

Dominic Beaulieu-Prévost

**ANALYSE DE VALIDITÉ DE LA DÉCLARATION (SVA), MENSONGE ET
FAUX SOUVENIRS: VALIDITÉ ET EFFICACITÉ CHEZ LES ADULTES**

Mémoire
présenté
à la Faculté des études supérieures
de l'Université Laval
pour l'obtention
du grade de maître en psychologie (M.Ps.)

École de psychologie
FACULTÉ DES SCIENCES SOCIALES
UNIVERSITÉ LAVAL

MARS 2001

RÉSUMÉ

Analyse de validité de la déclaration (SVA), mensonges et faux souvenirs: Validité et efficacité chez les adultes

Dominic Beaulieu-Prévost

Le SVA est utilisé pour évaluer la crédibilité d'un témoignage, surtout dans des cas d'abus sexuels. Les buts de cette étude étaient d'en évaluer l'efficacité et de vérifier si les effets observés étaient dus à l'intention de tromper ou à la source du souvenir (interne vs externe). 34 participant(e)s racontèrent 2 événements ou 2 rêves (1 réel et 1 inventé) après 2 jours ou plus de préparation. Seize des 19 critères du SVA furent évalués. Les récits réels contenaient plus d'incohérences et moins de détails inusités que les récits inventés. Aucun autre critère du SVA ne fut efficace. Les effets étaient causés exclusivement par l'intention de tromper. Les résultats indiquent qu'en contexte de variabilité et préparation réaliste, le SVA perd son efficacité.

Candidat:

Directeur de recherche:

Codirecteur de recherche:

AVANT-PROPOS

J'aimerais remercier un certain nombre de personnes qui ont d'une façon ou d'une autre rendu possible l'aboutissement de ce projet:

Geneviève, pour son support moral et technique, pour sa patience, et pour les compromis qu'elle a acceptés de faire durant les dernières années.

Jean-Roch Laurence, pour ses contributions académiques, commentaires, suggestions et encouragements tout au long du projet.

Robert Rousseau, pour son aide précieuse, ses commentaires, suggestions et corrections, et ce malgré les difficultés engendrées par la distance.

Iliana Lilova, pour la transcription et l'évaluation d'une partie des entrevues.

Antonio Zadra pour ses commentaires et son aide technique, particulièrement lors du premier projet.

Mes ami(e)s, ma famille et ma belle-famille, pour m'avoir supporté et ne pas m'avoir oublié malgré mon isolement de fin de rédaction.

Les gens du laboratoire d'hypnose de Concordia, pour leur support et leurs encouragements.

Mathieu Pilon, pour son aide technique occasionnelle durant le dernier mois.

Les trente-six participants, pour avoir contribué au projet.

TABLE DES MATIÈRES

	<u>Page</u>
RÉSUMÉ.....	ii
AVANT-PROPOS.....	iii
TABLE DES MATIÈRES.....	iv
LISTE DES TABLEAUX.....	vi
INTRODUCTION.....	1
Historique du SVA.....	3
Recherches utilisant des cas réels (quasi-expérimentales).....	6
Critiques sur les recherches quasi-expérimentales.....	13
Recherches analogues (expérimentales).....	15
Critiques sur les recherches expérimentales.....	25
Analyses critiques et revues de littérature.....	27
Vue d'ensemble sur la recherche.....	29
Présente étude.....	32
MÉTHODOLOGIE.....	36
Plan d'expérience.....	36
Instructions concernant les récits.....	36
Participants / Recrutement.....	37
Procédure.....	37
Entrevue téléphonique.....	37
Entrevue.....	38
Transcription et entrée de données.....	40
Formation des évaluateurs.....	40
Opérationnalisation des critères.....	41
RÉSULTATS.....	49
Accord inter-juges.....	49
Analyse de variance multivariée.....	50
Établissement de normes.....	54

DISCUSSION.....	55
Interprétation des résultats.....	55
Comparaison à la littérature.....	57
Hypothèses alternatives et limites.....	60
CONCLUSION.....	64
BIBLIOGRAPHIE.....	66
ANNEXE A - Description des éléments contenus dans la Validity Checklist.....	72
ANNEXE B - Définitions des critères du CBCA tels qu'utilisés dans l'étude.....	78
ANNEXE C - Formulaire de consentement (versions française et anglaise).....	87
ANNEXE D - Protocoles (anglais et français).....	92
ANNEXE E - Lettre envoyée aux participants (versions française et anglaise).....	103
ANNEXE F - Statistiques descriptives.....	110

LISTE DES TABLEAUX

	<u>Page</u>
Tableau 1: Brèves description des critères du Criteria-Based Content Analysis...	7
Tableau 2: Accords inter-juges des différents critères évalués mesurés par des corrélations de Pearson.....	51
Tableau 3: Moyennes (M) et écarts-types (ÉT) des événements réels (ÉR), événements inventés (ÉI), rêves réels (RR) et rêves inventés (RI) en fonction de chaque critère.....	53
Tableau F-1: Moyennes (M) et écarts-types (ÉT) des événements et des rêves en fonction de chaque critère.....	111
Tableau F-2: Moyennes (M) et écarts-types (ÉT) des déclarations réelles et des déclarations inventées en fonction de chaque critère.....	112
Tableau F-3: Moyennes (M) et écarts-types (ÉT) des événements réels (ÉR), événements inventés (ÉI), rêves réels (RR) et rêves inventés (RI) chez les francophones en fonction de chaque critère.....	113
Tableau F-4: Moyennes (M) et écarts-types (ÉT) chez les anglophones des événements réels (ÉR), événements inventés (ÉI), rêves réels (RR) et rêves inventés (RI) en fonction de chaque critère.....	114

Analyse de validité de la déclaration, mensonges et faux souvenirs: Validité et efficacité
chez les adultes

Dominic Beaulieu-Prévost

Du point de vue juridique, lors d'un procès ou d'une déposition, une des grandes difficultés est de déterminer la véracité d'un témoignage. En Amérique du Nord, un système judiciaire opposant les deux parties est utilisé pour aider le juge et/ou le jury à évaluer la véracité des différents témoignages et à émettre un jugement. Il arrive quelques fois que les circonstances entourant un procès rendent ce processus difficile soit par l'absence de témoin externe, par manque de preuve, ou à cause du jeune âge de la victime, par exemple. Dans ces situations, la cour ou une des deux parties ont parfois recours à un témoin expert dans l'espoir qu'il pourra apporter des éléments de preuve additionnels. Pour qu'un témoin soit reconnu comme expert, il faut démontrer que ce dernier a des connaissances professionnelles non-accessibles au commun des mortels et pouvant aider au processus décisionnel. Ainsi, plusieurs psychologues, médecins et autres professionnels ont donné leur avis sur la crédibilité d'un témoignage, et ce pour différents procès tant au Québec que dans le reste de l'Amérique du Nord. C'est à partir de ce "besoin juridique", et de l'intérêt pécuniaire qui en découle pour les témoins experts, que se sont développés, au cours des années, une multitude d'outils, de techniques et de procédures ayant la réputation de distinguer les vrais témoignages des faux.

D'un point de vue historique, il y a deux approches à l'évaluation de la véracité d'un témoignage. La première approche est basée sur la crédibilité de la personne et découle du postulat selon lequel il y a plus de chance qu'un témoignage soit vrai s'il origine d'un témoin crédible et qu'il y a plus de chances qu'un témoignage soit faux s'il origine d'un témoin non crédible. Cette approche repose sur une évaluation globale de la personne et de ses antécédents. L'utilisation de témoins de moralité, quoique différente d'une expertise, origine du même postulat de base puisque l'objectif de cette procédure est de s'assurer de la moralité, et par conséquent de la crédibilité, du témoin visé. L'évaluation psychologique, et tous les instruments psychométriques qui y sont rattachés, est probablement une des méthodes les plus utilisées de cette approche. Une seconde méthode consiste en l'évaluation des antécédents du témoin à plusieurs niveaux, familial, juridique, criminel, psychiatrique, ou autre.

Un des problèmes majeurs de cette première approche est qu'en se basant sur la crédibilité de la personne, on tend à discriminer certains groupes de témoins. Par exemple, si une évaluation des antécédents est effectuée, la déposition de témoins ayant un passé criminel

ou psychiatrique ou ayant eu des comportements marginalisés comme la prostitution, par exemple, sera probablement considérée a priori comme moins crédible, indépendamment de son contenu. Un autre problème tout aussi important soulevé par ce type d'approche concerne le postulat de base. Le lien n'est pas si clair entre la crédibilité d'un témoin et la véracité de son témoignage dans une circonstance particulière. Ce type de postulat ne tient pas compte de la possibilité que des personnes crédibles puissent ne pas dire la vérité ou que des personnes non-crédibles puissent dire la vérité.

Pour pallier ces difficultés, une deuxième approche basée sur l'évaluation du témoignage en tant que tel fut développée. Deux types de procédures sont associées à cette approche, soit des procédures coercitives et des procédures évaluatives. Les procédures dites coercitives ont pour but de rendre le témoin incapable de mentir. L'interrogatoire sous hypnose et la narcoanalyse, nom donné à l'entrevue à l'amobarbital de sodium, sont les deux meilleurs exemples de procédures coercitives. Malheureusement, leur utilisation est très controversée, entre autres à cause des problèmes de suggestibilité inhérents à leur usage (p. e. Day, 1998; Spanos, Burgess et Burgess, 1994; Gunn et Gudjonsson, 1988 dans Porter et Yuille, 1996). Les procédures dites évaluatives se basent sur le postulat que quelqu'un qui ment laisse échapper involontairement des indices qui différencient son témoignage de celui de quelqu'un qui dit la vérité. Ces procédures sont donc axées sur la détection de tels indices. Ces indices peuvent être physiologiques, comportementaux ou même verbaux.

L'entrevue au polygraphe représente le meilleur exemple de procédure physiologique. Dans cette technique, l'individu est interrogé sur des éléments critiques de son témoignage pendant que diverses mesures d'activation physiologiques sont enregistrées. L'analyse de ces enregistrements permettrait de détecter des réponses non véridiques. Par contre, malgré des débuts prometteurs, cette technique est assez controversée. Certaines recherches semblent démontrer qu'elle produit un nombre élevé de faux positifs (Honts et Perry, 1992), c'est-à-dire qu'un nombre élevé de déclarations valides sont faussement catégorisées comme trompeuses, et qu'elle peut faussement reconnaître quelqu'un qui a tendance à se sentir coupable comme coupable (p. e. Honts, Raskin et Kircher, 1994).

Les indices comportementaux ont aussi beaucoup fait parler d'eux. Par exemple, la mesure des délais de réponse à des exercices d'association libre fut proposée par Jung (1987, p.167) pour différencier les personnes coupables des personnes non coupables. De plus, des recherches ont été effectuées sur divers comportements tels les expressions faciales (voir Ekman et Friesen, 1969; Ekman, 1992). Le plus grand problème lié à l'utilisation des indices comportementaux en contexte légal est que malgré un certain espoir créé par la

recherche, aucune procédure n'est présentement disponible pour en faire l'évaluation. Il ne reste finalement que les indices verbaux.

Quelques procédures utilisant des indices verbaux pour évaluer la crédibilité d'un témoignage existent déjà. La plupart de celles qui existent (voir Sapir, 1987, dans Porter et Yuille, 1996) sont généralement peu utilisées en cour, sont mentionnées dans peu de publications et sont peu ou pas appuyées par la recherche. Par contre, une procédure verbale se distingue des autres. De loin la plus connue et la plus utilisée (Vrij, Kneller et Mann, 2000), elle s'appelle le "Statement Validity Analysis (SVA)", traduit ici par *analyse de la validité de la déclaration*. Dans le présent travail, le terme anglais et son acronyme, le SVA, seront en général préférés à la traduction française, et ce pour faciliter la compréhension, étant donné que ce sont les termes généralement utilisés dans les écrits scientifiques. Cette procédure qui origine de l'Allemagne de l'Ouest des années cinquante consiste en une entrevue semi-structurée dont la transcription écrite est ensuite analysée à partir d'une liste de critères précis. Deux instruments sont utilisés pour évaluer la véracité du témoignage dans cette procédure. Le premier instrument, appelé "Criteria-Based Content Analysis (CBCA)", consiste en une liste de dix-neuf critères utilisés pour analyser le contenu de la déclaration. Le deuxième instrument consiste en un ensemble de consignes rassemblées dans une liste de vérification appelée la "Validity Checklist (VCL)" et utilisées pour évaluer le contexte entourant la déclaration. Vu les limites et controverses reliées aux autres types de procédures, un nombre grandissant d'experts semblent utiliser l'analyse de la validité de la déclaration lorsqu'ils doivent évaluer la crédibilité d'un témoignage. Mais qu'en est-il de la validité et de l'efficacité de cette procédure? Pour mieux comprendre ce en quoi l'analyse de la déclaration consiste, un bref historique sera dressé dans les prochains paragraphes.

Historique du SVA

Créé au départ pour valider les déclarations d'abus sexuels effectuées par des enfants, le Statement Reality Analysis (SRA), prédécesseur du SVA, se doit d'être remis en contexte. Dans l'Allemagne des années 1950 et 1960, la déclaration d'un enfant, même assermentée, avait une valeur probante moindre que celle d'un adulte. Comme il était rare, lors de procès concernant une allégation d'agression sexuelle, qu'un adulte puisse corroborer le témoignage de l'enfant présumé agressé, ces procès se terminaient généralement par un non-lieu ou un verdict de non-culpabilité par manque de preuve (Casoni, 1999). Pour pallier ce problème, le système juridique ouest-allemand commença à accepter de considérer les témoignages d'enfants si et seulement si ces derniers étaient reconnus comme crédibles par un témoin expert. Cette position juridique, qui était à l'époque

commune à d'autres pays, comme par exemple la Suède, amena les différents professionnels concernés à chercher à développer des techniques permettant d'évaluer la véracité des témoignages d'enfants. Ainsi, plusieurs psychologues allemands et suédois de cette époque développèrent une expertise dans l'évaluation des déclarations d'enfants (Yuille, 1988).

C'est dans ces circonstances qu'Undeutsch introduisit, en 1954, la version originale du SRA (voir Undeutsch, 1989). Dans les mêmes années, une technique similaire fut aussi développée par un expert suédois nommé Trankell (Dittmann, 1998; Yuille, 1988). Undeutsch et Trankell présentèrent, dans leurs articles originaux et par la suite, des études de cas pour démontrer l'utilité de leur approche respective. Les deux auteurs insistèrent sur le fait que l'application du SRA demande une habileté clinique qui ne peut être acquise que par une pratique intensive. Par conséquent, leurs descriptions de la procédure étaient souvent moins que satisfaisantes. Étant donné que cette procédure répondait à un besoin juridique pressant de l'époque, elle n'a pas eu à être validée empiriquement avant de devenir un standard dans les cours allemandes et suédoises. Il a même fallu attendre quelques décennies avant qu'un souci de vérification empirique ne se manifeste. L'hypothèse à la base du SRA, voulant que les déclarations qui originent de souvenirs d'expériences réelles diffèrent en contenu et en qualité des déclarations inventées ou imaginées, est maintenant nommé l'hypothèse d'Undeutsch par la plupart des auteurs (p. e. Horowitz, 1991; Raskin et Esplin, 1991).

Undeutsch élaborait l'ensemble de sa procédure principalement grâce à sa vaste expérience clinique dans le domaine de l'évaluation des déclarations d'enfants. Il est donc plus exact de qualifier l'origine du SRA d'intuitive que d'empirique. Une des grandes difficultés lorsqu'une technique d'évaluation est basée sur l'expérience clinique est que le jugement clinique de l'expert est plus renforcé par le jugement de la cour que par la validité historique du témoignage évalué. Plus précisément, étant donné l'énorme difficulté généralement rencontrée à s'assurer de la validité historique du témoignage de l'enfant par des moyens indépendants du processus d'évaluation de la déclaration, sans compter que la cour a évidemment peu recours à des experts lorsqu'il y a confession ou rétractation, le jugement clinique d'un expert est probablement renforcé principalement par son habileté à convaincre juges et jurys et non par la véracité même du témoignage.

Jusqu'au début des années 1980, très peu de recherches furent effectuées sur le SRA. Pendant ce temps, soit entre les années 1960 et 1980, la réalité judiciaire se transforma peu à peu. Grâce entre autres aux changements de mentalité par rapport aux agressions sexuelles sur des enfants et au travail de certains chercheurs comme Ceci, Toglia et Ross

(1987), des réformes juridiques s'effectuèrent dans plusieurs pays à propos des témoignages d'enfants. Ainsi, presque toutes les cours européennes et nord-américaines abandonnèrent, au cours des années 1980, l'exigence de la corroboration du témoignage de l'enfant dans les cas d'agressions sexuelles (Casoni, 1999). La principale raison d'être du SRA, soit de valider le témoignage des enfants autrement irrecevable, n'était plus. Par contre, certains pays comme la Hollande exigent toujours des preuves corroborant les témoignages d'enfant, même si l'accusé confesse (Lamers-Winkelman et Buffing, 1996). Pour survivre à ces nouveaux systèmes qui avaient moins besoin d'eux, les experts en évaluation de la véracité de témoignages durent de plus en plus rendre des comptes en démontrant l'efficacité de leurs procédures.

C'est dans ce nouveau contexte que fut créée la seconde version du SRA. Dans le but de standardiser la procédure, de la rendre compatible à la réalité nord-américaine et de la rendre plus objective et plus apte à être validée empiriquement, une équipe internationale d'experts constituée de Max Steller (Allemagne), Guenter Köhnken (Allemagne), David C. Raskin (États-Unis) et John C. Yuille (Canada) modifièrent et définirent plus objectivement le *Statement Reality Analysis* vers la fin des années 1980 (Steller, 1989; Raskin et Esplin, 1991; Yuille, 1988). La nouvelle version prit le nom de *Statement Validity Analysis* (SVA). L'analyse de contenu opérationnalisée en dix-neuf critères (voir tableau 1 pour un résumé des critères) fut nommée *Criteria-Based Content Analysis* (CBCA) et des indications supplémentaires aidant à évaluer le contexte dans lequel la déclaration fut produite furent rassemblées sous le terme *Validity Checklist* (VCL). Un résumé des critères du CBCA est présenté au tableau 1 tandis que les différents éléments du VCL sont présentés dans l'annexe A. Le changement de terme, de *Reality* à *Validity*, se voulait une façon plus juste de décrire la procédure. Dans ce cas, le terme *Validity* fait référence à des déclarations basées sur des expériences personnelles, même si quelques détails de la déclaration peuvent ne pas être exacts, par opposition à des déclarations non-valides qui sont inventées par l'enfant ou sont la conséquence de l'influence d'autrui. Ainsi, une déclaration n'a pas à rapporter tous les détails d'un événement de façon exacte pour pouvoir être considérée valide. La validité d'un témoignage est donc essentiellement une opérationnalisation de sa véracité. Pour mieux apprécier les détails de la procédure, le lecteur est invité à consulter l'article de Raskin et Esplin (1991) la décrivant en profondeur. Un détail important à noter est que malgré que le SVA fût conçu à l'origine pour évaluer les déclarations d'enfants contenant des allégations sexuelles, de plus en plus d'auteurs (Steller et Köhnken, 1989; Porter et Yuille, 1996). considèrent que la technique peut aussi être utilisée pour évaluer les déclarations d'adultes contenant des allégations autres que sexuelles.

.....
Insérer tableau 1 environ ici
.....

Les recherches empiriques commencèrent réellement avec l'introduction du SVA et évaluèrent principalement la validité psychométrique du CBCA en comparant les résultats du CBCA à un critère indépendant de véracité. Deux types d'approches furent utilisées pour évaluer et valider cette procédure, soit une approche quasi-expérimentale utilisant des cas réels et une approche expérimentale utilisant des situations analogues à la situation réelle. Étant donné les différences entre les implications de ces deux approches, les recherches effectuées sur le SVA seront revues en fonction de l'approche utilisée.

Recherches utilisant des cas réels (quasi-expérimentales)

Les principales études sur des cas réels seront résumées dans les prochains paragraphes. Étant donné que la recherche sur le SVA ne fait que débiter, il y a une partie de ces études qui sont difficiles d'accès car elles ne furent présentées dans leur entier que dans le cadre d'une conférence scientifique. Par conséquent, des résumés publiés dans des articles furent utilisés lorsque la source d'origine n'a pu être consultée.

Avant que les premières études empiriques ne soient réalisées, certaines statistiques étaient disponibles par rapport à l'utilisation du SVA. Par exemple, Artzen (1982, dans Bradford, 1994) dit avoir évalué environ 24 000 cas, dont 92% étaient reliés à des allégations d'abus sexuel sur des enfants, et que, dans neuf cas sur dix, la décision de la cour à propos de la véracité du témoignage de l'enfant était la même que celle du psychologue utilisant le SVA. À prime abord, ces données semblent appuyer l'utilisation du SVA. Par contre, le critère utilisé comme indicateur de véracité pour les déclarations, soit la décision de la cour, est très critiquable, sans compter une certaine circularité de l'argument étant donné que la décision de la cour dépend en partie de l'évaluation du psychologue.

Tableau 1

Brève description des critères du Criteria-Based Content Analysis.^a

Caractéristiques générales de la déclaration

1. *Structure logique.* Une cohérence interne, une consistance et l'inclusion de différents détails indépendants qui décrivent la même suite d'événements augmentent la crédibilité.
2. *Production non-structurée.* Un narratif relativement non-structuré est perçu comme étant plus crédible, puisqu'il est supposé que les déclarations fabriquées sont caractérisées par un type de présentation continu, hautement structuré et chronologique.
3. *Quantité de détails.* Une abondance de détails est un indicateur de crédibilité.

Contenu spécifique de la déclaration

4. *Enchâssement contextuel.* Des narratifs qui relient les événements à des situations sociales et physiques et qui font des liens avec d'autres événements sont perçus comme plus crédibles.
5. *Descriptions d'interactions.* Des descriptions d'actions et réactions interreliées sont perçues comme des indications de crédibilité.
6. *Rappel de conversation.* La crédibilité est supportée lorsqu'il y a reproduction directe des paroles d'au moins une personne, surtout si des particularités verbales attribuées à quelqu'un d'autre sont rapportées.
7. *Références à des complications inattendues.* Rapporter des complications inattendues durant l'incident, comme par exemple des interruptions non-planifiées ou des difficultés, augmente la crédibilité.

Particularités du contenu

8. *Détails inusités.* Un contenu très concret et montrant un haut degré de vivacité supporte la crédibilité d'une déclaration. Puisque les détails inusités ont une faible fréquence d'occurrence, on ne s'attend pas à ce qu'ils apparaissent dans des histoires inventées.
9. *Détails superflus.* Des détails superflus sont perçus comme des indicateurs de crédibilité puisque des personnes cherchant à mentir n'inventeraient probablement pas des détails superflus et sans importance.
10. *Détails non compris mais rapportés de façon exacte.* Lorsqu'un enfant rapporte des actions ou mentionne des détails sans en comprendre la signification alors qu'ils sont compris par l'interviewer, la crédibilité de la déclaration est augmentée.
11. *Références à des incidents extérieurs.* Lorsqu'un témoin rapporte des conversations qui réfèrent à des événements extérieurs à l'allégation mais qui y sont reliées en termes de contenu, la crédibilité est augmentée.

12. *Références à ses propres états subjectifs.* Des mentions à propos de l'état mental du témoin durant l'événement, comme des pensées ou des émotions, augmente la crédibilité.

13. *Attribution d'un état psychologique à l'abuseur.* Attribuer des états mentaux et des intentions à l'abuseur est un signe de crédibilité dans une déclaration.

Contenus relatifs aux motivations de la déclaration

14. *Corrections spontanées.* Se corriger spontanément durant l'entrevue ou produire des rappels nouveaux ou plus clairs est vu comme un indicateur de crédibilité.

