} = 3.43, P < .001$	
10	K-SADS Rathman 2008 					11-14 ans (n=10)  (sur dossier) Info parent-ado K-SADS vs EC K % 0,78 90		
52	K-SADS Rothen 2009  mais avec parent avec Dx	7-17 ans  ?  ? Y Ado vs mère (n=235) 0,32 Ado vs père (n=189) 0,37 Mère vs père (n=128) 0,63 Baseline vs suivi Info ado (n=60) 0,57 Info parent (n=89) 0,48						
53	K-SADS Shahrivar		9-18 ans (n=32)  7 jrs			9-18 ans (n=102)  même jr		















¹⁴³ *Conners Abbreviated Questionnaire/Parent version*(Conners et Barkle, 1985)














	2010 ↔+ Version perse		♂ 2 Info ado K % Dx actuel 0,87 93 Dx vie 0,75 87			Info parent-ado K-SADS vs ÉC K Dx actuel 0,80 Dx vie 0,74		
54	K-SADS Shanee 1997 ↔+ Version hébreux	6-19 ans (n=57) 📅 14 jrs K ♂ 2 1,00 Mère-ado 0,94				6-19 ans (n=57) 📅 après 2 ms de Tx Info parent-ado K-SADS vs ÉC K 0,94		
55	K-SADS Weissman 1987, 👤 mais avec parent avec Dx vs sans Dx	6-23 ans (n=175) 📅 même jr ♂ 1 K Info mère vs ado 0,34						
	PADDI							
	PAI-A							









Annexe G : Tableau 6, État de stress post-traumatique (ESPT)

Voir la légende des Tableaux à l'Annexe A.

État de stress post-traumatique (ESPT)								
Q-2	Instrument Étude échantillon	Fidélité			Validité			
		Interjuge	Test-retest	Cohérence interne	Concordante	Concurrente	Convergente	Discriminante
	ADI							
2	APS Lemos 2011  Version portugaise		12-19 ans (n=68)  12 sem.  ? Info ado r 0,57	12-19 ans Portugal (n=656) ÉU (n=1827) α 0,81 0,83				
5	CAPA Costello 1998 Liste d'attente  Vs  d'une étude épidémiologi- que		8-17 ans ($\Delta n=58$)  15 jrs  2 K Info parent 0,54 Info ado 0,64					8-17 ans  n=58/  n=1015  Info à même le CAPA Info parent ou ado Concluant vs % 1+ évèn. imp. vie 77.3/43.1 1+ évèn faible 3ms 56.8/35,2 Souvenir douloureux 51,4/17,2 <i>Hyperarousal</i> 18,9/1,8 Évitement 21,6/7,4

								Dx 9,1/0,9
6	ChIPS Fristad 1998a  +				13-18 ans (n=19)  moins de 7 jrs Info ado ChIPS vs DICA K % 0,672 84	6-18 ans (n=47)  ? Info ado ChIPS vs EC K % 0,618 87		
7	ChIPS Fritad 1998b  dont n=18 endeuillés				6-18 ans (n=40)  même jrs Info ado ChIPS vs DICA K % - 100	6-18 ans (n=40)  (sur dossier) Info ado ChIPS vs EC K % - 100		
8	ChIPS Moreno 2006  vs  Version espagnole					6-18 ans (n=31)  30 jrs Info parent ChIPS vs EC K % 1,00 100		
9	ChIPS Paing 2010  +		Âge moyen 12.2 (n=12) ≡?  ? Info parent Tél. vs face à face % 100					
11	ChIPS Swenson 2007  +				12-18 ans (n=93)  3 jrs Info ado ChIPS vs K-SADS K % 0,42 80			
16	DICA Ezpeleta					13-17 ans (n=63)  ?		

	1997  Version espagnole					DICA vs ÉC K % Info ado 0,55 95 Info parent 0,66 98 Combinée 0,55 95		
17	DICA Granero 1998  Version espagnole		13-17 ans (n=52)  moyen 11 jrs  2 Info ado K 0,217					
19	DICA Reich 1995  Teste version informatique		13-18 ans (n=?)  approximatif 5 jrs  2 Info ado K 0,63					
	DISC							
	GAIN							
	ISCA							
47	K-SADS Ghanizadeh 2006  plus \cap Version Perse					4-19 ans (n=109)  ? Info parent-ado K-SADS vs ÉC K 0,879		
49	K-SADS Kaufman 1997  vs 		7-17 ans (n=20)  moyen 17,9 jrs  2 Info parent-ado K Dx actuel 0,67					

			Dx vie 0,60					
53	K-SADS Shahrivar 2010  Version perse					9-18 ans (n=102)  même jr Info parent-ado K-SADS vs EC K Dx actuel 0,17 Dx vie 0,20		
54	K-SADS Shanee 1997  Version hébreux	6-19 ans (n=57)  14 jrs K ‡ 2 1,00 Mère-ado 1,00				6-19 ans (n=57)  après 2 ms de Tx Info parent-ado K-SADS vs EC K 1,00		
56	PADDI Abrantes 2004 			13-18 ans (n=284) à 0,899 (7 items)				
57	PADDI Hoffman 2004  certains 			12-18 ans (n=279) à 0,872 (7items)				
	PAI-A							

Annexe H : Retour sommaire dans les bases pour les années 2013 à 2016 en utilisant le titre de l'instrument seulement

	Abréviation	Total occurrences des trois plateformes	Retenus sur titre seulement	
1	ADI	Ovid :	2	0
		Web of science :	1	0
		Proquest :	8	0
		Total :	11	0
2	APS	Ovid :	3	0
		Web of Science :	3	0
		Proquest :	9	2
		Total :	15	2
3	CAPA	Ovid :	10	0
		Web of Science :	13	1
		Proquest :	30	0
		Total :	53	1
4	ChIPS	Ovid :	7	0
		Web of Science :	1	0
		Proquest :	3	0
		Total:	11	0
5	DICA	Ovid :	8	0
		Web of Science :	5	0
		Proquest :	17	0
		Total :	30	0
6	DISC	Ovid :	10	0
		Web of Science :	64	2

		Proquest :	240 ¹⁴⁴	0
		Total :	314	2
7	GAIN	Ovid :	10	0
		Web of Science :	11	2
		Proquest :	40	0
		Total :	61	2
8	ISCA	Ovid :	3	0
		Web of Science :	2	0
		Proquest :	3	0
		Total :	52	0
9	K-SADS	Ovid :	2	0
		Web of Science :	155	6
		Proquest :	200 ¹⁴⁴	0
		Total :	357	6
10	PADDI	Ovid :	0	0
		Web of Science :	2	0
		Proquest :	0	0
		Total :	9	0
11	PAI-A	Ovid :	2	0
		Web of Science :	2	0
		Proquest :	13	0
		Total :	17	0
		Total :	930	13

¹⁴⁴ Les 50 premiers furent regardés.