15. *Aveu d'absence de souvenir.* Admettre une absence de souvenir augmente la crédibilité car il est supposé que les personnes rapportant un faux témoignage ne voudront pas admettre une absence de souvenir.

16. *Doutes à propos de sa propre déclaration.* Soulever des doutes à propos de sa propre déclaration augmente la crédibilité; les personnes rapportant de faux témoignages risquent moins de soulever de tels doutes.

17. *Désapprobation de sa propre participation.* Des mentions de détails non-favorables à l'incrimination ou qui amènent à s'incriminer augmentent la crédibilité de la déclaration.

18. *Le fait d'excuser l'abuseur.* Si la déclaration tend à favoriser l'accusé, par exemple en contenant des explications ou des exonérations, ou si le témoin ne fait pas usage de possibilités évidentes pour incriminer davantage l'accusé, la crédibilité est augmentée.

Éléments spécifiques concernant le délit

19. *Caractéristiques spécifiques du délit.* Des mentions d'éléments caractéristiques du type d'offense qui contredisent le sens commun comme des éléments spécifiques qui dévient des notions communes à propos du comportement sexuel, indiquent une plus grande crédibilité.

^a Traduit librement et adapté de Steller et Köhnken (1989).

Si l'on suppose la dépendance de ces deux décisions, on réalise que la décision de l'expert influence énormément la décision de la cour, ce qui démontre son efficacité à convaincre la cour mais pas nécessairement son efficacité à discerner la véracité de la déclaration. Et si l'on suppose l'indépendance des deux décisions, on peut se demander l'utilité d'une technique qui, neuf fois sur dix, confirme que la cour peut arriver à la même conclusion sans l'aide d'experts.

La première étude empirique répertoriée utilisant des cas réels fut réalisée par Esplin, Boychuk et Raskin (1988) d'après Anson, Golding et Gully (1993) et Raskin et Esplin, 1991a) ou par Esplin, Houed et Raskin (1988) d'après Landry et Brigham. Étant donné la similarité des descriptions de l'étude et le fait qu'elles furent présentées à la même conférence par le même premier auteur, elles seront considérées comme une seule étude. Quarante enfants âgés de 3 à 15 ans y furent interviewés par des psychologues formés à la technique d'entrevue associée au SVA. Dans vingt de ces cas (âge moyen = 9.1 ans), les allégations d'abus sexuel avaient été confirmées, tandis que dans les vingt autres (âge moyen = 6.9 ans), la validité des allégations était douteuse. Les cas confirmés avaient été validés par une confession de la part de l'accusé en dehors d'un contexte de diminution de peine, des évidences médicales sans équivoque ou les deux. Les cas douteux n'avaient aucune preuve corroborant l'incident. De plus, pour les cas douteux, soit la déclaration avait été refusée par la cour, soit aucun procès n'avait été intenté, souvent parce que l'enfant s'était rétracté. 65% des cas douteux incluaient un examen du suspect au polygraphe ayant un résultat positif. L'expérimentateur qui évalua la présence des 19 critères était formé à l'utilisation du CBCA et était aveugle quant à la validité de chaque cas. Landry et Brigham (1992) rapportent que onze des dix-neuf critères étaient présents presque exclusivement dans les déclarations "confirmées" et Anson, Golding et Gully (1993) rapportent que le taux de présence de tous les critères sauf trois soit les détails non-compris mais rapportés de façon exacte, les références à des incidents externes et les doutes à propos de sa déclaration, différenciaient les deux groupes de façon significative ($p < 0.01$) lorsque analysés par un test des signes binomial. Ils rapportent aussi qu'il n'y avait aucun chevauchement entre les scores totaux des deux groupes, les déclarations douteuses ayant des scores entre 0 et 10 (inclusivement) et les déclarations confirmées ayant des scores entre 16 et 34.

Cette étude semble supporter la validité de l'hypothèse d'Undeutsch à la base du SVA et même supporter la validité et l'efficacité de la majorité des critères du CBCA. Par contre, Wells et Loftus (1991) soulignèrent que l'intégrité de l'évaluateur avait peut-être été

compromise par la diffusion médiatique des cas ou des discussions professionnelles à leur sujet et qu'un élément de circularité était assurément présent dans cette étude. En réponse à la question de l'intégrité, Raskin et Esplin (1991b) ont écrit que ce n'était pas le cas car 30 des 40 cas provenaient d'un endroit autre que celui où vivait l'évaluateur. Étant donné la différence d'âge entre les deux groupes, il est aussi possible que la différence soit due à l'âge des enfants. L'équivalence entre les groupes était aussi compromise par rapport à la relation entre l'enfant et l'accusé. Il y avait plus de cas reliés à un divorce ou à un membre de la famille dans le groupe de cas douteux que dans le groupe de cas confirmés. De plus, comme pour Artzen (1982, dans Bradford, 1994), il est difficile de savoir si les déclarations "confirmées" et "douteuses" sont représentatives des déclarations valides et non-valides, quoique l'opérationnalisation est de loin plus rigoureuse que celle proposée par Artzen. Malgré ces limites et problèmes, cette étude représente probablement le support le plus solide pour l'utilisation du SVA.

Pour sa thèse de doctorat, Boychuk (1991, dans Lamers-Winkelman et Buffing, 1996), utilisa un paradigme expérimental similaire et trouva aussi que les critères du CBCA pouvaient distinguer les déclarations douteuses des déclarations confirmées. Par contre, elle trouva aussi un effet de l'âge sur quatre des critères soit l'aveu d'absence de souvenir, les descriptions d'interactions, l'attribution d'états mentaux à l'abuseur et la désapprobation de sa propre conduite. Malheureusement, puisque les informations présentées dans le présent mémoire furent tirées d'une source secondaire, les limites de l'étude peuvent difficilement être analysées.

Dans le but d'évaluer la fidélité inter-juge du CBCA, Anson, Golding et Gully (1993) firent une étude avec les déclarations de 23 enfants âgés de 4 à 12 ans. Toutes les déclarations provenaient d'entrevues vidéo de cas confirmés c'est-à-dire que ces cas incluaient une déposition de l'accusé confessant tous les actes sexuels déclarés par l'enfant. Aucune déclaration douteuse ne fut incluse dans l'analyse. Par contre, deux déclarations considérées comme inventées et deux considérées comme douteuses furent incluses dans l'échantillon pour éviter un trop grand biais durant l'évaluation. Ces quatre déclarations furent éliminées de l'échantillon avant l'analyse. Les enregistrements vidéo furent évalués pour la présence ou l'absence des critères du CBCA par deux des quatre évaluateurs ayant été formés préalablement à l'usage du CBCA. Les évaluateurs étaient aveugles quant au statut des déclarations (confirmée, douteuse ou inventée) et à la proportion de cas confirmés dans l'échantillon. En utilisant des corrélations de Pearson, la proportion d'accord entre les juges varia de 0.39 à 1.00 dépendamment du critère et était en moyenne de 0.75, ce qui était similaire à la valeur de 0.72 rapportée par Steller (1989, décrit dans la prochaine section). Par contre, lorsque des mesures statistiques corrigeant pour les

réponses aléatoires comme le kappa de Cohen et le coefficient d'erreur aléatoire de Maxwell furent utilisées, les auteurs constatèrent que sur les 19 critères, quatre critères (détails superflus, enchâssement contextuel, détails inusités et corrections spontanées) n'avaient qu'une fidélité marginale ($0.50 > \text{Maxwell} > 0.30$) tandis que six autres critères (aveu d'absence de souvenir, référence à des incidents externes, descriptions d'interactions, référence à ses états subjectifs, production non structurée et caractéristiques spécifiques du délit) avaient une fidélité inadéquate ($\text{Maxwell} < 0.30$). La fidélité moyenne était modérée ($\text{Maxwell} = 0.49$). De plus, six critères (structure logique, enchâssement contextuel, descriptions d'interactions, rappel de conversation, excuser l'abuseur et caractéristiques spécifiques du délit) avaient une corrélation significative avec l'âge de l'enfant, et ce malgré des consignes explicites demandant aux évaluateurs d'en tenir compte. L'analyse montrait que la fréquence d'apparition variait énormément d'un critère à l'autre.

Il est important de souligner trois des principales limitations de cette étude, soit le fait que l'évaluation portait sur des enregistrements vidéo et non des transcriptions écrites, que le style d'entrevue n'était pas contrôlé et que seulement des cas confirmés furent utilisés. Par contre, puisque l'objet de l'étude était la fidélité et non la validité de la procédure, ces limitations ne sont probablement que mineures. Les auteurs ont suggéré de définir les critères de façon plus explicite dans un manuel de codage détaillé et de définir des normes par strates d'âge pour éviter de désavantager les jeunes enfants. Ils ont aussi précisé que les plus grands problèmes de fidélité étaient probablement dus aux critères qui ne sont pas définis précisément comme celui de description d'interactions, et à ceux basés sur un jugement d'intensité d'un comportement mais ne fournissant pas de normes tel que dans le cas de détails inusités.

Pour confirmer la relation entre l'âge de l'enfant et le CBCA trouvée entre autres par Boychuk (1991, dans Lamers-Winkelman et Buffing, 1996) et Anson, Golding et Gully (1993), Lamers-Winkelman et Buffing (1996) évaluèrent 103 transcriptions d'entrevue d'enfants de 2 à 11 ans déclarant avoir été abusés sexuellement. Les entrevues ne furent pas évaluées indépendamment pour leur validité. Six critères soit l'enchâssement contextuel, la description d'interactions, le rappel de conversation, les détails superflus, l'aveu d'absence de souvenir et les caractéristiques spécifiques du délit, ont montré une corrélation significative avec l'âge. En comparant aux résultats des deux études précédentes, les auteurs ont conclu que cinq critères soit enchâssement contextuel, description d'interactions, rappel de conversation, aveu d'absence de souvenir et caractéristiques spécifiques du délit, étaient assurément dépendants de l'âge de l'enfant et que le développement de normes tenant compte de l'âge de l'enfant était essentiel si l'on voulait assurer la validité de la procédure.

Dans le but de pousser plus loin l'étude de Anson, Golding et Gully (1993) sur la fidélité inter-juge, Horowitz, Lamb, Esplin, Boychuk, Krispin et Reiter-Lavery (1997) décidèrent d'évaluer la fidélité inter-juge et la fidélité test-retest du CBCA en utilisant un plus grand échantillon. Les transcriptions écrites de cent déclarations d'enfants de 2.6 à 19.8 ans ($\mu = 8.6$) furent évaluées pour la présence des critères du CBCA par trois évaluateurs à deux occasions avec un délai d'environ trois à dix mois entre les deux occasions. Les trois évaluateurs étaient formés à l'utilisation du CBCA. La fidélité inter-juges pour les scores totaux (Pearson) a montré une variation de 0.78 à 0.82 pour la première évaluation et de 0.86 à 0.89 pour la deuxième évaluation, tandis que la fidélité test-retest (Pearson) a varié entre 0.85 et 0.91, ce qui démontre une fidélité hautement satisfaisante. Par contre, lorsque la fidélité inter-juges a été évaluée pour chaque critère et que des mesures correctrices ont été utilisées, des résultats similaires à ceux de Anson, Golding et Gully (1993) furent observés. Trois critères, soit les détails superflus, l'aveu d'absence de souvenir et les corrections spontanées, ont montré une fidélité inadéquate tandis que deux autres critères soit la référence à des incidents extérieurs et la mention de détails inusités, n'ont pas atteint le seuil d'acceptabilité (Maxwell > 0.50) établi par Anson, Golding et Gully (1993). Aucun de ces cinq critères n'a donc atteint le seuil d'acceptabilité dans l'étude d'Anson, Golding et Gully (1993). Les auteurs ont alors conclu qu'il fallait définir plus clairement ces cinq critères avant qu'ils puissent être utilisés, ou tout simplement de les éliminer. Les lecteurs intéressés à approfondir sur la controverse engendrée par cette recherche sont invités à lire la critique de Tully (1998) et les réponses faites à Tully par Horowitz (1998) et Lamb (1998).

Hershkowitz, Lamb, Sternberg et Esplin (1997) cherchèrent à savoir si le type de questions posées durant l'entrevue affecte le nombre de critères du CBCA présents dans le contenu des déclarations d'enfants disant avoir été abusés sexuellement. Seul les quatorze critères ayant une fidélité acceptable selon les conclusions de Horowitz, Lamb, Esplin, Boychuk, Krispin et Reiter-Lavery (1997) ont été utilisés dans l'étude. Vingt transcriptions d'entrevues avec des enfants de 4 à 13 ans ($\mu = 8.4$ ans) ont été évaluées selon les quatorze critères en fonction des interventions de l'interviewer ayant précédé chaque réponse. Les interventions étaient classées en cinq types: des invitations à une réponse ouverte, des facilitateurs, des énoncés directifs, des énoncés dirigés et des énoncés suggestifs. Comme prévu, les invitations ont évoqué des réponses contenant plus de détails et rencontrant plus de critères que tout autre type d'intervention. De plus, elles ont évoqué des réponses plus longues que les interventions directives, dirigées ou suggestives. Ces résultats supportent, à prime abord, l'utilisation fréquente d'invitations à une réponse ouverte lors d'entrevues destinées à l'utilisation du CBCA. Par contre, étant donné qu'aucune distinction n'est faite

entre les déclarations réelles et les fausses déclarations, il est impossible de savoir si cet effet affecte seulement les déclarations réelles comme c'est souhaité ou si l'effet affecte les deux types de déclarations de façon indifférenciée.

Pour résoudre cette question, Hershkowitz (1999) a repris l'expérience de Hershkowitz et coll. (1997) avec 12 déclarations "confirmées" et 12 déclarations "douteuses". Les résultats ont montré que les invitations à des réponses ouvertes évoquent plus de mots, de détails et de contenu se qualifiant comme critère du CBCA, en comptant les répétitions, que les énoncés directifs, dirigés et suggestifs, mais seulement pour les cas confirmés. Par contre, le nombre total de critères différents ne différenciait pas de façon significative les cas confirmés des cas douteux. Ces résultats semblent donc supporter l'utilité des invitations à réponse ouverte malgré le fait que du même coup ils ne réussissent pas à démontrer la validité de la procédure. Il est important de noter que ces résultats n'amènent pas à la conclusion que les entrevues utilisant plus d'invitations ouvertes produisent plus d'énoncés en accord avec les critères. Ils montrent plutôt que, durant une entrevue, les énoncés conformes aux critères ont plus tendance à être produits après une invitation qu'après un autre type d'intervention dans les déclarations confirmées tandis qu'ils semblent plus ou moins distribués également entre les types d'interventions dans les déclarations douteuses. De plus, cette étude tenait compte des répétitions, ce qui n'est pas du tout le cas de la procédure formelle du CBCA. Une étude semblable effectuée par Craig, Scheibe, Raskin, Kircher et Dodd (1999) a produit des résultats similaires. Comme la recherche sur les styles d'entrevue n'est que partiellement reliée aux objectifs de la présente étude, le lecteur est invité à consulter la source originale pour approfondir le sujet.

Critiques sur les recherches quasi-expérimentales

Les recherches effectuées avec des cas réels nous amènent à certaines conclusions sur l'utilisation du SVA comme procédure d'évaluation de la crédibilité. Pour ce qui est de la fidélité (test-retest et inter-juges) du score global du CBCA, ces études semblent démontrer qu'elle est adéquate. Cependant, lorsque les critères sont analysés séparément et que les corrections statistiques sont apportées, la fidélité est remise en question pour un minimum de cinq des dix-neuf critères lesquels devraient être éliminés ou redéfinis avant d'utiliser la procédure comme outil d'évaluation. Le manque de clarté et de précision dans la définition de ces critères semble être la cause principale de ce problème et une meilleure opérationnalisation des critères semble être la seule façon d'améliorer la situation. Étant donné que, par définition, la validité d'une mesure ne peut dépasser sa fidélité, il est important de se rappeler ces problèmes lorsque la validité du SVA est évaluée.

En ce qui a trait à la validité, les résultats de Esplin, Boychuk et Raskin (1988) et de Boychuk (1991, dans Lamers-Winkelmann et Buffing, 1996) semblent attribuer à la procédure une efficacité remarquable, mais il est bon de souligner qu'une autre étude, celle de Hershkowitz (1999), ne réussit pas à démontrer la validité de la procédure. Aussi, un nombre important de problèmes méthodologiques ont été soulevés par certains dont Wells et Loftus, (1991), sans compter les conséquences sur la validité du problème de fidélité mentionné plus haut. L'ensemble de ces interrogations génère une certaine controverse quant à la possibilité de généraliser ces résultats. Les principales critiques concernent probablement l'utilisation de cas "confirmés" et "douteux" pour définir la validité de façon opérationnelle. Premièrement, pour valider le SVA avec des déclarations réelles, dont la validité est impossible à évaluer avec certitude, il faut utiliser une mesure indépendante et indirecte de la validité et comparer les résultats du SVA aux catégories définies par la mesure indépendante. En conséquence, la validité d'un témoignage fut définie par la quantité et la qualité des évidences corroboratives et/ou le jugement de la cour. Malheureusement, les éléments servant à la mesure indépendante ne sont pas si indépendants que ça de la qualité du contenu de la déclaration. De plus, la validité d'un témoignage est confondue avec sa crédibilité, malgré qu'il y ait plusieurs facteurs possibles pour qu'un témoignage valide soit peu crédible, les plus évidentes étant les limites des capacités intellectuelles et verbales de l'enfant, ou encore son âge. Il est probable que, par exemple, la crédibilité d'une déclaration affecte les décisions du juge ou du procureur, ou même celle de l'accusé de se confesser ou non. La difficulté de l'enfant à convaincre son entourage ou les autorités entre autres à cause du manque de crédibilité de son témoignage ou du manque d'évidences additionnelles, pourrait même l'amener à se rétracter. Quoique les évidences utilisées par Esplin, Boychuk et Raskin (1988) pour classifier les déclarations comme confirmées sont convaincantes, il n'est pas si évident qu'un manque d'évidences équivaut à un manque de validité et l'on peut supposer qu'une partie des déclarations "douteuses" est totalement valide.

Un problème de taille se pose aussi en rapport aux circonstances dans lesquelles la procédure est utilisée. Lorsqu'un nombre suffisant d'évidences corroborent le témoignage de l'enfant, un témoin expert est rarement nécessaire, même en Allemagne. Au contraire, l'utilité d'une expertise réside dans son habileté à amener des évidences supplémentaires lorsque les évidences de base sont minces ou contradictoires, ce qui veut dire que les résultats obtenus en analysant des déclarations "confirmées" sont peu utiles pour l'application légale de la technique. Par conséquent, pour qu'une technique soit utile, elle doit parfois arriver à des conclusions différentes, mais généralement correctes, de celles produites par l'évaluation habituelle des évidences. Il n'y a aucune utilité pour une

technique qui amène aux mêmes conclusions que le système juridique, c'est-à-dire qui va toujours dans le sens des évidences sans rien y ajouter. L'étude d'Esplin, Boychuk et Raskin (1988) ne répond malheureusement pas à cette exigence, la différence de score étant tellement grande entre les deux groupes. Comme le problème de confusion entre crédibilité et validité est inhérent à toutes les études utilisant des cas réels mais absent des études expérimentales, il est légitime de penser que si la procédure évalue principalement la validité, son efficacité sera plus grande dans les études expérimentales. Par ailleurs, si elle évalue principalement la crédibilité, et non la validité, son efficacité sera plus grande dans les expériences utilisant des cas réels. Cette comparaison sera effectuée après le survol des études expérimentales sur le SVA.

Un autre fait à considérer est que les scores de CBCA semblent réagir fortement à l'âge de l'enfant, et parfois plus fortement qu'à la validité de la déclaration. Cet effet semble constant à travers les études prenant en compte cette variable. Par contre, les critères spécifiques dépendants de l'âge diffèrent parfois d'une étude à l'autre. Cinq critères en particulier, soit l'enchâssement contextuel, la description d'interactions, le rappel de conversation, l'aveu d'absence de souvenir et les caractéristiques spécifiques du délit, sont considérés à risque, du moins jusqu'à ce que des normes stratifiées par strates d'âge soient disponibles.

Finalement, il est difficile de conclure avec certitude à la supériorité des invitations à une réponse ouverte dans le cadre du SVA car premièrement, dans l'étude de Hershkowitz (1999), le CBCA fut incapable de différencier les déclarations douteuses des déclarations confirmées, deuxièmement, l'efficacité des invitations n'a rien à voir avec l'efficacité des entrevues utilisant plus d'invitations, et troisièmement, le fait que le taux d'invitations n'est pas manipulé, par exemple en ayant une partie des d'entrevues contenant plusieurs invitations et l'autre en contenant moins, le lien de causalité est difficile à faire.

Recherches analogues (expérimentales)

Pour pallier les difficultés liées aux études avec des cas réels, certains chercheurs ont entrepris d'évaluer le SVA en contexte expérimental. Les principales études de ce type seront résumées dans les prochains paragraphes.

Yuille (1989) a mené à terme une des premières études analogues du SVA. Pour en évaluer la validité, il a utilisé les déclarations de 49 enfants âgés de 6 à 9 ans. Ces enfants avaient à raconter deux histoires, soit une vraie et une fausse (mais plausible), et avaient deux jours pour se préparer avant d'être interviewés par un expérimentateur aveugle quant à la véracité des histoires. Les déclarations furent choisies par les enfants eux-mêmes. Elles ont été évaluées selon les critères du CBCA par deux évaluateurs formés à cette technique et qui

étaient eux aussi aveugles quant à la véracité des récits. Une déclaration était classée comme valide si elle contenait les cinq premiers critères (structure logique, production non-structurée, quantité de détails, enchâssement contextuel et descriptions d'interactions) plus deux autres. Après discussion entre les évaluateurs et atteinte d'un consensus par rapport à leurs désaccords sur 4% des classifications, 90.9% des vraies histoires et 74.4% des fausses histoires furent correctement classées. Ces résultats supportent en général la validité et la fidélité inter-juges des scores globaux du CBCA pour les enfants de ce groupe d'âge. Par contre, ils soulignent aussi un certain biais vers les faux positifs, soit une tendance à reconnaître comme crédibles des déclarations trompeuses. Les principales limites de cette étude sont que le type d'événements racontés ne fut pas contrôlé et était très différent des allégations d'abus sexuel rencontrées en cour. On peut quand même prétendre que malgré l'utilisation légale qui est faite de cette technique, ses postulats de bases impliquent une certaine possibilité de généraliser les résultats puisqu'ils s'appuient sur des principes généraux du comportement humain et définissent des qualités intrinsèques différenciant les déclarations valides des déclarations non-valides peu importe le contexte.

Steller (1989) a, lui aussi, voulu évaluer la validité du CBCA en contexte expérimental. Dans le but d'améliorer l'analogie à une déclaration d'abus sexuel, il interrogea des enfants à propos d'expériences répondant aux trois conditions suivantes: (1) le témoin était impliqué directement dans l'événement, (2) l'événement avait produit des émotions négatives chez le témoin, et (3) l'événement impliquait une perte de contrôle chez le témoin. Il demanda aux 88 enfants participant à l'étude, la moitié étant en première année et l'autre en quatrième année, de raconter deux histoires, soit une réelle et une inventée, qu'ils devaient choisir dans une liste de thèmes permis fournie par l'expérimentateur dans le contexte d'une compétition de conteurs d'histoires. La véracité des histoires réelles était confirmée en interviewant les parents. Les déclarations ont alors été évaluées selon dix-huit des dix-neuf critères du CBCA par des évaluateurs aveugles quant à la véracité des histoires. Des analyses de variance ont été effectuées sur les 176 histoires recueillies. Une première analyse a démontré que, pour que l'hypothèse d'Undeutsch soit corroborée, les histoires basées sur des thèmes non-médicaux, comme être attaqué par un chien ou battu par un autre enfant, devaient être éliminées et seules les histoires médicales, comme une prise de sang ou une visite chez le dentiste, devaient être conservées pour l'analyse. Un total de 79 histoires véridiques et de 48 histoires inventées a donc été conservé. L'analyse des histoires à thème médical montre que onze des dix-huit critères (structure logique, quantité de détails, enchâssement contextuel, complications inattendues, détails inusités, détails superflus, détails non-compris mais rapportés de façon exacte, références à des

incidents extérieurs et expériences subjectives) ont été observés plus fréquemment dans les déclarations véridiques que dans les descriptions inventées. Aucun des critères n'expliquait plus de 15% de la variance entre ces deux groupes.

Ces résultats supportent, en général, la validité du CBCA en contexte expérimental malgré le fait que l'efficacité, c'est-à-dire l'importance de l'effet, semble faible. Par contre, le fait que certaines histoires ont dû être éliminées de l'analyse pour que l'hypothèse de départ soit confirmée pose un sérieux problème du point de vue méthodologique. Il faut se rappeler que lorsque toutes les histoires recueillies sont utilisées dans l'analyse, les résultats infirment la validité et l'efficacité du CBCA. Les résultats obtenus pourraient donc être dus à une particularité des histoires médicales non-généralisable à des histoires d'abus sexuel. De plus, il n'est fait aucune mention dans l'article de l'utilisation de mesures statistiques, comme la correction de Bonferroni, pour tenir compte du fait que les données tirées des histoires médicales furent utilisées dans plus d'une analyse statistique. Il faut donc interpréter ces résultats avec prudence.

Dans le but de pousser plus loin l'expérience de Steller (1989), Landry et Brigham (1992) ont évalué la validité du CBCA avec une population adulte et une procédure plus complexe, dans laquelle chaque incident était enregistré dans un format vidéo. Les 70 participants ont raconté deux incidents personnels répondant aux trois conditions établies par Steller (1989). Un des incidents était réel et l'autre était inventé à partir d'un thème précis donné par l'expérimentateur. Les thèmes inventés étaient choisis parmi des thèmes ayant déjà été mentionnés comme réels par d'autres participants. Les participants recevaient leurs thèmes deux jours avant l'entrevue. Ils ont reçu la consigne de préparer des descriptions d'environ une à deux minutes. Les 140 enregistrements vidéos étaient évalués par sept évaluateurs par rapport à trois critères, soit l'intensité apparente du caractère traumatisant de l'incident, l'intensité émotionnelle apparente du participant lors de l'entrevue et l'intensité apparente de la perte de contrôle durant l'incident. De ces 140 enregistrements, 12 vignettes ont été sélectionnées comme stimuli de base de l'expérience, soit 6 réelles et 6 inventées dont l'intensité a été évaluée comme modérément intense sur les trois facteurs. Le but de cette sélection était de s'assurer que les vignettes utilisées étaient de qualité équivalente. Ces 12 vignettes ont été évaluées par 114 nouveaux participants sous une des quatre conditions suivantes. La moitié des participants avait reçu une formation de 45 minutes dans l'utilisation du CBCA avant d'évaluer les vignettes tandis que l'autre moitié n'avait reçu aucune formation. De plus, la moitié des participants de chaque groupe évaluait les enregistrements vidéos des entrevues tandis que l'autre moitié en évaluait les transcriptions écrites. Un délai de trois minutes par vignette était donné aux participants pour faire l'évaluation. Cinq des 19 critères du CBCA n'ont pas été utilisés, soit parce qu'ils étaient

spécifiques aux déclarations d'abus sexuels, ce qui était le cas pour les détails non-compris, le fait d'excuser l'abuseur et les caractéristiques spécifiques du délit, ou soit parce que les participants ne semblaient pas les comprendre ou les appliquer de façon satisfaisante lors d'une étude pilote, ce qui était le cas pour la production non-structurée et les références à des incidents extérieurs. Lors de l'évaluation, les participants devaient indiquer si chaque critère était absent, présent ou fortement présent et évaluer la probabilité que le récit soit véridique à l'aide d'une échelle de 0 à 9. Les cotes de 5, correspondant à des jugements incertains, étaient éliminées de l'analyse. Elles formaient 10.1% des réponses. Seuls les participants ayant été formés au CBCA et évaluant l'enregistrement vidéo ont eu un taux de précision significativement plus élevé que celui résultant du hasard, soit un taux de 58.1%. L'efficacité des critères individuels a été évaluée par une analyse de variance multivariée.

L'analyse a démontré que les critères étaient plus souvent évalués comme présents et fortement présents pour les déclarations réelles que pour les déclarations inventées. Par contre, les résultats sont beaucoup moins spectaculaires et démontrent une moins grande efficacité que ceux de Esplin, Boychuk et Raskin (1988). L'analyse de chaque critère a montré que deux des critères, soit les complications inattendues et la désapprobation de sa propre conduite, ne différenciaient pas de façon significative les déclarations réelles et inventées tandis que deux autres, soit la structure logique et l'attribution d'états subjectifs, étaient plus souvent présents dans les déclarations inventées. Les dix autres critères différenciaient de façon significative les deux types de déclaration.

Ces données amènent un appui ambivalent à la validité du SVA, du moins chez des participants adultes. Le fait que l'évaluation globale n'était efficace qu'avec l'enregistrement vidéo nous amène à penser que l'utilisation d'indices supplémentaires, comme les comportements non-verbaux par exemple, a pu se produire et que les critères du CBCA, à eux seuls, étaient incapables d'amener les participants à évaluer la véracité des déclarations de façon efficace. En ce qui a trait aux critères individuels, dix d'entre eux différencient de façon statistiquement significative les déclarations réelles et inventées dans la direction prévue par le CBCA. Par contre, il est bon de noter que quatre de ces critères font partie de ceux que Horowitz et al. (1997) avaient suggéré d'éliminer ou redéfinir dû à leur fidélité inadéquate.

Il faut quand même souligner certaines limites de l'étude. Par exemple, malgré le fait que la technique utilisée pour sélectionner les vignettes diminue les risques d'avoir des groupes non-équivalents, elle diminue aussi les chances de détecter tout effet qui covarie naturellement avec une des trois variables contrôlées. De plus, un problème digne de mention est lié à la façon dont les récits ont été choisis. De par la nature du paradigme

expérimental, il y a confusion entre la véracité du récit la modalité de choix. Plus précisément, il est possible que les déclarations basées sur un thème imposé soient qualitativement différentes des déclarations basées sur des thèmes choisis. Par exemple, si un participant se sent mal à l'aise avec un thème particulier, il risque de ne pas choisir de l'utiliser s'il en a la possibilité. Ce malaise pourrait se traduire, par exemple, par une utilisation restreinte de détails superflus. Le fait de limiter la durée des déclarations peut aussi amener les participants à résumer l'incident au lieu de tout rapporter et donc à produire des déclarations incomplètes. Une autre limitation à considérer est le fait que la variable véracité est confondue avec l'ordre de présentation car, lors de l'entrevue, les déclarations réelles furent toujours rapportées avant les déclarations inventées. Par conséquent, l'interviewer n'était pas aveugle quant à la validité de la déclaration. Il est donc possible que ce dernier ait biaisé involontairement les résultats du fait qu'il connaissait à l'avance la validité de la déclaration rapportée. Finalement, on peut soulever l'argument à l'effet qu'une formation de 45 minutes est insuffisante pour assurer une bonne compréhension de la technique nécessaire pour assurer une fidélité adéquate durant l'évaluation ou par ailleurs, qu'un délai de trois minutes est insuffisant pour évaluer les 14 critères.

Joffe et Yuille (1992, dans Lamers-Winkelmann et Buffing, 1996) ont cherché à savoir si le CBCA pouvait différencier les récits basés sur des expériences des récits basés sur des instructions détaillées ("coaching"). Les 142 participants, des enfants de deuxième et quatrième année, étaient assignés à l'une des trois conditions suivantes: (a) les témoins de l'événement étaient impliqués dans un événement monté par les expérimentateurs, (b) les enfants fortement entraînés n'étaient pas témoins de l'événement mais en recevaient une description détaillée et on leur faisait mention de certains éléments caractéristiques de plusieurs critères du CBCA, et (c) les enfants légèrement entraînés ne recevaient qu'un bref résumé de l'événement. Chez les enfants de quatrième année, les scores globaux du CBCA ont permis de différencier de façon significative les déclarations des témoins de celles des enfants légèrement entraînés mais non les déclarations des témoins de celles des enfants fortement entraînés. Chez les enfants de deuxième année, aucune différence n'a été trouvée entre les déclarations des trois groupes. Cette étude semble démontrer que l'entraînement d'un enfant à l'aide des critères du CBCA peut amener sa déclaration à être indifférenciable d'une déclaration réelle en plus de démontrer une fois de plus l'effet de l'âge sur les critères du CBCA. Malheureusement, puisque les informations ont été tirées d'une source secondaire, les limites de l'étude peuvent difficilement être analysées.

Dans le but de vérifier l'impact des différences culturelles sur l'efficacité du CBCA, Ruby et Brigham (1998) ont effectué une expérience avec un paradigme expérimental similaire à

celui utilisé par Landry et Brigham (1992). Ils ont comparé les déclarations de six participants caucasiens à celles de six participants afro-américains. Leurs hypothèses principales étaient qu'étant donné que le CBCA a été créé dans un milieu européen et caucasien, (1) il serait plus efficace pour reconnaître les déclarations réelles chez les caucasiens que chez les afro-américains, et (2) les déclarations réelles faites par des afro-américains rencontreraient moins de critères du CBCA que celles faites par des caucasiens. Les déclarations étaient choisies selon les mêmes critères que dans l'étude de Landry et Brigham (1992), à la différence qu'aucun thème n'était imposé, et l'entrevue se déroulait dans les mêmes conditions. Chaque participant avait à raconter premièrement l'événement réel et deuxièmement l'événement inventé. Les 24 déclarations étaient divisées en deux groupes contenant un nombre égal de déclaration de chacune des quatre catégories et tout élément de la déclaration identifiant l'ethnie du participant était éliminé des transcriptions écrites. Cent soixante-neuf participants ont été recrutés pour évaluer ces transcriptions. De ces 169 participants, 26 n'ont pas été formés et 143 ont reçu une formation de 45 minutes dans l'utilisation du CBCA. Après la formation, les participants avaient à évaluer un échantillon de déclaration destiné à évaluer l'efficacité de la formation. Les 24 évaluateurs ayant eu une précision de moins de 80% étaient éliminés de l'analyse. Chaque évaluateur avait à évaluer les douze transcriptions de déclarations d'un des deux groupes selon les mêmes critères que ceux utilisés par Landry et Brigham (1992) sauf pour l'évaluation globale de véracité qui se faisait à l'aide d'une échelle de 1 à 7. Quatre des critères du CBCA ont été éliminés de l'analyse car ils étaient spécifiques aux déclarations d'abus sexuels, soit les détails non-compris, le fait d'excuser l'abuseur, les attributions d'états psychologiques à l'abuseur et les caractéristiques spécifiques du délit.

L'analyse des résultats a démontré qu'aucun des critères était plus présent dans les déclarations réelles des caucasiens que dans celles des afro-américains. Au contraire, cinq des critères, soit la quantité de détails, le rappel de conversation, la référence à des états subjectifs, les corrections spontanées et la désapprobation de sa propre conduite, étaient plus souvent présents dans les déclarations réelles des afro-américains. Chez les caucasiens, les évaluateurs, formés ou non, étaient significativement moins efficaces qu'une classification aléatoire pour différencier les déclarations réelles des inventées. Aucune différence significative n'a été décelée chez les afro-américains. Pour ce qui est des critères individuels lorsque les déclarations réelles et inventées sont analysées ensemble, les déclarations des caucasiens exhibaient plus souvent les critères de structure logique et d'enchâssement contextuel que les déclarations des afro-américains. À part les critères de quantité de détails, de complications inattendues et de désapprobation de sa propre

conduite, tous les autres critères, au nombre de dix, étaient présents plus souvent dans les déclarations des afro-américains que dans celles des caucasiens.

Lorsque l'origine ethnique n'était pas considérée dans l'analyse, sept critères, soit la production non-structurée, les descriptions d'interactions, les complications inattendues, les détails inusités, les détails superflus, les corrections spontanées et l'aveu d'absence de souvenir, étaient présents plus souvent dans les déclarations réelles que dans les déclarations inventées. Par contre, cinq critères, soit la structure logique, l'enchâssement contextuel, les références à des incidents extérieurs, les références à des états subjectifs et la désapprobation de sa propre conduite, étaient présents plus souvent dans les déclarations inventées que dans celles réelles. Il est bon de noter que quatre des critères présents plus souvent dans les déclarations réelles et un de ceux plus souvent présents dans les déclarations inventées font partie de ceux qu'il fut suggéré d'éliminer ou redéfinir dû à leur fidélité inadéquate (Horowitz et al, 1997). Les analyses ont aussi démontré que le nombre de critères présents dans une déclaration n'était absolument pas relié à sa véracité, $r = -0.0001$, $p > 0.05$. En plus de n'offrir qu'un support très mitigé à la validité du CBCA chez les adultes, cette étude semble démontrer que la technique est grandement affectée par l'origine ethnique du témoin.

Pour finir, il est malgré tout important de souligner certaines limites de l'étude, la première étant reliée à la sélection des participants produisant des déclarations. Cette sélection ne s'est pas faite aléatoirement ou à l'aveugle. Elle était plutôt mue par un désir de trouver un échantillon caucasien équivalent, sur certains éléments, à l'échantillon afro-américain disponible. Il est donc possible que certains biais expérimentaux aient guidé cette sélection et éliminé ou transformé une certaine variabilité naturelle entre les déclarations. D'autres limites, mentionnées pour l'étude de Landry et Brigham (1992) s'appliquent aussi à l'étude de Ruby et Brigham (1998). Par exemple, le fait de limiter la durée des déclarations peut amener les participants à résumer l'incident au lieu de tout rapporter et donc à produire des déclarations incomplètes. Aussi, on peut soulever l'argument à l'effet qu'une formation de 45 minutes est insuffisante pour assurer une bonne compréhension de la technique et une fidélité adéquate durant l'évaluation ou par ailleurs, qu'un délai de trois minutes est insuffisant pour évaluer les 15 critères. Par contre, il se pourrait que les contrôles aient été plus sévères dans cette étude pour s'assurer que la formation était adéquate et bien assimilée. Finalement, comme dans l'étude de Landry et Brigham (1992), la variable véracité est confondue avec l'ordre de présentation, ce qui a pour conséquence que l'interviewer ne pouvait pas être aveugle quant à la véracité de la déclaration rapportée. Les résultats peuvent donc être dus à un simple effet de séquence ou à un biais de l'interviewer.

Porter et Yuille (1996) ont décidé d'évaluer la validité et l'efficacité du SVA et de trois autres techniques réputées pouvoir différencier les déclarations réelles de celles inventées, soit le reality monitoring de Johnson et Raye (1981), le programme de formation de Sapir (1987) et la diversité lexicale de Hollien (1990), dans une simulation de crime réaliste. Soixante participants ont été assignés de façon aléatoire à une des quatre conditions. Les participants dans une des trois premières conditions avaient comme instructions d'aller voler un billet de cent dollars à l'intérieur d'un des bureaux du département en évitant de se faire remarquer par le garde de sécurité et de revenir au laboratoire avec l'argent et la chemise dans laquelle l'argent était. Les participants dans la quatrième condition avaient comme instructions d'aller chercher une chemise ne contenant pas d'argent pour un professeur et de revenir au laboratoire. Une clé et les indications pour se rendre étaient données à chaque participant et il était précisé qu'ils n'avaient que 10 minutes pour effectuer leur tâche. En revenant au laboratoire, tous les participants étaient informés que la deuxième partie du contrat était d'évaluer l'efficacité des interrogateurs d'une compagnie à reconnaître qui était coupable du vol et qu'un bonus de cinq dollars leur serait remis s'ils réussissaient à convaincre l'interrogateur de la véracité de leur déclaration. Les participants dans la quatrième condition, celle de l'alibi véridique, avaient comme instructions de dire la vérité. Les autres participants étaient dans une des trois conditions suivantes: (1)-tromperie partielle- ces participants avaient à déclarer qu'ils avaient été chercher une chemise pour un professeur mais devaient nier savoir que de l'argent était dans le bureau, (2)-tromperie complète- ces participants avaient à produire une déclaration complètement fausse sous la forme d'un alibi plausible, ou (3)-confession véridique- ces participants avaient à faire une déclaration complètement véridique. Les entrevues étaient transcrites sur support informatique et les évaluateurs avaient reçu une formation de trois jours sur le SVA. La condition "tromperie complète" fut éliminée de l'analyse pour le critère "détails superficiels" à cause de la difficulté à définir ce critère dans ce cas. Huit des critères du CBCA, soit le rappel de conversation, l'attribution d'états mentaux à l'abuseur, le fait d'excuser l'abuseur, les caractéristiques spécifiques du délit, l'enchâssement contextuel, la description d'interactions, les détails non-compris et la désapprobation de sa propre conduite, n'ont pas été évalués car ils ne s'appliquaient pas au contexte de cette étude.

De tous les critères évalués, seul trois d'entre eux provenant tous du CBCA, soit la quantité de détails, la structure logique et l'aveu d'absence de souvenir, différenciaient les groupes de façon significative. Une analyse plus en profondeur a démontré que les déclarations provenant d'une des deux conditions véridiques, soit l'alibi véridique et la confession véridique, contenaient plus de détails et avaient des scores plus élevés pour le critère structure logique que les déclarations provenant d'une des deux conditions de tromperie,

soit la tromperie partielle et la tromperie complète. De plus, les déclarations de la condition "alibi véridique" contenaient plus d'aveu d'absence de souvenir que celles de la condition "tromperie partielle". Il est intéressant de se rappeler que Horowitz, Lamb, Esplin, Boychuk, Krispin et Reiter-Lavery (1997) suggéraient d'éliminer ou de redéfinir le critère "aveu d'absence de souvenir" à cause de sa fidélité inadéquate. Cette étude n'amène donc un support empirique qu'à trois critères du CBCA dans la population adulte, ou à seulement deux d'entre eux si l'on tient compte de la critique de Horowitz et al. (1997), et ce malgré l'effort particulier déployé pour augmenter le réalisme de la simulation et par conséquent la validité externe de l'étude. Certaines limites ont été soulignées par les auteurs. Par exemple, le fait d'évaluer des déclarations de suspects et non de victimes enlève l'élément de perte de contrôle considéré comme nécessaire par Steller (1989). Il est aussi important de se rappeler que huit critères du CBCA n'ont pas été inclus dans l'étude et qu'ils ne sont donc pas concernés par les conclusions de l'étude.

Vrij, Kneller et Mann (2000) dans une recherche analogue à celle de Joffe et Yuille (1992, dans Lamers-Winkelmann et Buffing, 1996) ont cherché à connaître l'effet d'informer des menteurs à propos du CBCA sur leur habileté à tromper des évaluateurs utilisant le CBCA. Un vidéo d'une durée approximative de 2 minutes et qui présentait le vol d'un sac de caméra y était utilisée comme événement-cible. Les 45 participants âgés de 19 à 46 ans étaient assignés aléatoirement à une des trois conditions suivantes: (1)-déclaration réelle- ces participants visionnaient le vidéo et étaient informés que l'expérience portait sur la mémoire, (2)-tromperie non-informée- ces participants avaient comme instruction de tromper l'interviewer en lui donnant une description du vidéo si convaincante qu'il croirait qu'ils l'avaient réellement vu. Une brève description écrite du vidéo leur était présentée et une certaine période de temps leur était allouée pour qu'ils puissent construire leur histoire. (3)-tromperie informée- ces participants recevaient les mêmes instructions que ceux du deuxième groupe mais, en plus, recevaient des conseils basés sur les critères du CBCA destinés à les aider à tromper l'interviewer. Après cette première étape, tous les participants furent interviewés par un expérimentateur aveugle quant au groupe du participant selon une procédure standard. Les entrevues étaient transcrites par écrit et évaluées par deux évaluateurs indépendants formés dans l'utilisation du CBCA. Sept des 19 critères, soit le rappel de conversation, les détails non-compris, les références à des incidents extérieurs, les références à des états subjectifs, la désapprobation de sa propre conduite, le fait d'excuser l'abuseur et les caractéristiques spécifiques du délit, ont été ignorés dans cette étude car ils étaient considérés inappropriés au type d'événement présenté.

Comme les auteurs l'avaient prédit, le score total du CBCA était significativement plus élevé pour les déclarations de la condition "tromperie informée" que pour la condition

"tromperie non-informée". Les menteurs informés incluait plus de détails et plus de doutes à propos de leur déclaration et faisaient plus d'attributions à l'état mental de l'accusé que les menteurs non-informés. Aucune différence significative au niveau des scores totaux ne fut observée entre les déclarations réelles et les fausses déclarations informées tandis que seul le critère détails inusités, plus fréquent dans les déclarations réelles, différenciait de façon significative ces deux types de déclarations. Quatre critères individuels, soit la quantité de détails, les détails inusités, l'attribution d'états mentaux et les corrections spontanées, différenciaient les déclarations réelles des fausses déclarations non-informées. Ces différences étaient conformes aux prévisions tirées du CBCA.

L'évaluation du taux de précision du CBCA a été effectuée de deux façons. Premièrement une analyse discriminante a été effectuée en utilisant les scores totaux des déclarations. En tout, 53% des déclarations réelles, 47% des fausses déclarations informées et 13% des fausses déclarations non-informées ont été classifiées comme des déclarations réelles. Cela indique que malgré que l'analyse ait pu classifier avec un certain succès les fausses déclarations non-informées, elle fut incapable de différencier les déclarations réelles des fausses déclarations informées. Deuxièmement, les expérimentateurs ont fait appel à un expert dans l'utilisation du CBCA qui ignorait les hypothèses et manipulations expérimentales spécifiques à l'étude dans le but de voir s'il pourrait différencier les déclarations réelles et fausses avec plus de succès. L'expert a réussi à classifier comme valides 80% des déclarations réelles, 40% des fausses déclarations non-informées et 73% des fausses déclarations informées. En d'autres mots, la majorité des menteurs informés ont réussi à tromper la classification de l'expert. Après que la classification soit effectuée, les auteurs ont informé l'expert à propos de la nature de l'expérience et de l'information donnée à certains participants pour les aider à tromper les évaluateurs et lui ont demandé de réviser sa classification en connaissance de cause. Dans cette nouvelle classification, 73% des déclarations réelles, 40% des fausses déclarations non-informées et 60% des fausses déclarations informées ont été classifiées comme valides. Donc, même après avoir été informé de la nature de l'expérience, la majorité des menteurs informés ont réussi à tromper la classification de l'expert.

Cette étude semble démontrer que des personnes informées par rapport aux critères du CBCA peuvent très facilement tromper une classification faite à partir du CBCA, que cette classification soit faite à partir des scores totaux, de l'analyse par un expert ou de critères spécifiques. Il est assez surprenant qu'un tel effet soit démontré à l'aide d'un stimulus unique car un tel type de stimulus par opposition à des événements personnels diminue de beaucoup la variabilité intra-groupe et facilite la reconnaissance des écarts dans les fausses déclarations par rapport aux déclarations réelles. De plus, il est impossible pour les

évaluateurs et les interviewers d'être totalement aveugle quant à la condition à laquelle le participant est assigné car un participant mentionnant un ou plusieurs détails en désaccord avec la majorité des récits, comme par exemple une veste qui est bleu au lieu d'être rouge, sera probablement reconnu comme menteur. Si le participant veut éviter cette possibilité, il peut limiter son récit à certains éléments et éviter d'inventer des événements risqués, comme des détails inusités non mentionnés dans la description. Ce problème peut par conséquent amener les déclarations réelles et fausses à être anormalement différentes l'une de l'autre et donc améliorer l'efficacité des techniques basées sur l'analyse du contenu. Dans cette étude, seulement quatre critères sur douze, soit la quantité de détails, les détails inusités, l'attribution d'états mentaux et les corrections spontanées, ont différencié significativement les déclarations réelles des fausses déclarations non-informées. On doit se rappeler que deux de ces quatre critères, soit les détails inusités et les corrections spontanées, sont considérés comme ayant une fidélité inadéquate par Horowitz et al. (1997).

Critiques sur les recherches expérimentales

Lorsque comparées avec les recherches effectuées avec des cas réels, les recherches effectuées avec des cas réels nous amènent un nouveau regard sur l'utilisation du SVA comme procédure d'évaluation de la véracité. Pour ce qui est de la fidélité de la mesure, aucune étude aussi poussée que celles de Anson, Golding et Gully (1993) et de Horowitz et al. (1997) n'a été réalisée. Des analyses de base, comme la proportion d'accord et des corrélations de Pearson, ont été effectuées dans la plupart des études dans le but de s'assurer d'avoir un minimum de qualité dans l'évaluation des critères mais rien de plus. Les résultats de ces analyses rejoignent ceux des études quasi-expérimentales en démontrant une fidélité de modérée à élevée. Il ne faut pas oublier qu'aucune analyse correctrice ne fut utilisée pour tenir compte de l'effet de la chance. Une des raisons qui a probablement amené les chercheurs à ne pas faire beaucoup d'études expérimentales sur la fidélité est que les paradigmes quasi-expérimentaux utilisant des cas réels sont beaucoup plus adaptés à ce type de recherche étant donné que tous les critères peuvent être évalués et qu'il est moins nécessaire de s'assurer de la validité des déclarations lors d'études de fidélité que lors d'études de validité.

Au niveau de la validité, les résultats sont beaucoup moins probants que ceux des études quasi-expérimentales. Aucune de ces études n'a démontré un effet aussi spectaculaire que celui de Esplin, Boychuk et Raskin (1988). La validité et l'efficacité du score global du CBCA, lorsque analysées, sont remises en question par plus d'une étude, comme par exemple celles de Landry et Brigham (1992), de Ruby et Brigham (1998) et de Porter et Yuille (1996), malgré le support qu'elles reçoivent de certaines autres comme celle de

Yuille (1989). Cela peut probablement être expliqué par le fait que, dans la plupart des études analysant les critères individuels, rarement plus de la moitié des critères sont supportés par les résultats. En effet pour ce qui est des critères individuels, le portrait est assez complexe. Entre 3 et 10 critères sont supportés par les différentes études. Si on élimine les cinq critères dont la fidélité est considérée inadéquate d'après, entre autres les recherches de Horowitz et al. (1997), le maximum tombe à six critères. Cela diminue de beaucoup l'apparente validité démontrée par certaines études. De plus, certains critères comme la structure logique étaient parfois présents plus souvent dans les déclarations inventées que dans les déclarations réelles.

Si l'on en croit ces résultats, le CBCA n'est pas un ensemble de critères très unifié. De plus, il est intéressant de noter que, tant pour les études quasi-expérimentales que pour les études expérimentales, les résultats les plus probants, soit ceux de Esplin, Boychuk et Raskin (1988) et de Yuille (1989) ont été démontrés par les auteurs du SVA. Malgré qu'il est facile d'y voir l'influence d'un biais expérimental, il est aussi possible que cette différence soit due à une différence dans la façon d'interpréter les critères du CBCA. Comme certains critères manquent de précision, il est possible que les auteurs opérationnalisent ces critères d'une façon particulière. Vrij, Kneller et Mann (2000) apportent un certain appui à cette hypothèse en démontrant qu'en contexte expérimental la classification des experts en SVA arrive à un résultat différent d'une classification utilisant le score total du CBCA. Peu importe l'explication, la seule façon d'y voir clair serait d'opérationnaliser explicitement chaque critère. Il faut aussi souligner certaines limites aux études expérimentales. Premièrement, à cause de contraintes méthodologiques, les études expérimentales peuvent rarement analyser l'efficacité de tous les critères. Certains des critères, comme les caractéristiques spécifiques du délit s'appliquent difficilement à des souvenirs ne faisant pas référence à des allégations sexuelles. Par conséquent, entre 7 et 3 critères étaient éliminés dans les études recensées. Aussi, contrairement aux études quasi-expérimentales, une partie significative des études expérimentales récentes utilisent une population adulte. Cette différence pourrait partiellement expliquer l'apparente baisse d'efficacité dans les études expérimentales.

Le fait que le CBCA est influencé par l'âge a aussi été confirmé à l'intérieur d'une étude expérimental (Yuille, 1989). De plus, deux autres facteurs influençant le CBCA ont été démontrés. Premièrement, il semble que lorsqu'un témoin a une connaissance, même partielle, des critères utilisés pour évaluer la validité de sa déclaration, il peut facilement tromper la classification de l'évaluateur. Cet effet a été remarqué autant chez les enfants (Yuille, 1989) que chez les adultes (Vrij, Kneller et Mann, 2000). Finalement, l'étude de Ruby et Brigham (1998) semble démontrer que l'origine ethnique du témoin a un effet

majeur sur le CBCA mais que cet effet est très variable d'un critère à l'autre et difficile à prédire. Ces nouvelles données complexifient de nouveau la tâche de normalisation qui sera nécessaire pour assurer la validité de la procédure avec différentes populations.

Dans la section "Critiques sur les recherches quasi-expérimentales", le problème de confusion entre crédibilité et validité fut abordé. Il fut entre autres déduit que si le SVA évalue principalement la validité, son efficacité sera plus grande dans les études expérimentales, tandis que si elle évalue principalement la crédibilité, et non la validité, son efficacité sera plus grande dans les expériences utilisant des cas réels. Pour ce qui est des études recensées, il semble que le SVA discriminait en général avec plus d'efficacité les déclarations considérées inventées des déclarations considérées réelles dans les études quasi-expérimentales que dans les études expérimentales. Cela supporterait l'hypothèse que le SVA évalue, en partie du moins, la crédibilité des déclarations. Par contre, il est encore trop tôt pour se prononcer car d'autres facteurs peuvent expliquer ces différences. Le type de population étudié, soit des enfants ou des adultes, en est un exemple.

Finalement, il est bon de noter l'importance du choix des événements-cibles lors de l'élaboration de ces études. Certaines des études expérimentales recensées ont utilisé un vidéo unique ou un événement simulé tandis que les autres ont utilisé des événements choisis par les participants. Lorsqu'un événement unique est utilisé comme base pour les déclarations, ces dernières sont beaucoup plus uniformes que lorsque chaque déclaration est basée sur un événement différent. Cette uniformité produit une variabilité anormalement réduite entre les déclarations, ce qui facilite grandement la détection de l'effet des manipulations mais augmente la possibilité de détecter des artefacts expérimentaux, c'est-à-dire des effets particuliers à l'événement-cible qui ne pourraient être généralisés à d'autres types d'événements. De telles conditions sont souhaitables lorsqu'un domaine de recherche en est à ses débuts car la sensibilité accrue de ces paradigmes permet de remarquer des effets qui seraient camouflés par une trop grande variabilité intra-groupe. Par contre, lorsque l'objectif est d'appliquer les résultats en milieu naturel, comme c'est le cas pour la recherche sur le SVA, il faut interpréter ces recherches avec prudence car un effet statistique significatif dans un paradigme à variabilité réduite peut être de grandeur insignifiante dans un milieu à grande variabilité. L'efficacité du CBCA et de chacun de ces critères dépend donc grandement du paradigme dans lequel il est évaluée. Par conséquent, un critère valide et efficace dans un paradigme à variabilité réduite peu donner une fausse impression d'efficacité en milieu naturel.

Analyses critiques et revues de littérature

Plusieurs auteurs ont publié des analyses critiques ou des revues de littérature sur le SVA, soit pour alimenter la controverse ou souligner un point négligé dans la littérature, soit pour essayer de faire le point sur les connaissances actuelles à ce propos. Avant de faire le point, dans ce présent ouvrage, sur les recherches présentées, un court aperçu des conclusions d'auteurs importants sera présenté dans les prochaines lignes. Pour éviter la répétition, les critiques qui ont déjà été présentées dans les sections précédentes ne seront pas présentées de nouveau. L'objectif principal de cette revue étant de pouvoir apprécier l'opinion actuel sur le sujet dans une perspective historique et non de faire une méta-analyse, les arguments présentés par chacun seront résumés. Le lecteur intéressé à approfondir ses connaissances sur le sujet est invité à consulter les publications originales pour apprécier la complexité du débat.

En 1994, Doris cherche à faire le point sur la recherche sur le CBCA. Il considère la recherche sur les critères à utiliser pour évaluer la validité de déclarations comme un processus jeune et croit qu'un consensus ne pourra être atteint que par la recherche empirique. Il soulève aussi la question de la transférabilité d'un tel processus à un système de justice nord-américain étant donné les différences de procédures. Il soutient que, par nécessité, on peut utiliser la procédure sous sa forme présente et se fier sur le jugement clinique des experts l'utilisant mais qu'à long terme on se doit d'en vérifier empiriquement la valeur scientifique.

La même année, Bradford (1994) présente le SVA comme une procédure prometteuse. Malgré qu'il croit qu'il est trop tôt pour l'utiliser comme preuve légale à l'intérieur d'un procès, il pense que le SVA peut être utile dans les étapes préliminaires pour guider les décisions des différentes instances impliquées. Il indique que la procédure ne devrait pas être utilisée avec de courtes déclarations, ce qui remet en question son utilisation avec de jeunes enfants, et souligne le fait qu'aucune étude n'a établie de normes tenant compte de l'âge de l'enfant, de son sexe ou de son expérience, ce qui amène comme conséquence que l'évaluation est probablement très influencée par les connaissances et l'habileté de l'évaluateur. Cet article a été suivi de commentaires par deux auteurs différents.

En réponse à l'article de Bradford (1994), Jones (1994) souligne que les résultats de la recherche sont équivoques et que la prudence est de mise face à une technique qui pourrait n'être qu'une fausse promesse d'une solution désirée par tous. Contrairement à Bradford, Jones espère que le SVA ne sera utilisé par aucune des différentes instance juridiques pour guider leurs décisions, du moins jusqu'à ce que des recherches futures en démontrent l'efficacité de façon claire et sans ambiguïté.

Davies (1994) répond aussi à l'article de Bradford par la prudence. Compte tenu des lacunes qu'il observe dans la recherche qui existe à ce moment, il considère important que certaines questions soient approfondies. Premièrement, la fidélité inter-juges doit être établie et des règles claires concernant la façon d'appliquer les critères doivent être développées. Deuxièmement, la question de la pondération de chaque critère doit être résolue. Davies appuie Bradford (1994) dans sa prudence face à l'utilisation du SVA chez les jeunes enfants et ajoute que l'origine ethnique est aussi un facteur à considérer. De plus, il souligne l'interdépendance du contenu de la déclaration et de la technique d'entrevue et donc l'importance d'utiliser le SVA conjointement avec une procédure d'entrevue rigoureusement standardisée. Il conclut en notant que le SVA a un long chemin à faire avant de devenir une science applicable mais considère que l'ensemble des critères peut quand même être vu comme une sorte de grille intuitive pouvant aider à la prise de décision.

En 1997, Ruby et Brigham publient une revue critique de la recherche sur le CBCA. Selon leurs conclusions, le critère "quantité de détails" semble le plus utile. Cette revue considère comme prometteurs les dix critères suivants puisqu'ils furent significatifs dans quatre des études analysées: structure logique, production non-structurée, enchâssement contextuel, description d'interactions, rappel de conversations, complications inattendues, détails inusités, référence à des états subjectifs, corrections spontanées et aveu d'absence de souvenir. Malgré cela, ils soulignent que dans deux études, le critère "structure logique" était plus fréquent pour les fausses déclarations. Le critère désapprobation de son propre comportement n'a reçu aucun support empirique dans les neuf études l'ayant analysé tandis que trois autres critères, soit l'attribution d'états mentaux, les doutes à propos de sa déclaration et le fait d'excuser l'abuseur, n'ont reçu du support que dans une seule étude. Les auteurs concluent en notant que les résultats suggèrent que le CBCA est loin d'être un instrument unitaire et cohérent et qu'il n'est pas prêt à être utilisé à l'intérieur d'une expertise pour corroborer la déclaration d'un témoin.

Tout récemment, Casoni (1999) a soumis le CBCA à un examen critique de sa validité de construit à la lumière des connaissances théoriques et empiriques actuelles dans le but d'en évaluer la pertinence au point de vue juridique et clinique. L'auteure résume son constat en ces mots:

"En résumé, l'analyse critique de ces listes de critères permet de mettre en lumière le fait qu'un grand nombre de propositions méta-théoriques ayant donné naissance à certains critères ont des assises conceptuelles lacunaires comme le soulignent spécifiquement Bekerian et Dennett (1992). De même, comme il en a été longuement discuté, l'absence de balises pour permettre la pondération des critères, la sur-représentation de critères équivalents sans justification valable et enfin l'absence de lignes interprétatives basées sur des justifications théoriques, cliniques ou empiriques démontrées diminuent beaucoup la

pertinence de recourir à de semblables grilles. [...] Pour le moment, [...] la pertinence de recourir à des outils conçus à la manière des grilles de validation proposées par Undeutsch au début des années 1950 est révolue dans le domaine de l'évaluation des allégations de sévices sexuels." (Casoni, 2000)

Vue d'ensemble sur la recherche

Après avoir évalué globalement l'ensemble de la littérature sur le SVA, dont celle sur le CBCA, il peut sembler difficile de faire le point sur la situation et la direction à prendre. Malgré tout, certains éléments ressortent de l'analyse. Entre autres, lorsque les études expérimentales sont comparées aux études quasi-expérimentales, elles semblent contenir moins de problèmes méthodologiques majeurs. Ces forces méthodologiques ont malheureusement comme corrélat une certaine perte en validité externe. Ce problème peut quand même être compensée par l'utilisation d'opérations convergentes. S'il est démontré qu'un critère est efficace dans plusieurs situations différentes, cela assure une certaine possibilité de généraliser des résultats.

Avant de tirer des conclusions sur la validité et l'applicabilité du SVA, il est important de clarifier les circonstances dans lesquelles il sera utilisé. Certains désaccords entre les différents auteurs semblent venir d'une différence dans la façon dont ils définissent le rôle du SVA en contexte légal. Cette confusion vient probablement du fait que le rôle du SVA a effectivement changé depuis ses origines. À l'origine, le Statement Reality Analysis, prédécesseur du SVA, était nécessaire à chaque fois que le témoignage de l'enfant était la seule évidence du crime. La procédure avait donc comme objectif principal d'assurer que la valeur légale de la déclaration de l'enfant soit considérée équivalente à celle d'un adulte. Une telle procédure n'avait pas besoin d'être opérationnalisée de façon stricte et de démontrer empiriquement et sans ambiguïté sa validité et son efficacité avant de pouvoir être utilisée puisqu'elle s'apparente plus à une formalité juridique qu'à un instrument psychométrique. Par contre, du moment que le SVA est utilisé comme évidence additionnelle de la véracité d'un témoignage déjà accepté par la cour, son efficacité et sa validité doivent être évaluées de la même façon que pour tout autre test psychométrique si l'on veut justifier sa valeur en tant que preuve.

Cette précision peut aider à mieux comprendre, par exemple, la raison pour laquelle certains auteurs comme Tully (1998) ne reconnaissent pas la légitimité des études recommandant d'éliminer ou de redéfinir certains critères. Si le SVA est défini comme une procédure légale, il est inutile en Amérique du Nord et dans la plupart des systèmes juridiques actuels parce que les témoignages d'enfants sont maintenant admissibles comme évidences et que diverses procédures, comme le témoignage vidéo ou devant miroir unidirectionnel, ont été créées pour adapter le système juridique à ce type de témoin. La

recherche actuelle devient donc inutile. Par contre, si le SVA est considéré comme candidat potentiel au titre d'instrument pouvant différencier les témoignages véridiques des témoignages non-véridiques, la même rigueur scientifique se doit d'être appliquée qu'avec les autres instruments psychométriques avant qu'il puisse servir d'évidence additionnelle. En ce sens, le présent auteur rejoint Porter et Yuille (1996), Casoni (1999) et Vrij, Kneller et Mann (2000) en soutenant que la présente littérature ne permet pas de considérer le SVA, dans sa forme actuelle, comme une procédure apte à être utilisée pour évaluer la véracité d'un témoignage et en recommandant de ne pas l'utiliser en contexte légal.

Comme il fut mentionné à la section "Critiques sur les recherche expérimentales", le CBCA est loin d'être un ensemble uniforme de critères d'importance égale. Il ressemble plus à un amalgame de critères de fréquence, de valeur, d'efficacité, de validité et de fidélité différentes dont certains sont probablement des fardeaux n'aidant pas à catégoriser les déclarations selon leur véracité. De plus, l'efficacité du score total du CBCA comme indicateur de véracité est équivoque. On peut se demander où est l'intérêt de conserver un tel carcan, particulièrement si l'on considère que la procédure ne devrait pas être utilisée dans sa forme actuelle et que certains des critères ont besoin d'être redéfinis ou éliminés. Il semble qu'une meilleure approche pour maximiser le potentiel heuristique de l'hypothèse d'Undeutsch serait une approche évaluant les "ingrédients actifs" de la procédure. Une telle approche permettrait d'évaluer l'apport de chaque critère et de les opérationnaliser de différentes façons, et ce pour optimiser leur efficacité et par conséquent, l'efficacité globale de l'évaluation. Cette approche permettrait même d'incorporer de nouveaux critères potentiellement efficaces. Déjà, Porter et Yuille (1996) ont fait un pas dans cette direction en évaluant dans la même étude, les critères de différentes techniques.

Un des faits intéressants à noter est que dans les études expérimentales sur le SVA, seul les déclarations volontairement inventées sont étudiées. Dans ces conditions, il est difficile de savoir si les différences de contenu perçues sont causées par des facteurs motivationnels comme l'intention de tromper ou par des facteurs cognitifs comme la source de l'événement, soit un événement imaginé ou perçu, car la véracité et l'origine des déclarations sont confondues. Cette différence peut paraître mineure mais si les effets détectés sont causés par les différences d'intentions, soit tromper ou dire la vérité, les déclarations non-véridiques provenant de témoins croyant en l'authenticité de leur récit seront catégorisées comme valides. Cela voudrait dire que si l'origine de la déclaration est un souvenir faussement attribué à une source externe, par opposition à une source interne comme un rêve, cette déclaration sera considérée comme valide. Cette différence a donc des implications importantes au niveau légal malgré qu'elle n'a pas réellement été étudiée jusqu'à présent.

L'évaluation de l'influence des facteurs cognitifs et motivationnels sur les différents critères nécessite un certain cadre théorique. Le "source monitoring", présenté par Johnson, Hashtroudi et Lindsay (1993), est un paradigme théorique idéal pour ce genre d'étude. C'est un cadre théorique basé sur la nature reconstructive de la mémoire qui permet, par exemple, de comprendre comment le souvenir d'un événement qui n'est jamais arrivé peut exister. Le "source monitoring" soutient l'hypothèse que l'origine du souvenir n'est pas directement retrouvée comme une étiquette spécifiant la source du souvenir mais plutôt que les souvenirs activés sont évalués et attribués à des sources particulières par un processus décisionnel accompli durant le rappel (Johnson, Hashtroudi et Lindsay, 1993). L'ensemble des processus impliqués dans les attributions à propos de l'origine de souvenirs, de connaissances et de croyances est appelé "source monitoring". Deux types de souvenirs sont généralement reconnus: les souvenirs d'événements générés intérieurement, comme des rêves et des événements imaginés, et les souvenirs d'événements dérivés de l'extérieur, soit les événements perçus. Une erreur de "source monitoring" arrive lorsqu'une source est faussement attribuée à un souvenir d'une autre source. Selon ce cadre théorique, un souvenir d'un événement qui n'est jamais arrivé serait dû au souvenir d'un événement interne faussement attribué à une source externe. Comme pour les études présentées précédemment, l'étude des souvenirs faussement attribués à une source externe peut se faire dans un paradigme quasi-expérimental ou expérimental avec des conséquences méthodologiques presque identiques.

Présente étude

L'objectif à long terme de la présente étude étant de pouvoir contribuer à la recherche sur le développement d'une procédure pouvant différencier les déclarations véridiques des déclarations non-véridiques, une approche évaluant les "ingrédients actifs" du SVA sera privilégiée à une approche évaluant le CBCA comme entité globale. Cette distinction est très importante, surtout si l'on prend en considération l'hypothèse, soulignée dans la section "Critiques sur les recherches expérimentales", selon laquelle il est possible que les auteurs du SVA utilisent certaines règles implicites non-mentionnées dans la littérature pour préciser les différents critères. Évaluer l'efficacité d'un expert à distinguer les types de déclarations, ce qui revient à évaluer sa formation et/ou ses croyances, pourrait amener un certain appui à l'utilisation du SVA en contexte légal, en supposant que les résultats soient probants, malgré la difficulté de savoir si les règles implicites utilisées par l'expert correspondent aux règles utilisées par la majorité des experts. Par contre, ce type d'étude pourrait difficilement amener à optimiser l'efficacité de la procédure étant donné que la

contribution individuelle de chaque critère y est difficile à évaluer. Les définitions exhaustives des différents critères du CBCA telles que présentées dans la littérature seront utilisées comme point de départ. Le lecteur qui voudrait s'y référer est invité à consulter l'annexe B. Lorsque l'opérationnalisation nécessitera l'ajout de règles additionnelles, elles seront précisées dans la méthodologie pour assurer que le nombre de règles implicites soit réduit au maximum.

Un deuxième objectif de la présente recherche est de débiter l'étude de l'influence de la source de la déclaration sur les différents critères du CBCA, et ce dans le but de comprendre si les effets démontrés sont causés par des facteurs motivationnels ou cognitifs. Pour éviter des problèmes méthodologiques majeurs et pour assurer un plus grand contrôle des variables, la recherche utilisera une approche expérimentale au lieu de cas réels. La réaction des différents critères sera évaluée face à des déclarations véridiques et des déclarations trompeuses à propos d'événements de source externe et internes. Des souvenirs de rêves seront utilisés comme événements de source interne. Un élément qui justifie cette décision est que d'après des recherches utilisant le source monitoring, les rêves seraient des cibles faciles pour des erreurs d'attributions vu l'absence d'effort cognitif, qui est un critère d'attribution interne, durant ce processus (Johnson, Kahan et Raye, 1984). Les événements-cibles seront choisis par les participants, selon des critères préétablis, et un temps de préparation de quelques jours leur sera attribué. Cela diminuera les chances de découvrir des effets simplement dus au manque de préparation du participant ou de gonfler artificiellement l'efficacité des critères par l'utilisation d'une méthodologie à variabilité intra-groupe réduite comme c'est le cas avec les méthodologies à événement-cible unique. Cela rendra aussi les résultats plus facilement applicables à un contexte légal.

Étant donné la difficulté de généraliser les résultats d'un groupe d'âge à l'autre, l'étude se fera sur un seul groupe d'âge, soit les adultes. Cette décision est motivée par plus d'un facteur. Premièrement, malgré que le SVA fut créé dans le but d'évaluer des déclarations d'enfants, certains experts l'utilisent maintenant avec des déclarations d'adultes et/ou ne contenant pas d'allégations sexuelles (Laurence, 1999). De plus, ce type d'utilisation est appuyé par l'opinion d'auteurs du SVA comme Steller et Köhnken (1989) et Porter et Yuille (1996) même si les principales données sur la validité et la fidélité de la procédure proviennent d'études avec des enfants. Malgré tout, une certaine tendance à évaluer la validité du SVA chez les adultes peut être observée à partir des années 1990 chez des auteurs tels Ruby et Brigham (1998) et Porter et Yuille (1996). La présente recherche s'inscrit donc dans cette tendance. C'est aussi, à la connaissance de l'auteur la première étude utilisant une variabilité réaliste ET un temps de préparation réaliste chez les adultes.

Les deux autres études utilisant un événement-cible différent pour chaque déclaration et un temps de préparation réaliste (Landry et Brigham, 1992; Ruby et Brigham, 1998) ont utilisé une procédure de sélection des déclarations réduisant la variabilité. De plus, la variable véracité y était confondue avec l'ordre de présentation et l'influence de l'interviewer. Il est espéré que des études similaires à celle-ci et différenciant tromperie et source seront effectuées avec des groupes d'âge différents. Finalement, le facteur véracité/tromperie sera une variable intra-sujet, ce qui permettra de s'assurer que les effets démontrés ne soient pas dus à une différence entre les sujets. Cela aura aussi comme avantage d'augmenter la puissance statistique des analyses qui seront effectuées.

Il est important de bien saisir le paradigme utilisé pour interpréter correctement les résultats qui seront obtenus. En comparant des déclarations réelles et inventées à propos d'événements et de rêves, deux facteurs principaux peuvent influencer les résultats. Ces facteurs sont l'intention de tromper, associée à la véracité du récit, et la source du souvenir, qui est externe pour les événements perçus et interne pour les événements inventés et les rêves. Par contre, un troisième facteur entre en jeu dû au fait que les souvenirs de rêves et les souvenirs d'événements ne diffèrent pas seulement au niveau de leur source d'origine mais aussi au niveau de leur statut. De façon plus explicite, il est possible que les participants aient des croyances, probablement justifiées, donnant un statut particulier aux souvenirs de rêves, comme par exemple les considérer plus flous, moins logiques et/ou plus émotionnels, ce qui peut amener certains critères à différencier les récits de rêves des récits d'événements sans que l'effet soit directement relié à la source du souvenir. Il est donc important de ne pas confondre un effet dû au statut du souvenir et un effet dû à sa source.

Malgré tout, il est quand même possible de différencier les effets des différents facteurs, principalement en comparant les conditions qui ne diffèrent que par rapport à un seul de ces facteurs. Par exemple, si une différence significative est démontrée entre les récits de rêve réels et les récits de rêve inventés, elle ne peut pas être due à l'effet de la source car les deux types de récits proviennent de souvenirs de source interne. De plus, la différence ne peut pas être due à l'effet du statut particulier attribué aux rêves car les deux types de récits sont à propos de rêves. Une telle différence peut donc être attribuée à l'effet de la tromperie. Par conséquent, un effet significatif pour la variable véracité en l'absence d'interaction démontre l'influence de la tromperie car cela suppose que les récits inventés et réels sont différents, même pour les rêves. De même, une différence entre les récits d'événement inventés et les récits de rêve inventés sera attribuée à l'effet du statut particulier attribué aux rêves étant donné que la véracité et la source des récits sont des variables fixes dans ce cas. Par conséquent, un effet significatif pour la variable "type de récit" (événement

ou rêve) en l'absence d'interaction démontre l'influence du statut attribué aux rêves car cela suppose que les récits de rêves sont différents des récits d'événements, même lorsqu'ils sont inventés. Finalement, un effet peut être attribué à la source du souvenir si et seulement si une différence significative est démontrée entre les récits d'événements réels et les récits de rêves réels en l'absence de différence significative équivalente entre les récits d'événement inventés et les récits de rêve inventés. Un tel effet prendrait la forme d'une interaction vérité X type de récit.

Comme l'objectif de la recherche est d'évaluer la validité et l'efficacité des critères du CBCA, il est difficile de prévoir lesquels seront valides. Si on en croit l'opinion des auteurs du SVA, la plupart des critères devraient probablement être plus fréquents dans les déclarations réelles et dans les déclarations de source externe. Par contre, la revue de littérature qui vient d'être effectuée démontre un tableau beaucoup plus ambiguë. Il est donc probable que seulement une partie des critères s'avèreront efficaces avec cette méthodologie.

MÉTHODOLOGIE

Plan d'expérience

L'expérience consiste en deux conditions distinctes. Les participants assignés à la première condition ont à raconter deux événements, soit un réel et un inventé, tandis que ceux assignés à la deuxième condition ont à raconter deux rêves, soit un réel et un inventé. Aucun participant n'a à raconter deux récits inventés ou deux récits réels mais cette information ne leur est jamais spécifiée avant le debriefing et ce, pour s'assurer que leur opinion par rapport à la crédibilité du premier récit raconté n'affecte pas outre mesure leur performance au deuxième récit. L'ordre dans lequel les deux récits doivent être présentés est contrebalancé pour éviter de confondre l'effet de la véracité du récit avec l'effet de l'ordre de présentation. Les participants sont assignés de façon aléatoire entre les deux conditions et l'expérimentateur principal reste aveugle quant à la véracité de tous les récits jusqu'au moment de l'entrée des données. Étant donné que des francophones et des anglophones participent à la recherche, l'appartenance linguistique est contrôlée de manière à avoir un nombre équivalent de participants de même appartenance assigné à chaque condition.

Instructions concernant les récits

Certaines instructions régissent le choix des récits par les participants. Pour éviter d'inclure des souvenirs ayant été hautement sujet à la distorsion et la réinterprétation, les souvenirs ayant été ou étant associés à une démarche psychothérapeutique et les souvenirs d'enfance ne peuvent être choisis. Pour s'assurer que les récits racontés soient comparables entre eux et qu'ils aient une certaine validité écologique, il est aussi précisé de choisir des événements (ou des rêves) qui se déroulent sur une courte période de temps et dont le contenu soit marquant, mémorable et émotionnellement significatif. Pour les récits inventés, il est aussi précisé de créer un récit crédible. De plus, pour motiver les participants à être crédibles tout au long de l'expérience, il est stipulé qu'un dédommagement de cinq dollars leur est attribué et qu'une compensation additionnelle de cinq dollars est ajoutée à moins que, durant un des deux interrogatoires, l'interviewer soit totalement convaincu que le récit raconté est inventé. Finalement, pour que les rêves soient plus comparables aux événements, il est demandé de choisir des rêves comportant peu ou pas d'éléments irréels. Ces instructions sont envoyées par courrier postal aux participants quelques jours avant leur rendez-vous pour leur permettre un minimum de deux à trois

jours pour lire la lettre et se préparer à l'entrevue. Le lecteur intéressé à connaître les instructions précises est encouragé à consulter l'annexe E.

Participants / Recrutement

Trente-six étudiant(e)s universitaires (quatorze francophones et vingt-deux anglophones) ont participé à l'expérience. Ils ont été recrutés dans des classes de psychologie de l'université Concordia et de l'université de Montréal. On leur a expliqué que le but de l'expérience était d'évaluer une technique utilisée en psychologie légale pour différencier les vrais témoignages des faux témoignages. Comme il fut vérifié durant chaque debriefing, tous les participants étaient naïfs et ignorants quant à la technique utilisée pour évaluer les récits.

Procédure

Entrevue téléphonique

Après avoir été recruté en classe, chaque participant est appelé par téléphone par l'expérimentateur principal (voir annexe D pour les protocoles utilisés). Durant cette entrevue, il est expliqué au participant que l'expérience consiste en une entrevue durant laquelle il lui sera demandé de raconter deux événements ou deux rêves, soit réels ou inventés et que cinq dollars de dédommagement sont prévus pour sa participation. On lui spécifie aussi qu'il devra essayer d'être le plus crédible possible lorsqu'il racontera ses deux récits, le but étant de convaincre l'interviewer de la véracité de chaque récit, et que cinq dollars additionnels lui seront donnés si, durant l'entrevue, l'interviewer qui est aussi l'expérimentateur principal trouve les deux histoires crédibles.

Si la personne accepte de participer, il lui est dit qu'une lettre incluant les instructions à suivre pour choisir ses deux récits lui sera envoyée par Dr Jean-Roch Laurence, chercheur en charge du projet et un rendez-vous est pris pour l'entrevue. Un délai est prévu, durant la prise de rendez-vous, pour s'assurer que le participant ait un minimum de deux à trois jours pour lire la lettre et se préparer à l'entrevue. La lettre envoyée au participant (annexe E) contient les instructions de base données durant l'entrevue téléphonique ainsi que les instructions spécifiques concernant le choix des récits. De plus, un formulaire, qui doit être apporté à l'entrevue, est envoyé avec la lettre. Ce formulaire consiste essentiellement en une page où le participant doit résumer, pour lui-même, les éléments importants de ses récits.

La fonction de ce formulaire est de s'assurer que le participant s'est réellement préparé et sert à vérifier, après l'entrée des données, que les instructions ont bien été suivies.

Entrevue

Les entrevues individuelles, d'une durée d'environ 40 minutes, peuvent être divisées en trois parties, soit l'introduction, l'interrogatoire et le debriefing. Durant la première partie, l'expérimentateur s'assure que les instructions de bases et les conditions de participation sont bien comprises par le participant. Il y précise entre autres que si, lors de l'interrogatoire, il est totalement convaincu que le récit raconté est inventé, l'interrogatoire sera arrêté et seulement le dédommagement de base de cinq dollars sera donné. Par contre, cette information est trompeuse car l'expérimentateur n'interrompt aucun des récits, même s'il croit qu'il est inventé, et car tous les participants se voient attribué un dédommagement de dix dollars peu importe leur performance. L'objectif de cette manipulation est évidemment de motiver les participants à être crédibles sans toute fois les pénaliser. Cette partie se termine avec la lecture et la signature du formulaire de consentement (voir annexe C).

La deuxième et principale partie de l'entrevue consiste en l'interrogatoire durant lequel le participant a à raconter chacun de ses récits. Cet interrogatoire, dont la fonction est d'éliciter un rappel libre le plus complet possible tout en évitant de suggérer des détails, est basé sur le format proposé dans la littérature sur le SVA (Yuille, 1988; Raskin et Esplin, 1991) et est adapté au contexte particulier demandant le rappel de deux récits. Pour mieux en comprendre le déroulement, l'interrogatoire sera sub-divisé en quatre étapes. Il est à noter que l'interrogatoire est enregistré au complet sur cassette audio pour en permettre la transcription et l'analyse.

Durant la première étape, le participant est invité à décrire son premier rêve ou événement de la façon la plus complète possible en commençant par le début. Les pauses sont tolérées mais lorsqu'elles sont trop longues, l'expérimentateur essaie de réamorcer le récit par une question comme "Vous avez dit X, qu'est-il arrivé après?" ou "Et après, qu'est-il arrivé?" Lorsque le participant termine son récit, l'expérimentateur lui demande s'il se rappelle d'autres choses par rapport à cet événement ou ce rêve. Lorsque le participant répond qu'il ne se rappelle de rien d'autre, il lui est demandé de décrire à nouveau de la façon la plus complète possible l'événement ou le rêve dont il vient de parler, avec la consigne supplémentaire d'essayer de n'oublier aucun détail, même s'ils paraissent insignifiants. La même procédure est utilisée que lors de la première fois. La fonction de ce deuxième rappel, suggéré entre autres par Yuille (1988), est de maximiser le nombre de

détails racontés et de disposer d'un plus grand échantillon verbal. Aucune question spécifique n'est posée durant cette étape de l'interrogatoire.

La deuxième étape est identique à la première sauf qu'elle a trait au deuxième récit rapporté par le participant. Avant chacune de ces deux étapes, le participant est laissé seul dans la salle d'entrevue pendant environ 5 minutes avec comme instruction de se préparer à l'interrogatoire pour le récit en question. Durant cette période de l'entrevue, et seulement durant cette période, le participant peut consulter le formulaire ou il a écrit des notes sur les récits qu'il a à raconter. La fonction de cette coupure est de laisser le participant anticiper l'interrogatoire, de s'assurer qu'il se rappelle de son récit et qu'il n'inverse pas l'ordre de présentation et de diminuer l'impact des facteurs extérieurs à l'expérience sur le rappel, comme par exemple un retard, une discussion récente ou le sujet du dernier cours.

L'objectif de la troisième étape est de donner la chance à l'expérimentateur de poser des questions spécifiques pour demander des précisions ou des clarifications à propos du premier récit. Seulement des questions non-directives sont posées selon un format standard pour éviter de suggérer des réponses ou des façons de répondre. Comme pour toutes les autres étapes, l'expérimentateur ne se permet pas de commentaires ou réflexions à propos du récit. Lorsque c'est possible, les questions sont des variations d'une des structures suivantes: "Vous m'avez parlé de X, pourriez-vous me le décrire plus précisément?", "Vous m'avez dit que X, pourriez-vous m'expliquer ce que vous entendez par ça?", "Vous rappelez vous de... (ce que vous portiez / les conditions météorologiques / la couleur de X / ...)?" Les questions ouvertes sont préférées et celles pouvant être répondues par oui ou non sont évitées lorsque possible, le but étant toujours d'éliciter le rappel le plus complet possible. Le nombre de questions posé dépend du nombre de précisions et de clarifications que l'expérimentateur croit nécessaire pour compléter le récit mais un minimum de trois questions est posé. Lorsque l'expérimentateur n'a plus d'autres questions à poser, il demande au participant s'il a d'autres choses à rajouter à propos de l'événement ou du rêve ou s'il y a des éléments qu'il aimerait préciser. Finalement, la quatrième étape est identique à la troisième sauf qu'elle a trait au deuxième récit raconté. Les questions non-dirigées sont posées après les deux rappels libres simplement pour éviter que les questions ayant trait au premier récit influencent le rappel libre du deuxième récit.

L'entrevue se termine par un debriefing durant lequel l'information trompeuse, à propos de la compensation additionnelle et de la possibilité que l'expérimentateur interrompt l'interrogatoire, et les raisons pour en faire usage sont exposées. L'expérimentateur répond alors à toute autre question ou préoccupation que le participant peut avoir concernant la

recherche au meilleur de ses connaissances et habiletés. Un deuxième formulaire de consentement (voir annexe C) est présenté et lu au participant pour assurer un consentement libre et éclairé. Si le participant décide de ne pas signer le deuxième formulaire, l'information et les enregistrements reliés à sa participation sont immédiatement détruits. Finalement, le participant reçoit un dédommagement de dix dollars, même s'il ne signe pas le formulaire de consentement, et il est remercié pour sa participation à la recherche.

Transcription et entrée de données

Le verbatim de chaque entrevue est par la suite transcrit sur traitement de texte. Tous les mots dits par le participant sont transcrits de façon intégrale mais les questions ou instructions standards dites par l'expérimentateur sont généralement résumées, à moins que leur formulation soit particulière. Tous les mots ou expressions inaudibles ou incompréhensibles rencontrés sont représentés par un trait continu sur la transcription. De plus, les pauses et les sons émis par le participant tels "amh", "ahh" et "mmm" sont transcrits au meilleur de l'habileté du transcripateur. Les enregistrements restent confidentiels et sont effacés après une période maximale d'un an. L'analyse subséquente de la validité de la déclaration se fait à partir des textes produits.

Avant d'encoder les transcriptions des entrevues, un test d'accord inter-juges est effectué. Pour se faire, trois récits de rêve et trois récits d'événement sont encodés indépendamment par l'expérimentateur principal et une seconde expérimentatrice et sont comparés entre eux. Le résultat est ensuite analysé. Les transcriptions de témoignage sont finalement encodées par l'expérimentateur principal et les données sont analysées après avoir été entrées sur support informatique.

Formation des évaluateurs

Toutes les déclarations sont encodées par l'expérimentateur principal et une partie des déclarations est aussi encodée par une seconde expérimentatrice pour évaluer l'accord inter-juges. L'expérimentateur principal, qui est aussi l'auteur de ce mémoire, était étudiant à la maîtrise en psychologie au moment de l'encodage tandis que la seconde expérimentatrice était en dernière année du baccalauréat en psychologie. Il est important de savoir que l'objectif de la recherche est de valider les critères du CBCA tels que décrits et opérationnalisés dans les écrits scientifiques et non tels qu'implicitement compris et utilisés par les chercheurs ayant participé à l'élaboration du SVA. Comme expliqué auparavant, cela permet de s'assurer que les critères tels que définis soit les mêmes que les critères tels que compris par les expérimentateurs et qu'un minimum de règles implicites soit utilisé.

La formation des deux expérimentateurs a donc été basée principalement sur les écrits disponibles et ils ne suivirent aucun programme de formation donné par un expert officiellement accrédité par les auteurs de la technique. Par contre, l'auto-formation fut encadrée par Dr Jean-Roch Laurence, un chercheur en psychologie légale connaissant la technique mais n'ayant jamais été officiellement formé dans son utilisation. Cette formation s'est étendue sur une période de neuf mois et représente plus de 150 heures. Durant ce temps, les expérimentateurs ont lu les écrits scientifiques disponibles en anglais et en français sur le sujet (p. e. Raskin et Esplin, 1991; Porter et Yuille, 1996; Landry et Brigham, 1992; Steller, 1989; Yuille, 1988) et des textes utilisés par les experts et dans les formations officielles (voir Yuille, 1990; Yuille, 1990a; et Dittmann, 1988), en ont discuté tout en tenant compte par écrit de la façon dont les critères étaient compris et opérationnalisés, se sont pratiqués à encoder sur plus d'une cinquantaine d'extraits de témoignages différents, ont comparé et discuté leur façon d'encoder, ont amélioré leur accord inter-juges et ont travaillé sur un autre projet de recherche évaluant la même technique (Lilova, 2000) qui a été déposé avant l'encodage pour la recherche actuelle et qui utilisait des souvenirs créés expérimentalement, soit le visionnement d'un cambriolage à main armée.

Opérationnalisation des critères

Dès le début de l'opérationnalisation des critères, il devint évident que plusieurs d'entre eux étaient définis de manière très vague et subjective dans les écrits actuels, ce qui laissait place à beaucoup d'interprétation dans la manière de les opérationnaliser. Malgré tout, les définitions officielles les plus récentes ont été suivies de la façon la plus conservatrice et méticuleuse possible et, lorsqu'une certaine précision ou interprétation était nécessaire pour l'opérationnalisation, elle y est mentionnée. Par contre, les définitions des critères n'ont pas été transformées pour les rendre plus facile à opérationnaliser, le but premier étant d'évaluer la technique et non de l'améliorer. Dans les prochains paragraphes, la façon dont chaque critère est opérationnalisé est expliquée de la façon la plus claire possible. Pour une définition théorique de chaque critère, le lecteur est référé aux définitions données dans l'introduction.

1. Structure logique

Ce critère est évalué selon une échelle de Lickert qui va de 0 à 2. L'expérimentateur évalue le récit dans son ensemble et compare les deux rappels entre eux pour tirer sa conclusion. Un récit dépourvu d'incohérences reçoit un 2. S'il contient de légères incohérences

affectant peu ou pas le déroulement ou le sens de l'histoire, il reçoit un 1. Un récit contenant une incohérence majeure qui le rend illogique reçoit un 0.

2. Production non-structurée

Ce critère utilise aussi une échelle de 0 à 2. L'expérimentateur inclut les deux rappels pour faire l'évaluation. Par contre, la période de questions n'est pas incluse car le type et le nombre de questions posées qui sont différent pour chaque récit affectent probablement ce critère. Les éléments suivants sont utilisés comme principaux marqueurs d'un rappel non-structuré: ne pas raconter dans l'ordre chronologique, revenir en arrière, faire des digressions, changements spontanés de direction dans la déclaration ("spontaneous shifts of focus"). Un récit structuré reçoit un 0. Un récit en général structuré mais comportant un des marqueurs de façon modérée seulement reçoit un 1. Un récit non-structuré reçoit un 2.

3. Quantité de détails

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque élément du récit. Chaque élément n'est compté qu'une seule fois. Les objets, les personnes et leurs attributs sont considérés comme des éléments du récit. Il en est de même pour les actions. Les parties du corps, à l'exception des cheveux, ne sont pas considérés comme des éléments distincts de la personne. Par exemple, "homme noir" et "homme à la peau noir" sont considérés comme deux éléments chacun. Chaque phrase citée équivaut à un élément. Étant donné que la définition semble faire une différence entre la quantité de détails de base et la quantité de détails additionnels lorsque plus de détails sont demandés, deux variables sont utilisées pour ce critère. La première variable est définie par le nombre de détails dans le premier rappel. C'est la variable principale pour évaluer ce critère. La deuxième variable, appelée réminiscence, est définie par le nombre de détails additionnels dans le deuxième rappel.

4. Enchâssement contextuel

Ce critère utilise une échelle de 0 à 2. Seul les deux rappels sont utilisés dans l'évaluation car le type et le nombre de questions posé par l'interviewer peut affecter le résultat de manière inégale. Les principaux indices utilisés pour évaluer ce critère sont les indications de temps et d'endroit, les endroits ou circonstances particuliers, comme un anniversaire de naissance, et les éléments de routine inclus dans le récit, comme aller se brosser les dents. De plus, pour les rêves, il a été décidé que les informations se rapportant au contexte dans lequel le récit fut rêvé se devaient aussi d'être considérés comme des indices

d'enchâssement contextuel. Un récit contenant peu ou pas d'indices contextuels, comme par exemple contenant seulement "il y a longtemps" et "chez nous", reçoit un 0. Un récit contenant l'information minimale seulement, comme par exemple "au début du semestre passé", "en début d'après-midi" et "à mon chalet", reçoit un 1. Un récit contenant, en plus de l'information minimale, un ou plusieurs indices particulier créant un contexte particulier et/ou incorporant des éléments de routine reçoit un 2.

5. Descriptions d'interactions

Ce critère utilise une échelle quantitative. Seul les deux rappels sont utilisés dans l'évaluation de ce critère. Un point est attribué pour chaque séquence d'interaction de trois éléments ou plus. Le participant doit faire partie de l'interaction, et non seulement l'observer, pour qu'elle soit comptée. Un élément d'interaction est défini comme une action ou une réaction dirigée vers l'autre interlocuteur et explicitement décrite dans le récit. Les actions doivent être des comportements, ce qui inclut les éléments de conversation, comme par exemple "je lui ai dit que ça ne valait pas la peine", et les expressions d'émotions, comme par exemple "je l'ai regardé d'un air de dégoût", mais exclut les expériences purement subjectives, comme par exemple "ça m'a rendu triste", étant donné que de telles expériences ne peuvent être directement perçues par l'interlocuteur. Dans le cas où plusieurs actions sont entreprises de façon consécutive par la même personne sans être entrecoupées de réactions de la part de la personne à qui sont adressées ces actions, comme par exemple "je lui ai dit "Bonjour!" et puis "Comment ça va?" et j'ai mis ma main sur son épaule", elles sont considérées comme un seul élément de communication. Par définition, les trois premiers éléments d'une interaction seront donc, dans l'ordre, une action effectuée par le premier interlocuteur, la réponse à cette action effectuée par le deuxième interlocuteur et la réponse à cette réponse par le premier interlocuteur. Une deuxième variable, définie comme le nombre d'interactions de deux éléments ou plus est aussi créée pour une deuxième analyse. Mais puisqu'elle inclut des données déjà utilisées dans la première variable, une correction statistique se doit d'être utilisée pour éviter les chances de créer une erreur de type 1.

6. Rappel de conversation

Ce critère est évalué selon une échelle quantitative. Un point est attribué pour chaque élément de conversation cité directement et intégralement. Seul les deux premiers rappels sont utilisés pour évaluer ce critère. Contrairement au critère #5, ces éléments de conversation n'ont pas à être suivis d'une réponse pour être comptés. Les citations de phrases dites à soi-même sont aussi comptées. Comme la plupart des catégories, celle-ci

n'est pas exclusive, ce qui fait qu'une même phrase pourrait remplir plus d'un critère. Par exemple, "je lui ai dit "Arrête de me faire peur!"" rempli les critères #6 et #12. Par simple curiosité scientifique, le nombre d'éléments de conversation cités indirectement et intégralement, comme "je lui ai dit qu'il devait aller dans la cuisine", est aussi évalué et regroupé sous une autre variable.

7. Complications inattendues

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque complication inattendue décrite dans le récit. Seulement les deux premiers rappels sont utilisés dans l'évaluation de ce critère.

8. Détails inusités

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque détail inusité décrit dans le récit. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère.

9. Détails superflus

Ce critère utilise une échelle de 0 à 2. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère. Les définitions officielles semblent utiliser les expressions "superflus" et "périphériques" de façon indifférenciée, malgré que le concept de "détail périphérique" renvoie à une notion sensiblement différente dans la littérature sur le souvenir et le témoignage oculaire. Le lecteur intéressé est incité à lire les travaux d'Elizabeth Loftus pour un approfondissement de ce concept. Pour les fins de cette étude, le concept de "détail superflu" tel qu'il est décrit dans la littérature sur l'analyse de la validité de la déclaration est utilisé et non le concept de "détail périphérique". Une simple description de l'endroit où se situe le récit sans détail caractéristique ne remplit donc pas ce critère, sauf si cette description va au-delà de ce qui est utile pour comprendre le récit. Un récit ne comportant aucun élément superflu reçoit un 0. Un récit comportant un ou deux éléments superflus peu détaillés reçoit un 1. Un récit comportant au moins un élément superflu très détaillé ou plusieurs éléments superflus reçoit un 2.

10. Détails non compris mais rapportés de façon exacte

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque détail non compris mais rapporté de façon exacte. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère.

11. Références à des incidents extérieurs

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque référence à un incident extérieur. Seulement les deux premiers rappels sont utilisés dans l'évaluation de ce critère. Il est considéré comme exclusif par rapport au critère #2, et cela pour en préciser leur définitions mutuelles.

12. Références à ses propres états subjectifs

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque énoncé référant aux états subjectifs du participant durant l'événement ou le rêve. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère. Les énoncés référant aux états subjectifs du participant durant l'entrevue sont exclus. Chaque émotion et chaque pensée explicitement nommées reçoivent un point. Cependant, lorsqu'une pensée est une explication d'une émotion vécue, comme c'est le cas pour "j'ai eu peur car je pensais qu'elle allait mourir", elle n'est pas comptée comme un élément différent de l'émotion puisqu'elle se réfère au même état subjectif. Les énoncés spécifiant que le participant savait ou ne savait pas quelque chose, comme par exemple "je ne savais pas le numéro" ou "je savais comment me rendre", et ceux spécifiant les goûts du participant, comme par exemple "j'aimais beaucoup le chocolat", sont exclus de cette analyse car ils réfèrent plus à une connaissance ou une préférence qu'à une expérience subjective spécifique.

13. Attributions d'états subjectifs à l'abuseur

Ce critère utilise une échelle quantitative. Étant donné que les récits ne décrivent pas nécessairement des abus, un point est attribué pour chaque état subjectif attribué à autrui. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère. Les règles utilisées sont les mêmes que pour le critère #12, mais avec une précision supplémentaire: les énoncés attribuant une intention à autrui, comme par exemple "il cherchait à m'attraper" ou "elle avait décidé de partir", sont exclus étant donné qu'ils sont souvent utilisés simplement pour décrire un comportement.

14. Corrections spontanées

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque correction ou addition spontanée. Seulement les deux premiers rappels sont utilisés dans l'évaluation de ce critère. Toute correction explicite est comptée, même si elle se produit juste après l'élément corrigé, à moins que l'élément corrigé n'ait pas été dit, comme c'est le cas pour "alors la j'ai... ...ben j'ai plutôt été manger". Les corrections de temps de verbe ou de

structure de phrase sont exclues. Une addition est définie comme un groupe d'éléments d'information qui satisfont au critère #3 et qui ajoute de la précision à une description donnée auparavant dans le même rappel. N'est pas considéré comme une addition une description de personne ou d'objet précédemment mentionné à moins que le participant ait commencé à décrire cet élément après l'avoir mentionné et que la seconde description ajoute des détails à la première. L'inclusion d'une action dans le récit lorsque le récit est rendu chronologiquement plus avancé, comme c'est le cas dans "Ah oui! Avant d'aller au dépanneur j'avais éteint la radio.", est considéré comme une addition. En plus de la variable principale tirée de ce critère, soit le nombre total de corrections et additions spontanées par récit, deux autres variables sont comptabilisées, soit le nombre de corrections spontanées par récit et le nombre d'additions par récit. Par contre, comme ces deux variables utilisent des données déjà utilisées, une correction statistique se doit d'être utilisée pour éviter les chances de créer une erreur de type 1.

15. Aveu d'absence de souvenir ou de connaissance

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque aveu d'absence de souvenir ou de connaissance. L'absence de souvenir se doit d'être explicite, non-ambiguë et reliée à un aspect de l'événement ou du rêve. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère.

16. Doutes à propos de sa propre déclaration

Ce critère utilise une échelle quantitative. Un point est attribué pour chaque énoncé rapportant un doute à propos de la déclaration. Seul les deux premiers rappels sont utilisés dans l'évaluation de ce critère. Deux types de formulation différentes semblent indiquer que le participant énonce un doute par rapport à sa déclaration. Dans le premier type, plus complet et basé sur une structure similaire à "je crois que... ..mais je ne suis pas sûr", le participant exprime clairement l'énoncé comme une croyance et non un fait et spécifie explicitement qu'il doute de sa véracité historique. Le deuxième type d'énoncé qui semble exprimer un doute mais de façon moins explicite utilise une structure similaire à "je crois que...", ce qui ne spécifie pas clairement le doute mais plutôt le sous-entend. Différentes expressions peuvent se substituer à "je crois", et ce dans les deux types d'énoncés. Les principaux sont "je pense", "j'imagine" et "j'ai l'impression". Les équivalents anglais "I believe", "I think", "I imagine", "I guess" et "...but I'm not sure" sont les principaux indices utilisés dans les récits anglais. Lorsque l'énoncé est utilisé à la fin du récit pour signifier que le récit est terminé, comme par exemple "je crois que c'est tout.", il est exclu de l'analyse. Étant donné que ces deux types d'énoncé semblent représenter deux degrés

différents de conviction, deux variables sont tirées de ce critère. Puisque les définitions officielles semblent inclure tout énoncé exprimant un doute, la variable principale est défini comme le nombre total d'énoncés rapportant un doute à propos de la déclaration. Une deuxième variable, n'utilisant que les énoncés exprimant le doute de façon explicite, c'est-à-dire ceux du premier type, est aussi comptée. Par contre, comme cette variable utilise des données déjà utilisées, une correction statistique se doit d'être utilisée pour éviter les chances de créer une erreur de type 1.

17. Désapprobation de sa propre participation

Étant donné la nature des récits, ce critère n'est pas analysé. Par contre, un critère similaire y est substitué de façon exploratoire. Selon un raisonnement similaire, il est possible que les participant aient plus tendance à éviter de présenter une image négative d'eux-mêmes dans les récits inventés que dans les récits réels. L'image de soi est définie comme la façon dont le participant réagit aux événements extérieurs et est distincte des événements eux-mêmes. Cette variable est aussi apparentée aux échelles de succès et d'échec telles que définies par Hall et Van de Castle (1966). Une échelle de 0 à 2 est utilisée pour ce critère. Un récit présentant une image positive du participant, c'est-à-dire qui inclut de façon explicite des exemples de circonstances où le participant résout un problème avec succès, réussit à se sortir d'une impasse et/ou pose des gestes socialement encouragés mais hors-normes, comme par exemple pardonner une grave offense ou se sacrifier pour autrui, reçoit un 0. Un récit présentant une image négative du participant, c'est-à-dire qui inclut de façon explicite des exemples de circonstances où le participant ne réussit pas à résoudre un problème, où son comportement résulte en une détérioration de sa situation ou où il pose des gestes socialement désapprouvés et hors-normes, comme par exemple frapper quelqu'un sans raison valable ou faire du chantage à une amie, reçoit un 2. Un récit présentant une image neutre du participant, c'est-à-dire qui inclut peu ou pas d'exemples explicites d'image négative ou positive ou qui inclut un nombre relativement équivalent de chacun, reçoit un 1.

18. Le fait d'excuser l'abuseur

Étant donné la nature des récits, ce critère n'est pas analysé. Par contre, il est partiellement représenté par la variable image de soi définie au critère #17.

19. Caractéristiques spécifiques du délit

Étant donné la nature des récits, ce critère n'est pas analysé.

En résumé, les seize variables principales sont: structure logique, production non-structuré, quantité de détails (premier rappel), enchâssement contextuel, interactions, rappel de conversation, complications inattendues, détails inusités, détails superflus, détails non compris, références à des incidents extérieurs, états subjectifs, attributions, corrections & additions spontanées (total), absence de souvenir et doutes (total). Les sept variable additionnelles sont: réminiscence, interactions (2+éléments), citations indirectes, corrections spontanées, additions spontanées, doutes (explicite) et image de soi.

RÉSULTATS

Deux des trente-six participants ont dus être éliminés de l'analyse car l'enregistrement de leur entrevue était incompréhensible. Le nombre final de participants par catégorie était donc de 7 francophones racontant des événements, 6 francophones racontant des rêves, 11 anglophones racontant des événements et 10 anglophones racontant des rêves. De plus, tous les récits eurent un score de 0 au critère #10, détails non compris mais rapportés de façon exacte. Cette variable fut donc retirée de l'analyse, ce qui ramène à quinze le nombre de variables principales. Il est aussi important de noter que, pour la variable structure logique, la note 0, représentant une incohérence majeure, ne fut attribuée à aucun récit. À moins d'indication contraire, chaque analyse a été exécutée avec une marge d'erreur de 0.05.

Accord inter-juges

L'accord inter-juges final a été effectué sur des récits anglais seulement car l'expérience d'encodage des deux expérimentateurs était plus grande dans cette langue qu'en français. Trois récits de rêve et trois récits d'événements choisis aléatoirement ont été utilisés pour l'analyse. Pour les variables ayant une petite fréquence d'apparition rendant difficile l'évaluation d'un accord inter-juges avec seulement six récits, six récits additionnels ont été utilisés. Une corrélation de Pearson a été effectuée pour déterminer la valeur de l'accord inter-juges. Deux variables ont été éliminées de cette analyse puisqu'elles utilisent les mêmes type d'éléments d'information que des variables déjà évaluées et que l'accord inter-juges déterminé pour les premières variables peut directement s'appliquer à elles. Ainsi, la variable réminiscence n'a pas été évaluée directement car la variable quantité de détails (1er rappel) était évaluée. De même, pour le critère #14, soit corrections spontanées, les corrections et les additions ont été évaluées séparément. La variable corrections & additions spontanées (total) a donc été éliminée de l'analyse. Cela amène le total de variables analysées à vingt. Comme on peut voir au tableau 2, l'accord inter-juges pour treize des vingt variables est plus grand ou égal à 0.70. Par contre, pour les variables structure logique, complications inattendues, détails inusités et références à des incidents extérieurs, il fut impossible d'évaluer l'accord inter-juges, malgré l'utilisation de douze récits au lieu de six, puisqu'une note de zéro, ou de deux dans le cas de structure logique, a été attribué à tous les récits évalués. Pour les variables production non-structurée, détails superflus et doutes (explicite), seulement un ou deux récits sur douze par évaluateur se sont fait attribuer une note différente de zéro. L'accord inter-juges n'a donc pas pu être évalué pour ces critères-là.

.....
Insérer tableau 2 environ ici
.....

Analyse de variance multivariée

Les résultats ont été analysés suivant un plan d'expérience $2 \times 2 \times (2)$. La variable indépendante intra-sujet est la véracité du récit et les deux variables indépendantes inter-sujets sont la source du récit et la langue utilisée durant l'entrevue. Malgré que l'ordre de présentation variait d'un participant à l'autre, il n'a pas été inclus dans les analyses car l'objectif était simplement de contrebalancer les récits réels et inventés. Ce plan fut utilisé sur les quinze variables principales. Par la suite, une analyse de variance univariée utilisant les mêmes variables indépendantes a été effectuée sur chacune des sept variables additionnelles.

Les moyennes et écarts-types pour chacune des quatre conditions sont illustrés au tableau 3 tandis que les statistiques descriptives additionnelles sont présentées en annexe F. Une analyse de variance multivariée a été effectuée avec les quatre variables indépendantes et les quinze variables dépendantes principales. Cette analyse, effectuée sur la version 10 du logiciel SPSS (Statistical Program for the Social Sciences), a dévoilé un effet principal du facteur véracité pour les critères structure logique, $F(1, 30)=9.182, p<0.05$, et détails inusités, $F(1, 30)=4.712, p<0.05$, un effet principal du facteur source du récit pour les critères quantité de détails, $F(1, 30)=5.304, p<0.05$, enchâssement contextuel, $F(1, 30)=13.211, p<0.05$, complications inattendues, $F(1, 30)=7.905, p<0.05$, détails inusités, $F(1, 30)=7.482, p<0.05$, et absence de souvenir, $F(1, 30)=4.672, p<0.05$, et finalement une interaction véracité*source du récit*langue utilisée pour le critère références à des incidents extérieurs, $F(1, 30)=6.220, p<0.05$. L'analyse des effets simples de l'interaction a dévoilé une interaction source du récit*langue utilisée pour les déclarations réelles, $F=4.812, p<0.05$. Aucun effet simple de second ordre n'a été significatif. Pour les fins de l'interprétation de cette interaction, il est important de noter que ce critère a été noté comme présent dans seulement 4 des 68 récits dont 3 des 34 récits réels.

Tableau 2

Accords inter-juges des différents critères évalués mesurés par des corrélations de Pearson.

CRITÈRE	CORRÉLATION
3 Quantité de détails (1er rappel)	0.96
4 Enchâssement contextuel	0.92
5 Interactions(3+)	0.80
Interactions (2+)	0.91
6 Rappel de conversation	0.84
Citations indirectes	0.95
12 États subjectifs	0.98
13 Attributions	0.92
14 Corrections spontanées	0.93
Additions spontanées	0.75
15 Absences de souvenir	0.70
16 Doutes (total)	0.92
17 Image de soi	1.00

Note. $p < 0.05$ pour chacune de ces corrélations.

L'analyse effectuée sur les sept variables additionnelles a dévoilé un effet principal du facteur véracité pour le critère réminiscence, $F(1, 30)=5.255$, $p<0.05$, un effet principal du facteur source du récit pour le critère citation indirecte, $F(1, 30)=6.966$, $p<0.05$, un effet principal du facteur langue pour le critère doutes (explicite), $F(1, 30)=7.358$, $p<0.025$ (correction Bonferroni), et finalement une interaction véracité*langue utilisée pour le critère image de soi, $F(1, 30)=11.068$, $p<0.05$. Les effets simples de cette interaction ont été analysés par un test de T sur la variable véracité pour chaque niveau de la variable langue. Chez les francophones, un effet simple a été dévoilé, $T=3.323$, $p<0.05$. Aucun effet significatif n'a été dévoilé chez les anglophones.

.....

Insérer tableau 3 environ ici

.....

L'analyse nous démontre que les récits véridiques contenaient plus souvent des incohérences mineures ($\mu=1.62$, $s=0.60$) et moins souvent des détails inusités ($\mu=0.09$, $s=0.29$) que les récits inventés ($\mu=0.97$, $s=0.17$, et $\mu=0.29$, $s=0.46$ respectivement). Par ailleurs, les récits véridiques contenaient plus d'éléments nouveaux au deuxième rappel ($\mu=12.97$, $s=14.03$) que les récits inventés ($\mu=8.35$, $s=6.05$). Ni l'effet global des quinze critères tirés du CBCA ni aucune autre variable individuelle ne différenciait de façon significative ces deux types de récits. De plus, l'effet global des quinze critères principaux sur la variable véracité n'était pas significatif.

Par rapport aux récits de rêves, les récits d'événements étaient enchâssés dans un contexte plus riche ($\mu=1.28$, $s=0.78$, et $\mu=0.44$, $s=0.62$), contenaient plus de détails ($\mu=84.22$, $s=43.98$, et $\mu=56.19$, $s=21.46$), plus de complications inattendues ($\mu=0.19$, $s=0.40$, et $\mu=0.00$, $s=0.00$), moins de détails inusités ($\mu=0.08$, $s=0.28$, et $\mu=0.31$, $s=0.47$) et moins d'aveu d'absence de souvenir ou de connaissance ($\mu=0.33$, $s=0.53$, et $\mu=1.47$, $s=2.68$) que les récits de rêves. De plus, les récits d'événements contenaient plus de citations indirectes que les récits de rêves. Ni l'effet global des quinze critères tirés du CBCA ni aucune autre variable individuelle ne différenciait de façon significative ces deux types de récits.

Les récits français contenaient plus d'énoncés exprimant explicitement un doute à propos de la déclaration ($\mu=0.92$, $s=1.47$) que les récits en anglais ($\mu=0.12$, $s=0.33$). De plus, les récits inventés en français présentaient, en moyenne, une image de soi plus négative

Tableau 3

Moyennes (M) et écarts-types (ÉT) des événements réels (ÉR), événements inventés (ÉI), rêves réels (RR) et rêves inventés (RI) en fonction de chaque critère.

CRITÈRE	ÉR		ÉI		RR		RI	
	M	ÉT	M	ÉT	M	ÉT	M	ÉT
<u>Critères principaux</u>								
1. Structure logique ^a	1.61	0.61	1.94	0.24	1.62	0.62	2.00	0.00
2. Production non-structuré	0.44	0.51	0.50	0.71	0.25	0.45	0.06	0.25
3. Quantité de détails ^b	82.17	36.21	86.28	51.60	53.06	21.51	59.31	21.63
4. Enchâssement contextuel ^b	1.33	0.77	1.22	0.81	0.44	0.63	0.44	0.63
5. Interactions	1.06	1.11	0.83	1.34	0.50	0.89	0.50	0.52
6. Rappel de conversation	6.06	7.63	6.33	13.25	2.50	2.88	2.13	2.22
7. Complic. inattendues ^b	0.33	0.49	0.06	0.24	0.00	0.00	0.00	0.00
8. Détails inusités ^{a,b}	0.06	0.24	0.11	0.32	0.13	0.34	0.50	0.52
9. Détails superflus	0.28	0.46	0.22	0.43	0.06	0.25	0.13	0.34
11. Incidents extérieurs ^c	0.06	0.24	0.06	0.24	0.13	0.34	0.00	0.00
12. États subjectifs	4.56	3.05	4.61	3.50	3.19	2.76	3.31	2.52
13. Attributions	1.39	1.50	1.33	1.68	0.38	0.62	0.94	0.85
14. Corr. spontanées (total)	2.22	2.29	2.06	1.66	1.44	1.41	0.81	0.66
15. Absence de souvenir ^b	0.39	0.61	0.28	0.46	1.81	3.62	1.13	1.20
16. Doutes (total)	2.56	3.58	2.06	2.07	1.06	1.12	1.31	2.33
<u>Critères additionnels</u>								
Réminiscence ^a	16.67	15.71	9.78	6.04	8.81	10.89	6.75	5.83
Interactions (2+éléments)	1.72	1.64	1.89	1.71	1.00	1.10	1.06	0.68
Citations indirectes ^b	1.78	1.59	1.00	0.77	0.69	0.79	0.81	1.05
Corrections spontanées	1.61	2.17	1.22	1.31	0.69	1.08	0.38	0.50
Additions spontanées	0.61	0.61	0.83	0.92	0.75	1.24	0.44	0.63
Doutes (explicite)	0.33	0.69	0.61	1.46	0.44	0.73	0.31	1.01
Image de soi ^c	1.17	0.79	1.33	0.59	1.19	0.54	0.88	0.62

^a effet principal de véracité, $p < 0.05$

^b effet principal de type de récit, $p < 0.05$

^c interaction impliquant les variables véracité et/ou type de récit, $p < 0.05$

($\mu=0.77$, $s=0.44$) que les récits réels en français ($\mu=1.46$, $s=0.66$). Pour ce qui est de l'interaction rapportée à propos des références à des incidents extérieurs, aucune interprétation ne peut être tentée à cause du nombre restreint de cas où le critère était présent, l'effet simple d'interaction ne reposant que sur 3 cas sur 34. Aucune autre variable n'a été affectée de façon significative par la langue utilisée.

Établissement de normes

Étant donné que seulement deux des variables principales, soit structure logique et détails inusité, différenciaient de façon significative les récits réels des récits inventés, que la direction des deux effets était à l'inverse des prédictions du SVA et que l'effet global de la variable véracité dans l'analyse multivariée n'était pas significatif, aucun essai d'établir des normes d'applications pour l'analyse de la validité de la déclaration n'a été effectuée.

DISCUSSION

Interprétation des résultats

Pour aider à l'interprétation des résultats, chaque effet sera analysé plus en profondeur dans les prochaines lignes en fonction du paradigme de départ.

Pour le critère "structure logique", seul un effet principal de la variable "véracité" a été significatif. Selon le paradigme présenté en fin d'introduction, cet effet peut être attribué à l'intention de tromper car, lorsque les participants ont cherché à tromper l'interviewer, ils ont moins souvent produit une déclaration contenant des incohérences mineures que lorsqu'ils ont cherché à dire la vérité.

Pour le critère "détails inusités", un effet principal a été significatif pour les variables "véracité" et "source du récit" mais aucune interaction n'a été significative. L'effet peut donc être attribué à l'intention de tromper et au statut particulier attribué aux rêves, c'est-à-dire que premièrement, lorsque les participants ont cherché à tromper l'interviewer, ils ont plus souvent produit une déclaration contenant des détails inusités que lorsqu'ils ont cherché à dire la vérité, et que deuxièmement les participants qui avaient à raconter un récit en considérant le fait qu'il devait avoir le statut de rêve ont plus souvent produit une déclaration contenant des détails inusités que les participants qui avaient à raconter un récit en considérant le fait qu'il devait avoir le statut d'événement.

Pour le critère "réminiscence", seul un effet principal de la variable "véracité" a été significatif. L'effet peut donc être attribué à l'intention de tromper car, lorsque les participants ont cherché à tromper l'interviewer, ils ont produit une déclaration contenant en moyenne moins d'éléments nouveaux au deuxième rappel que lorsqu'ils cherchèrent à dire la vérité.

Pour les critères "enchâssement contextuel", "quantité de détails", "complications inattendues" et "citations indirectes", seul un effet principal de la variable "source du récit" a été significatif. Aucune interaction n'a été significative. Ces effets peuvent donc être attribués au statut particulier attribué aux rêves, c'est-à-dire que les participants qui avaient à raconter un récit en considérant le fait qu'il devait avoir le statut de rêve ont produit une déclaration qui contenait en moyenne moins d'éléments contextuels, de détails et de citations indirectes mais qui incluait plus souvent des complications inattendues et des aveux d'absence de souvenir ou de connaissance que les participants qui avaient à raconter un récit en considérant le fait qu'il devait avoir le statut d'événement.

Pour le critère "doutes explicites", seul un effet principal de la variable "langue" a été significatif. L'effet peut donc être attribué à l'influence de la langue ou du groupe culturel, c'est-à-dire que les participants s'exprimant en français ont eu plus tendance à produire des déclarations contenant des doutes explicites que les participants s'exprimant en anglais.

Pour le critère "image de soi", seule une interaction "véracité*langue" a été significative. Cet effet peut être attribué à l'intention de tromper et à l'influence de la langue ou du groupe culturel, c'est-à-dire que lorsque les participants s'expriment en français ont cherché à tromper l'interviewer, ils ont produit une déclaration présentant en moyenne une image d'eux plus négative que lorsqu'ils ont cherché à dire la vérité. Les différentes déclarations des participants s'exprimant en anglais ne différencient pas entre eux dans la valeur de l'image qu'ils présentaient d'eux.

Pour le critère "référence à des incidents extérieurs", une interaction "véracité*source du récit*langue" a été significative et a dévoilé une interaction "source du récit*langue" pour les déclarations réelles. Vu la fréquence réduite de ce critère, il est difficile d'interpréter ces résultats. Par contre, si on se risque à le faire, il semble premièrement que l'effet ne puisse être attribué au statut particulier attribué aux rêves car les déclarations inventées ne différaient pas entre elles. Étant donné l'interaction, on peut supposer que la source du souvenir, soit externe ou interne, est, du moins partiellement, responsable de l'effet.

En résumé, dans cette étude évaluant 16 critères tirés du CBCA et 7 variables additionnelles dans un contexte de variabilité réaliste et de temps de préparation réaliste, seule la validité et l'efficacité de deux des critères tirés du CBCA, soit la structure logique et les détails inusités, a été démontrée. Par contre, la direction de l'effet fut contraire aux prédictions du CBCA pour ces deux critères. La validité et l'efficacité d'un troisième critère, la réminiscence, a aussi été démontrée. L'effet de ce critère, indirectement inspiré du CBCA, a une direction en accord avec les prédictions qui en découlent, soit un plus grand nombre d'éléments dans les déclarations véridiques. La validité et l'efficacité d'un quatrième critère, l'image de soi, a été démontré, mais seulement pour les participants s'exprimant en français. L'intention de tromper, et non la source du récit, a été le facteur détecté par ces quatre critères. Aucun critère n'a réagi de façon significative à la source du récit. Le critère "référence à des incidents extérieurs" est un candidat potentiel pour détecter la source d'une déclaration. Par contre cette hypothèse est basée sur des effets ambigus et contestables. Un effet de la langue et/ou du groupe culturel a été démontré pour deux des critères additionnels inspirés du CBCA, soit les doutes explicites et l'image de soi. Aucun des 16 critères tirés du CBCA n'a été affecté significativement par la langue ou le groupe culturel des participants. Finalement, malgré que cela ait peu d'utilité clinique ou légale, il

est bon de mentionner que quatre des critères tirés du CBCA, soit l'enchâssement contextuel, la quantité de détails, les complications inattendues et les détails inusités, et un des critères additionnels, soit les citations indirectes, ont réagi au statut particulier attribué aux rêves. Les récits de rêves peuvent donc être considérés qualitativement différents des récits d'événements en rapport à ces critères-là.

Comparaison à la littérature

Au premier regard, il peut sembler étrange que la structure logique des déclarations inventées soit plus grande que celle des déclarations réelles. Malgré que ce résultat est à l'opposé des prédictions du SVA, des résultats identiques furent démontrés dans deux autres études, soit celles de Landry et Brigham (1992) et de Ruby et Brigham (1998). Une des particularités qui distingue ces études des autres études expérimentales effectuées avec des adultes est que les participants ont eu un temps de préparation de plus d'une journée, comme pour la présente étude. Il est donc très probable que l'effet dépende du temps que le participant ou le témoin dispose pour élaborer son récit. Un témoin ayant moins d'une heure pour inventer un récit crédible a probablement plus de possibilités d'inclure involontairement des incohérences logiques dans son récit simplement parce que peu de temps était disponible pour l'élaboration. Lorsqu'une déclaration est basée sur un événement réel, un script de base est déjà disponible et le récit a peu besoin d'être élaboré. Il y a donc moins de possibilités que des incohérences logiques s'y glisse. Mais lorsqu'un témoin a plus d'une journée pour inventer un récit crédible, il a très certainement le temps de le réviser plus d'une fois pour s'assurer qu'aucune incohérence s'y glisse et que le récit est parfaitement logique. Par contre, lorsque le témoin cherche à dire la vérité, il serait surprenant qu'il cherche à réviser son récit et le transformer pour lui donner une apparence plus logique. Ces résultats démontrent donc l'importance d'étudier les différents critères sous plusieurs conditions et surtout d'utiliser des conditions de variabilité et de préparation réalistes si l'on veut connaître l'utilité légale et clinique de ces critères.

Pour ce qui est des détails inusités, l'explication est moins claire. Les études de Ruby et Brigham (1998) et de Vrij et al. (2000) ont démontré un effet significatif de ce critère dans la direction prévue par le SVA. Par contre, pour l'étude de Vrij et al. (2000), le fait d'avoir utilisé un événement-cible unique est problématique par rapport à l'évaluation de ce critère. Comme il a été expliqué dans la section résumant cette étude, ce facteur peut avoir créé artificiellement une différence au niveau de ce critère, ce qui limite la possibilité de généraliser le résultat. Pour ce qui est de l'étude de Ruby et Brigham (1998), les

différences expérimentales ne peuvent expliquer clairement cette incongruence au niveau des détails inusités. Par contre, comme ce critère fait partie des cinq critères dont la fidélité est inadéquate (Anson, Golding et Gully, 1993; Horowitz et al., 1997), il est possible que ce résultat s'explique par le fait que les règles implicites utilisées par les évaluateurs pour combler les lacunes des définitions officielles étaient différentes dans ces deux études. Cela amène donc un appui additionnel à la nécessité d'opérationnaliser clairement et explicitement les critères évalués. Malgré ces résultats incongrus, il y a probablement un espoir pour ce critère. Par contre, l'influence des différents paradigmes expérimentaux utilisés se doit d'être étudiée en profondeur avant de pouvoir généraliser sa validité et son efficacité aux contextes légal ou clinique.

L'efficacité du critère "réminiscence" revêt un caractère très intéressant. Malgré que le critère officiel de "quantité de détails" a été absolument inefficace pour différencier les déclarations véridiques des déclarations non-véridiques, une reconceptualisation de ce même critère a été de loin un des critères les plus efficaces. Ce résultat appuie fortement le besoin criant de se débarrasser du carcan du CBCA pour les chercheurs ayant comme objectif d'arriver à une procédure pouvant réellement différencier les déclarations véridiques des déclarations non-véridiques en contexte légal ou clinique. Tandis que la conservation de ce carcan réduit la recherche à évaluer sans cesse les mêmes critères dont l'efficacité est loin d'être supportée par la littérature actuelle, une approche moins dogmatique à l'analyse de la validité des déclarations maximise la valeur heuristique de l'hypothèse d'Undeutsch en permettant dévaluer et d'améliorer l'efficacité des différentes variantes des critères actuels et d'intégrer de nouveaux critères suggérés par d'autres domaines de la littérature, comme par exemple la mémoire, la psycholinguistique ou les comportements non-verbaux. Cette approche a permis ici de trouver un critère précis, clair et explicite plus efficace que le critère d'origine en conditions de variabilité et temps de préparation réalistes.

La réaction des critères "image de soi" et "doutes explicites" est très intéressante à interpréter à la lumière des résultats obtenus par Ruby et Brigham (1998) sur les différences culturelles. Malgré que la différence entre les groupes culturels anglophone et francophone du Québec est probablement moins grande que celle entre les groupes ethniques afro-américain et caucasien des États-Unis, un certain appui a été trouvé dans la présente étude pour l'hypothèse de l'influence de l'appartenance culturelle. Par contre, il est aussi possible que l'effet ait été produit par une différence au niveau de la langue parlée. Malgré la similarité des deux concepts, il pourrait être intéressant d'approfondir cette question dans de futures études, surtout si un ensemble de critères efficaces commencent à être développés. Étant donné que la variable "image de soi" ait été capable de distinguer les

déclarations réelles des déclarations inventées chez les francophones, il y a peut-être une piste intéressante à suivre. Par contre, le fait qu'une différence culturelle relativement modeste fut suffisante à affecter ce critère démontre la nécessité de le redéfinir ou l'éliminer si l'objectif est de l'utiliser en contexte légal ou clinique. Il est important de noter qu'aucun des critères principaux directement tirés du CBCA n'a été affecté par l'appartenance à un des deux groupes culturels. Ainsi, des quatre variables différenciant les déclarations réelles des déclarations inventées, seulement l'image de soi, un critère vaguement inspiré du CBCA, a été affecté par l'appartenance culturelle et/ou la langue parlée dans la présente étude. L'effet a donc été de moins grande envergure que celui démontré par Ruby et Brigham (1998), ce qui s'explique facilement étant donné que la différence culturelle entre les groupes était moins grande dans la présente étude. Il est donc à prévoir que des différences culturelles plus marquées auraient plus tendance à affecter les différents critères.

Une des questions qui était à l'étude était de savoir si les effets détectés par les différents critères du CBCA étaient causés par un facteur motivationnel, soit l'intention de mentir, ou un facteur cognitif, soit la source du récit. Aucune donnée n'était alors accessible à ce propos. Cette étude démontra que, de tous les critères évalués, quatre d'entre eux, ou trois si le critère "image de soi" est éliminé, ont pu détecter l'influence de l'intention de tromper chez les participants. Par contre, aucun des critères étudiés n'a été capable de détecter si le récit originait d'une source interne ou externe. Cela nous amène donc à conclure que ces critères ne réagissent pas à la véracité du récit évalué mais bien à la véracité attribuée par le témoin au récit, c'est-à-dire qu'ils ne font que détecter si le témoin croit en son récit et qu'ils seraient par conséquent incapable de reconnaître comme non-véridique un récit qui serait le résultat de la suggestion ou d'une simple erreur d'attribution de source. La procédure partage donc certaines limites avec le polygraphe qui lui aussi ne détecte que le facteur motivationnel, et donc ne peut que détecter si le témoin croit en son récit. Pour correctement représenter son rôle, le Statement Validity Analysis devrait donc être rebaptisé le Statement Attributed Validity Analysis, ou l'analyse de la validité attribuée de la déclaration, malgré que son utilité et son efficacité soient de plus en plus remis en question. Comme aucune étude de ce genre n'a encore été effectuée chez les enfants, on ne peut être sûrs que ces conclusions s'appliquent à l'utilisation des différents critères avec des déclarations d'enfants. Par contre, comme ces résultats sont les seuls dont nous disposons sur le sujet, il est proposé de généraliser ces conclusions jusqu'à preuve du contraire, la prudence étant préférée à l'imprudence.

Hypothèses alternatives et limites

Assurément, certaines hypothèses alternatives peuvent expliquer les résultats obtenus et certaines critiques peuvent être faites par rapport à l'interprétation de ces mêmes résultats. L'objectif de la présente section est de faire un survol de ces différentes hypothèses et critiques pour s'assurer de bien saisir les limites de l'étude.

Premièrement, malgré qu'un des objectifs était d'évaluer les critères originaux du CBCA, la définition opérationnelle de certains d'entre eux a dû être précisée. Une approche objective et quantifiable avec des règles explicites fut préférée à une approche plus intuitive. Cette décision n'est pas une limite en soi mais il est possible qu'une opérationnalisation différente ait des résultats différents. Par contre cet argument ne devrait pas servir d'excuse à la performance du CBCA dans cette étude mais bien de motivation à évaluer des opérationnalisations alternatives dans de futures recherches. Une limite supplémentaire est liée au problème de fidélité de certains critères. Un des critères qui a été significatif, soit les détails inusités, et quatre des critères qui ne l'ont pas été, soit les détails superflus, la référence à des incidents extérieurs, les aveux d'absence de souvenir et les corrections spontanées, font partie des critères dont la fidélité a été jugée inadéquate dans les deux études qui en ont fait l'évaluation (Anson, Golding et Gully, 1993; Horowitz et al., 1997). Il est donc possible que les résultats obtenus dans la présente étude soient dus à une interprétation particulière de leur définition, ce qui limiterait la possibilité de généraliser ces résultats. Par contre, des mesures particulières ont été déployées dans cette étude pour rendre explicite toute opérationnalisation additionnelle. De plus, cette limite est inhérente à toutes les études évaluant ces critères.

Un des faits à se rappeler est que, pour des raisons techniques, trois critères, soit la désapprobation de sa propre participation, le fait d'excuser l'abuseur et les caractéristiques spécifiques du délit, n'ont pas été évalués. Les conclusions apportées dans le présent mémoire ne sont donc qu'indirectement reliées à ces critères. Par contre, il est intéressant de noter que des formulations plus récentes du SVA ont éliminé le critère "caractéristiques spécifiques du délit" du CBCA pour l'inclure dans la *Validity Checklist* (Raskin et Esplin, 1991) tandis qu'une troisième version, aussi approuvée par des auteurs du SVA, a éliminé quatre critères additionnels du CBCA, soit les aveux d'absence de souvenir, les doutes à propos de sa déclaration, la désapprobation de sa propre participation et le fait d'excuser l'abuseur, sans même les transférer à la *Validity Checklist* (Raskin, Esplin et Horowitz, 1991, dans Craig, Scheibe, Raskin, Kircher et Dodd, 1999). L'argument utilisé pour éliminer ces critères était qu'ils n'étaient pas reliés à l'hypothèse d'Undeutsch. Malgré ces modifications, peu d'études semblent avoir suivi, pas même celle-ci. La version originale du

CBCA est donc encore la plus étudiée. Cette résilience de l'ensemble officiel amène le présent auteur à souligner de nouveau l'importance d'étudier chaque critère comme une entité indépendante, et non comme une partie d'un ensemble fixe.

Un autre problème est lié au fait que la fréquence de certains critères était extrêmement basse. Cela a amené l'élimination du critère "détails superflus" pour l'analyse statistique et a rendu difficile l'interprétation du critère "référence à des incidents extérieurs". Pour compenser cette difficulté, un plus grand nombre de participants aurait pu être utilisé. Par contre, on est en droit de se demander si un critère atteint d'un grave problème d'atténuation d'échelle peut réellement être utile, surtout lorsque sa présence n'est pas exclusive aux déclarations véridiques. Par contre, étant donné le manque de normes, il est possible que l'opérationnalisation utilisée dans la présente étude ait été plus conservatrice que les différentes opérationnalisations des autres études. Il est aussi possible que cette basse fréquence soit due à une caractéristique intrinsèque des déclarations à propos d'abus sexuel pour lesquelles les critères ont été construits. Cette dernière hypothèse s'éloigne par contre du postulat de base de l'hypothèse d'Undeutsch voulant que l'efficacité des différents critères soit attribuable à une différence intrinsèque entre les déclarations véridiques et les déclarations non-véridiques. De plus, des études utilisant des cas réels ont aussi démontré une très grande variabilité dans la fréquence des critères. L'incidence des critères évalués par Horowitz et al. (1997), par exemple, variait de 0.01 à 0.97 tandis que, dans l'étude d'Anson, Golding et Gully (1993), elle variait de 0.00 à 0.91. Ce problème supporte à nouveau l'urgence de préciser l'opérationnalisation des différents critères.

De plus, l'expérimentateur n'était pas aveugle quant au type de récit, soit un événement ou un rêve. Il faut donc tenir compte de la possibilité d'un biais expérimental au niveau des résultats concernant l'influence de la source du récit, qui est externe ou interne, et du statut particulier attribué aux rêves. Si un tel biais était présent, il a pu produire des différences significatives en accord avec les attentes de l'expérimentateur lorsque aucune différence n'était réellement présente. Il est donc important d'interpréter avec une certaine prudence les effets significatifs en rapport au statut particulier attribué aux rêves. Par contre, comme aucune influence de la source du récit n'a été démontré de façon significative, cet argument n'a qu'une importance limitée puisqu'il est plus difficile à utiliser pour justifier une absence qu'une surabondance de résultats significatifs.

Comme pour toutes les études expérimentales effectuées auparavant, les événements-cibles utilisés n'étaient pas des récits d'abus sexuel. Il peut toujours être argumenté que les différents critères du CBCA ont été conçus pour être utilisés avec des déclarations à propos d'abus sexuels et que leur manque d'efficacité avec d'autres types de déclaration ne

peut être généralisé à la problématique des abus sexuels, mais si l'on en croit l'hypothèse d'Undeutsch et l'opinion des auteurs du SVA comme Yuille (1989) et Steller (1989), l'efficacité des différents critères n'est pas relié au type de déclaration évaluée et devrait être généralisable à plusieurs types de déclarations. De plus, dans la présente étude, certaines précautions ont été prises pour s'assurer un minimum d'équivalence entre les récits expérimentaux et des récits d'abus. Il a entre autres été demandé de rapporter des récits à propos d'événements marquants, anecdotiques, c'est-à-dire différents du quotidien, et ayant une certaine valeur émotionnelle. Par contre, il n'a pas été exigé de rapporter des récits à teneur émotionnelle négative. Cette différence peut avoir eu un impact sur les résultats mais aucune recherche n'a encore confirmé cette hypothèse.

Une dernière problématique importante à mentionner se rapporte au réalisme de la situation expérimentale. Malgré que, dans cette étude, un effort particulier a été pris pour s'assurer du réalisme du temps de préparation et de la variabilité entre les déclarations, certaines variables importantes ont dû être éliminées. Les participants recevaient, entre autres, comme précision de ne pas rapporter d'événements reliés à l'enfance, soit avant l'âge de 14 ans, ou pour lesquels ils avaient consultés en thérapie. Il était nécessaire d'utiliser ces consignes pour éviter d'évaluer des récits réels fortement altérés par le temps ou des influences extérieurs mais, du même coup, cela diminuait le réalisme de la situation en éliminant certaines possibilités rencontrées en contexte légal. De plus, malgré que l'utilisation d'un temps de préparation de quelques jours augmente le réalisme de la situation, cela ne tient pas compte de la grande variabilité à ce niveau en contexte réel. Malgré qu'un délai de quelques jours constitue probablement un délai minimal réaliste, des semaines de préparation, ou mêmes des mois, peuvent passer dans certains cas avant la première déposition. Il est donc impossible de tenir compte de tous les facteurs influençant une déclaration en contexte réel, ce qui laisse supposé que tout critère efficace en contexte expérimental sera très certainement moins efficace en contexte légal et dans les études utilisant des cas réel. L'efficacité de chaque critère doit donc être évaluée à la baisse pour les études expérimentales lorsque l'application légale est considérée, même lorsque des conditions dites réalistes sont utilisées. De telles conditions sont quand même probablement les meilleures approximations expérimentales d'un contexte réel.

L'application légale pose aussi un problème additionnel de taille qui est peu abordé dans la littérature actuelle. Ce problème est celui du groupe auquel le récit est comparé. Lors d'études expérimentales, les récits évalués sont comparés à des récits de véracité opposée raconté soit par les mêmes participants, dans un plan intra-sujet, soit par d'autres participants provenant d'un groupe jugé équivalent, dans un plan inter-sujet. Il y a donc un groupe équivalent auquel les récits peuvent être comparés. En contexte légal par contre, il

est difficile de trouver un groupe équivalent auquel comparer le récit évalué, premièrement parce que les caractéristiques spécifiques de la situation et du témoin peuvent être très complexes et deuxièmement parce que certaines caractéristiques de la situation et/ou du témoin sont souvent inconnues de l'évaluateur ou très difficile à évaluer, comme par exemple le temps de préparation dans le cas d'un récit inventé. Il est donc très facile d'utiliser un groupe de comparaison inadéquat lors d'une expertise. Une solution à ce problème passerait peut-être par l'utilisation d'une procédure semblable à celle d'une entrevue au polygraphe durant laquelle le niveau de base évalué en entrevue est utilisé comme point de comparaison. Par contre, une telle procédure devra être validée de façon empirique avant d'être utilisée et le protocole exact à suivre devra être défini.

CONCLUSION

Malgré la validité que certains des critères du CBCA ont démontré en milieu contrôlé et variabilité réduite, une analyse en profondeur de la présente étude et des écrits scientifiques se rapportant au SVA et au CBCA amène comme constat que la technique, sous sa forme actuelle, est loin d'avoir démontré de façon empirique son efficacité à différencier les déclarations véridiques des déclarations non-véridiques, surtout lorsque des correctifs visant à augmenter le réalisme expérimental sont utilisés.

En plus de démontrer l'inefficacité de la plupart des critères du CBCA à détecter l'influence de la tromperie, cette étude a permis de démontrer qu'aucun des critères n'était sensible à la source de la déclaration. Cela amène l'auteur à conclure que seul la véracité attribuée par le témoin à sa déclaration est évaluée par les quelques critères efficaces tirés du CBCA. Malgré que ces conclusions concernent principalement l'utilisation de la procédure chez une population adulte, les experts voulant utiliser le SVA pour évaluer des déclarations d'enfants devraient quand même en tenir compte. Les études chez les enfants ont généralement obtenu des résultats plus positifs mais encore très controversés.

De plus, différents problèmes tels le manque de connaissances sur la façon dont l'âge affecte chacun des critères jugés valides et l'absence de normes stratifiées par groupe d'âge empêchent l'application efficace d'une telle technique. Tant pour les adultes que pour les enfants, les normes devront probablement tenir compte de l'appartenance culturelle du témoin.

La plus grande efficacité des critères dans les études utilisant des cas réels en comparaison aux études expérimentales semble suggérer que les différents critères du CBCA sont plus affectés par la crédibilité que par la véracité des déclarations. Par contre, des études quasi-expérimentales devraient être effectuées avec des populations adultes pour approfondir cette question.

Pour ce qui est de la recherche, les différents critères du CBCA ont quand même une certaine valeur en tant que *prima mater* dans le développement d'une procédure évaluant la véracité attribuée d'une déclaration. Mais pour ce faire, ils doivent nécessairement être débarrassés du carcan du CBCA qui impose les critères à utiliser, la façon de les utiliser et leur définition officielle. Il est inutile de continuer les études sur les scores totaux du CBCA car plusieurs des critères semblent n'être que des fardeaux sans efficacité démontrée tandis que certains autres ont besoin de retourner sur les planches à dessins pour être précisés plus explicitement ou être reformulés. Une approche évaluant les

ingrédients actifs et incluant de nouveaux critères serait donc beaucoup plus appropriée. Il serait aussi important de continuer à différencier l'influence de l'intention de tromper de celle de la source du récit pour les critères prometteurs grâce à un paradigme expérimental similaire à celui utilisé dans la présente étude, et ce particulièrement chez les enfants.

Pour ce qui est de l'évaluation de la source du récit, les études sur le "source monitoring" suggèrent plusieurs critères qui seraient dignes d'être évalués. L'efficacité de certains de ces critères a même été démontrée empiriquement dans des contextes légèrement différents (voir Stern et Dunning, 1994). Il est donc temps de réunir ces domaines de recherche inutilement cloisonnés pour leur permettre de s'enrichir mutuellement.

Finalement, au niveau professionnel, des sommes importantes sont en jeu, sous forme d'honoraires. La légitimité d'une telle procédure dans les systèmes juridiques québécois, canadien et nord-américain doit être clarifiée étant donné que l'évaluation de la crédibilité d'un témoin ou d'un témoignage est généralement considérée comme étant le rôle du juge et du jury. Est-ce que l'arrivée de professionnels de la santé réclamant ce rôle servirait réellement la justice? Ou est-ce que ce sont plutôt des intérêts professionnels et pécuniaires qui seraient servis par une telle transformation, au détriment de l'intérêt des citoyens? La question doit être posée et le débat doit être publique.

BIBLIOGRAPHIE

- Anson, David A., Golding, Stephen L. et Gully Kevin J. (1993). Child sexual allegations: Reliability of criteria-based content analysis. Law and Human Behavior, 17, 3, 331-341.
- Arntzen, F. (1982). Dans Trankell, A. (Éd.), Reconstructing the past: The role of psychology in criminal trials. Norstedt, Stockholm, p 107-120.
- Arntzen, F. (1983). Psychologie der Zeugenaussage: System der Glaubwürdigkeitsmerkmale. (2nd Edition). Munchen: Beck.
- Bradford, Roger. (1994). Developing an objective approach to assessing allegations of sexual abuse. Child Abuse Review, 3, 2, 93-101.
- Boychuk, Tascha D. (1991). Criteria-based content analysis of children's statements about sexual abuse. Thèse de doctorat non-publiée. Arizona State University, Tempe.
- Casoni, Dianne. (1999). Les grilles de validation de témoignages d'enfants: Analyse critique. Dans Louis Brunet (Éd.), L'expertise psycholégale: balises méthodologiques et déontologiques. Sainte-Foy, Québec: Presses de l'Université du Québec.
- Ceci, Stephen J., Toglia, Michael P., et Ross, David F. (1987). Age differences in suggestibility: Narrowing the uncertainties. Dans Stephen J. Ceci, Michael P. Toglia, et David F. Ross (Éds.) Children's eyewitness memory. New York, NY: Springer-Verlag.
- Craig, Ron A., Scheibe, Rick, Raskin, David C., Kircher, John C. et Dodd, David H. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. Applied Developmental Science, 3, 2, 77-85.
- Davies, Graham M. (1994). Statement validity analysis: An art or a science? Commentary on Bradford. Child Abuse Review, 3, 2, 104-106.

- Day, D.J.A. (1998). Psychological correlates of the UFO abduction experience: the effects of beliefs and indirect suggestions on abduction accounts during hypnosis (thèse de doctorat non-publiée). Montréal: Université Concordia.
- Dittmann, Volker. (1998). L'expertise de la fiabilité du point de vue psychologico-psychiatrique dans les cas de soupçons d'abus sexuels. Tiré à part de Jeunesse et Droit Pénal, vol.16. Zurich, Suisse: Éd. Rüegger. 13p.
- Doris, John. (1994). Commentary on criteria-based content analysis. Journal of Applied Developmental Psychology, 15, 281-285.
- Ekman, Paul. (1992). Telling lies. New York, NY: Norton.
- Ekman, Paul, et Friesen, Wallace. V. (1969). Non-verbal-leakage and clues to deception. Psychiatry, 32, 1, 88-105.
- Esplin, Phillip W., Boychuk, Tascha D., et Raskin, David C. (1988, Juin). A field validity study of criteria-based content analysis of children's statements in sexual abuse cases. Article présenté au NATO Advanced Study Institute on Credibility Assessment à Maratea, Italie.
- Esplin, Phillip W., Houed, T., et Raskin, David C. (1988, Juin). Applications of statement validity assessment. Article présenté au NATO Advanced Study Institute on Credibility Assessment à Maratea, Italie.
- Gunn, J., et Gudjonsson, G. (1988). Using the psychological stress evaluator in conditions of extreme stress. Psychological Medicine, 18, 235-238.
- Hall, Calvin S., et Van de Castle, Robert L. (1966). The content analysis of dreams. New York, NY: Appleton-Century-Crofts.
- Hershkowitz, Irit. (1999). The dynamics of interviews involving plausible and implausible allegations of child sexual abuse. Applied Developmental Science, 3, 2, 86-91.
- Hershkowitz, Irit, Lamb, Michael E., Sternberg, Kathleen J. et Esplin, Phillip W. (1997). The relationships among interviewer utterance type, CBCA scores and the richness of children's responses. Legal and Criminological Psychology, 2, 2, 169-176.

- Hollien, Harry. (1990). The acoustics of crime: The new science of forensic phonetics. New York, NY: Plenum Press.
- Honts, Charles. R., et Perry, Mary V. (1992). Polygraph admissibility: Changes and challenges. Law and Human Behavior, 16, 357-379.
- Honts, Charles. R., Raskin, David. C., et Kircher, John. C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. Journal of Applied psychology, 17, 252-259.
- Horowitz, Steven W. (1991). Empirical support for statement validity assessment. Behavioral Assessment, 13, 293-313.
- Horowitz, Steven W. (1998). Reliability of criteria-based content analysis of child witness statements: Response to Tully. Legal and Criminological Psychology, 3, 2, 189-191.
- Horowitz, Steven W., Lamb, Michael E., Esplin, Phillip W., Boychuk, Tascha D., Krispin, Orit et Reiter-Lavery, Lisa. (1997). Reliability of criteria-based content analysis of child witness statement. Legal and Criminological Psychology, 2, 1, 11-21.
- Joffe, R., et Yuille, John C. (1992, mai). Criteria-based content analysis: An experimental investigation. Affiche présentée au NATO Advanced Study Institute on the Child Witness in Context: Cognitive, Social, and Legal Perspectives, Lucca, Italie.
- Johnson, Marcia K., Kahan, T. I., et Raye, C. L. (1984). Dreams and reality monitoring. Journal of Experimental Psychology: General, 113, 329-344.
- Johnson, Marcia K., et Raye, C. L. (1981). Reality monitoring. Psychological bulletin, 88, p 67-85.
- Johnson, Marcia K., Hashtroudi, S. et Lindsay, D. S. (1993). Source monitoring. Psychological Bulletin, 114, 1, 3-28.
- Jones, David P. H. (1994). Objective approaches to assessing allegations of sexual abuse: Commentary on Bradford. Child Abuse Review, 3, 2, 101-104.

- Jung, Carl Gustav. (1987). L'homme à la découverte de son âme. Paris: Éd. Albin Michel.
- Lamb, Michael E. (1998). Mea culpa but caveat emptor! Response to Tully. Legal and Criminological Psychology, 3, 2, 193-194.
- Lamers-Wilkelman, F. et Buffing, F. (1996). Children's testimony in the Netherlands: A study of statement validity analysis. Criminal Justice and Behavior, 23, 2, 304-321.
- Landry, Kristine L. et Brigham, John C. (1992). The effect of training in criteria-based content analysis on the ability to detect deception in adults. Law and Human Behavior, 16, 6, 663-676.
- Laurence, Jean-Roch. (septembre 1999). Communication personnelle.
- Lilova, Iliana. (2000). The validity of statement validity analysis. (projet de baccalauréat non-publié). Montréal, Canada: Université Concordia.
- Porter, Stephen et Yuille, John C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. Law and Human Behavior, 20, 4, 443-458.
- Raskin, David C. et Esplin, Phillip W. (1991). Statement validity assessment: Interview procedures and content analysis of children's statements of sexual abuse. Behavioral Assessment, 13, 265-291.
- Raskin, David C., et Esplin, Phillip W. (1991a). Assessment of children's statements of sexual abuse. Dans John Doris (Éd.), The suggestibility of children's recollections (p. 168-171). Washington, DC: American Psychological Association.
- Raskin, David C., et Esplin, Phillip W. (1991b). Commentary: Response to Wells, Loftus, and McGough. Dans John Doris (Éd.), The suggestibility of children's recollections (p. 168-171). Washington, DC: American Psychological Association.
- Raskin, David C., Esplin, Phillip W., et Horowitz, Steven W. (1991). Investigative interviews and assessments of children in sexual abuse cases. Unpublished manuscript, Department of Psychology, University of Utah.

- Ruby, Charles. L. et Brigham, John C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations: A critical review. Psychology, Public Policy, and Law, 3, 705-737.
- Ruby, Charles. L. et Brigham, John C. (1998). Can CBCA distinguish true and false statements in black speakers? Law and Human Behavior, 22, 4, 369-388.
- Sapir, A. (1987). The LSI course on scientific content analysis (SCAN). Phoenix, Arizona: Laboratory for scientific interrogation.
- Spanos, N. P., Burgess, C. A. et Burgess, M. F. (1994). Past-life identities, UFO abductions, and satanic ritual abuse: The social construction of memories. The International Journal of Clinical and Experimental Hypnosis, 42, 433-445.
- Steller, Max. (1989). Recent development in statement analysis. Dans John C. Yuille (Ed.), Credibility assessment. (p 135-154) Dordrecht, The Netherlands: Kluwer.
- Steller, Max, et Köhnken, Guenter. (1989). Criteria-based content analysis. Dans David C. Raskin (Ed.), Psychological methods in criminal investigation and evidence (p 217-245). New York, NY: Springer-Verlag.
- Stern, Lisa Beth et Dunning, David. (1994). Distinguishing accurate from inaccurate eyewitness identifications: A reality monitoring approach. Dans David Frank Ross, J. Don Read, et al. (Éds). Adult eyewitness testimony: Current trends and developments. New York, NY: Cambridge University Press.
- Szewcyk, H. & Frohlich, H. H. (1984). Die psychologische Begutachtung der Glaubwürdigkeit. Dans H. Dettborn, H Frolich, et H. Szewcyk (Éds.). Forensische Psychologie. (pp.289-333). Berlin: VEB-Verlag.
- Trankell, A. (1971). Der Relitatzgehalt von Zeugenaussagen: Methoden der Aussagepsychologie. Gottingen: Vandenoek Ruprecht.
- Tully, Bryan. (1998). Reliability of criteria-based content analysis of child witness statements: Cohen's kappa doesn't matter. Legal and Criminological Psychology, 3, 2, 183-188.

- Undeutsch, Udo. (1967). Beurteilung der Glaubwürdigkeit von Aussagen. Dans Udo Undeutsch (Éd.), Handbuch der Psychologie: Vol. II: Forensische Psychologie. (pp.26-181). Gottingen: Hogrefe.
- Undeutsch, Udo. (1989). The development of statement reality analysis. Dans John C. Yuille (Éd.), Credibility assessment. (p 101-120) Dordrecht, The Netherlands: Kluwer.
- Vrij, Aldert, Kneller, Wendy et Mann, Samantha. (2000). The effect of informing liars about criteria-based content analysis on their ability to deceive CBCA-raters. Legal and Criminological Psychology, 5, 1, 57-70.
- Wells, Gary L., et Loftus Elizabeth (1991). Commentary: Is this child fabricating? Reaction to a new assessment technique. Dans John Doris (Éd.), The suggestibility of children's recollections (p. 168-171). Washington, DC: American Psychological Association.
- Yuille, John C. (1988). The systematic assessment of children's testimony. Psychologie Canadienne, 29, 3, 247-262.
- Yuille, John C.(1989). Credibility assessment. Dordrecht, The Netherlands: Kluwer.
- Yuille, John. C. (1990). SVA: Critères relatifs au contenu pour l'analyse de la déclaration. (document traduit de l'anglais au français par Hubert Van Gijseghem). Document non-publié. Vancouver, Canada: University of British Columbia. 13p.
- Yuille, John. C. (1990a). L'analyse de la validité de la déclaration (SVA): Liste de vérification. (document traduit de l'anglais au français par Hubert Van Gijseghem). Document non-publié. Vancouver, Canada: University of British Columbia. 4p.

ANNEXE A

Description des éléments contenus dans la liste de vérification (Validity Checklist).

L'ANALYSE DE LA VALIDITÉ DE LA DÉCLARATION (SVA)

LISTE DE VÉRIFICATION

JOHN C. YUILLE¹

University of British Columbia

¹Traduction de Hubert Van Gijseghem, Université de Montréal

NOTE: Certains éléments de cette liste de vérification ont été adaptés à partir de la Liste de Vérification de Steller, Raskin, Yuille & Esplin.

ÉLÉMENTS ISSUS DE L'ENTREVUE MÊME

A. LE COMPORTEMENT DE L'ENFANT

1. Le langage:

Est-ce que le langage utilisé par l'enfant était conforme à son niveau de développement? Si l'enfant a utilisé des termes (ex: pour indiquer des parties du corps, ou pour décrire des activités sexuelles) qui s'avèrent inappropriés, ou a-t-il appris ces termes? Il est important de savoir, afin de confirmer l'existence de l'abus, si ces termes ont été appris dans le contexte de l'abus. Par ailleurs, si les termes proviennent de d'autres sources, cela peut indiquer que l'enfant a subi une forme quelconque de "coaching".

2. Le savoir:

Est-ce que les connaissances de l'enfant concernant la sexualité étaient conformes à son niveau de développement? Sinon, comment l'enfant a-t-il acquis ces connaissances? Cet item peut augmenter la crédibilité si ce savoir a été acquis dans le contexte de l'abus. Par contre cela diminue la crédibilité si le savoir provient d'un "coaching" ou de d'autres sources.

3. L'affect:

La présence d'un affect approprié donne de la crédibilité aux allégations. Un affect inapproprié (i.e. des rires lors de la description d'un incident pénible) demande d'être investigué quant aux raisons de la présence d'un tel affect. L'absence d'affect peut diminuer la crédibilité de la déclaration. Cependant, un affect plat peut être la conséquence d'un abus

sexuel. Si peu d'affect est présent, l'interrogateur devrait vérifier les sentiments de l'enfant par rapport à un présumé abus. Aussi, un enfant qui a été interrogé à plusieurs reprises peut démontrer un affect plat par rapport à l'abus sexuel.

4. Expressions gestuelles spontanées:

Lors de la description de l'incident abusif un enfant peut spontanément utiliser des gestes expressifs. Par exemple, un enfant qui aurait été attaché lors de l'abus, pourrait lever ses bras pour représenter la posture dans laquelle il était. La présence d'une telle gestuelle augmente la crédibilité. Par contre l'absence de cette gestuelle ne diminue la crédibilité que dans le cas d'un enfant gestuellement expressif par ailleurs.

5. La susceptibilité à la suggestion:

Si l'enfant apparaît avoir été influencé, celui qui évalue la déclaration doit examiner les contenus et les types d'entrevues auxquels l'enfant a été préalablement soumis.

6. Croquis:

Si l'enfant a produit un croquis des parties du corps (particulièrement des organes génitaux) ou de positions sexuelles, cela devrait être considéré sous l'item 2 mentionné ci-dessus. Des croquis indiquant la position des objets (l'arrangement de l'ameublement dans une pièce) peut augmenter la crédibilité si l'enfant n'a eu accès à cette pièce que lors du présumé abus. Sous aucune considération l'investigateur ne devrait interpréter symboliquement les dessins de l'enfant.

7. Comportement avec les poupées:

Si, lors de l'entrevue l'enfant a utilisé des poupées pour l'aider à décrire l'abus, cela devrait être considéré sous l'item 2 mentionné ci-dessus. Les poupées ne peuvent être considérées comme élément d'évaluation de la crédibilité qu'en autant qu'il est clairement établi qu'elles n'ont pas été utilisées d'une manière suggestive. Par exemple, s'il n'y avait pas eu de dévoilement avant l'utilisation des poupées, leur utilisation pour évaluer la crédibilité de l'enfant s'avère TRÈS problématique. Cependant, s'il y avait eu dévoilement avant l'utilisation des poupées, alors le comportement de l'enfant avec ces poupées peut aider à la

validation de l'item 2.

8. Comportement sexualisé:

L'enfant peut démontrer des comportements sexualisés vis-à-vis l'interrogateur ou même envers lui-même. Quoique cela ne soit pas fréquent, un tel comportement peut indiquer que cet enfant a subi un comportement sexualisé de la part d'un adulte. Il est important de s'assurer que l'interrogateur n'a pas suscité un tel comportement.

B. CARACTÉRISTIQUES DE L'ENTREVUE

9. La conformité de l'entrevue:

Est-ce que l'entrevue a été menée d'une manière appropriée? Est-ce que l'interrogateur a établi une relation avec l'enfant? L'enfant pouvait-il relater les faits dans ses propres mots? L'interrogateur semblait-il avoir un "agenda caché" lors de l'entrevue? S'il y a eu des aspects inadéquats dans l'entrevue, vont-ils jusqu'à compromettre l'utilisation de l'Analyse de la Validité de la Déclaration (SVA)?

10. Questions successives et directives:

Est-ce que l'interrogateur a utilisé des questions suggestives ou directives? Leur utilisation si tel est le cas, est-elle suffisante à compromettre l'utilisation de l'Analyse de la Validité de la Déclaration (SVA)?

11. Pression et coercition:

Y-a-t-il eu utilisation de moyens de pression ou de coercition lors de l'entrevue? Est-ce suffisant à compromettre l'utilisation de l'Analyse de la Validité de la Déclaration (SVA)?

CONSIDÉRATIONS CONCERNANT LES MOTIFS DU DÉVOILEMENT

12. Contexte du dévoilement initial:

La validité de la déclaration est rehaussée si le premier récit de l'abus a été livré spontanément. (i.e. raconté à un parent, un professeur ou un médecin (etc.) n'étant pas en conflit avec le présumé abuseur).

Si le dévoilement initial provient d'un parent et s'inscrit dans le contexte d'une séparation ou dans les cas de disputes concernant la garde des enfants, ou si il y a quelques indices que l'accusation est faite dans un but quelconque, (i.e. vengeance d'un enfant envers un professeur), alors la validité doit être remise en question.

13. Pressions pour dévoiler:

Y a-t-il des indices selon lesquels l'enfant a subi des pressions pour faire le dévoilement initial? Les pressions peuvent être intentionnelles ou non, elles peuvent aller de questionnements répétitifs aux menaces si l'enfant ne dévoile pas et ce malgré le fait que l'enfant nie.

AUTRES ÉVIDENCES

14. Évidences médicales:

Y a-t-il des évidences de transformations physiques, blessures ou maladie qui iraient dans le sens des allégations?

15. Compatibilité avec d'autres déclarations de l'enfant:

Quel est le lien entre cette déclaration et les autres déclarations de l'enfant? Des variations en ce qui a trait aux détails périphériques augmentent la crédibilité. Des variations dans les détails centraux suscitent des interrogations concernant la validité.

16. Compatibilité avec des déclarations faites par d'autres témoins:

Quel est le lien entre la déclaration de l'enfant et les déclarations provenant de d'autres témoins?

17. Évidences matérielles:

Y a-t-il des évidences matérielles (i. e: linges, photos, jouets sexuels,. etc.) qui supportent ou diminuent la validité de la déclaration?

18. Évidences comportementales compatibles avec l'abus:

Ce que l'on appelle les indices subtiles de l'abus (i.e. changements dans les rythmes de sommeil, de nutrition, pseudomaturité, agir sexuel, etc.) ne peuvent être considérés pour démontrer qu'il y a eu abus. Cependant, si nous concluons que la déclaration est valide, ces comportements peuvent être perçus comme compatibles avec l'abus.

ANNEXE B

Définitions des critères du CBCA tels qu'utilisés dans l'étude.

Le texte principal définissant les dix-neuf critères provient d'un document de John Yuille (1990) utilisé comme outil de référence par les experts formés dans l'utilisation de l'analyse de la validité de la déclaration. Les références à des textes de langue allemande qui y sont faites ont été reprises du document de Yuille et sont incluses dans la bibliographie du présent travail simplement pour procurer l'information au lecteur. Lorsque des éléments de définition supplémentaires provenant d'autres sources sont jugés nécessaires, ils sont inclus à la fin de la définition entre crochets. De plus, lorsque le terme utilisé dans le présent document pour décrire les nombreux critères diffère de celui employé dans le document de Yuille (1990), il est écrit entre crochets à côté de l'autre terme. En plus de ces critères, trois règles proposées par Steller (1989) ont été utilisées: (1) la répétition d'un même élément dans différentes parties de la déclaration n'augmente pas la note au critère qui y est relié, (2) un énoncé peut remplir plus d'un critère à la fois, et (3) seulement le contenu de la déclaration relié à l'incident rapporté est considéré. Les quatre sous-titres en majuscules correspondent aux catégories générales utilisées dans le document.

CARACTÉRISTIQUES GÉNÉRALES DE LA DÉCLARATION

1. Cohérence: [Structure logique]

Ce critère est présent quand les différents éléments de la déclaration de l'enfant font un tout cohérent et consistant.

Undeutsch (1967, p.138) utilise les termes de cohérence interne et de consistance pour définir ce critère, faisant référence au fait que la déclaration dans son ensemble doit "se tenir". La déclaration devrait être cohérente sur le plan logique et avoir une consistance interne suffisante.

Trankell (1971, p.126) préfère le terme "homogénéité", indiquant qu'un récit est homogène quand "Il peut être démontré que les détails indépendants d'un même témoignage décrivent tous la même suite d'événements".

Arntzen (1983, p.51) définit le concept d'homogénéité comme étant cet aspect d'une déclaration "par lequel les détails contextuels font un tout, et ce, sans contradictions ou inconsistances. De plus ces détails peuvent être mis ensemble pour former un tout cohérent dans lequel les contenus du témoignage apparaissent eux aussi consistants.

La cohérence est un indice de crédibilité. Le fait qu'une déclaration puisse faire référence à des complications inattendues (critère 7) et/ou à des détails inusités (critère 8) n'en diminue pas pour autant sa cohérence.

2. Verbalisations spontanées: [Production non-structurée]

Le rappel "tout en ayant les caractéristiques de la cohérence" devrait se faire de manière non-structurée et spontanée. Par exemple, chacun des rappels de l'événement devrait contenir quelques corrections spontanées, "comme l'ajout ou l'élimination de détails périphériques selon l'événement ponctuel rappelé". Notons qu'une déclaration non crédibile se caractérisera par une reproduction figée d'un rappel à l'autre.

La façon dont l'enfant livre son témoignage nous donne un indice de la crédibilité. Une déclaration crédible est celle:

- dont la partie du libre rappel est livrée de façon spontanée. Une déclaration fabriquée sera plus susceptible d'avoir une structure rigide, et répétera de façon identique les descriptions de l'événement.

Artzen (1983, p.80) utilise le terme "verbeux" pour décrire un "témoignage incontrôlé", un témoignage qui se présente d'une manière "incohérente, sans suite et désorganisée. Les faits sont fragmentés et éparpillés à travers la déclaration et surgissent d'une manière désordonnée; Il y a des sauts dans l'évocation des époques et la séquence des événements racontés ne suit pas l'ordre chronologique. Les déclarations fabriquées présentent les événements sous forme chronologique.

Malgré le fait qu'une déclaration non préparée et désorganisée est une indication de crédibilité, les parties de la déclaration doivent pouvoir se rassembler d'une façon consistante, de telle sorte qu'elle réponde au critère 1.

[Raskin et Esplin (1991) considèrent aussi les digressions et les changements spontanés de direction dans la déclaration ("spontaneous shifts of focus") comme des indices de production non-spontanée.]

3. Détails en quantité suffisante: [Quantité de détails]

Toute chose demeurant égale par ailleurs, plus il y a de détails dans une déclaration, plus elle est susceptible d'être crédible. Les fausses déclarations sont habituellement pauvres en détails spécifiques, et le fait de demander à l'enfant de donner plus de détails n'apportera habituellement que fort peu de détails supplémentaires. Par ailleurs, nous devons tenir

compte du niveau de développement de l'enfant, puisque les très jeunes enfants livrent moins de détails concernant un événement que les enfants plus âgés. L'évaluation doit alors prendre en considération et refléter les habiletés cognitive, expressive et mnésique de l'enfant. On peut évaluer la capacité de l'enfant à livrer des détails lors de la phase de mise en relation.

Selon Arntzen (1983, p.27) le postulat qui sous-tend ce critère est "qu'il est impossible pour la majorité des témoins d'embellir un faux témoignage avec de nombreux détails". En d'autres mots, plus la déclaration de l'enfant contient de détails, plus il est facile de comparer ces détails avec d'autres informations disponibles. Une autre indication de crédibilité est l'ajout de détails périphériques aux détails centraux.

Il est à noter que la répétition d'un même détail ne confirme pas ce critère; ce ne sont que les faits et détails nouveaux qui contribuent à le confirmer.

CONTENUS SPÉCIFIQUES DE LA DÉCLARATION

4. L'enchâssement contextuel:

Il arrive parfois que le rappel d'un événement contienne des informations concernant le moment et/ou l'endroit où s'est produit l'événement. L'enchâssement d'un événement dans son contexte spatio-temporel est un indice de crédibilité. Undeutsch (1967, p.139) remarque que les "incidents réels reposent sur des assises spatio-temporelles". Szewczyk & Frohlich (1984, p.313) notent que si on peut insérer les contenus d'une déclaration dans un plus large contexte, cela ajoute certes à sa validité. Arntzen (1983, p.35) indique qu'une déclaration crédible contient "différents entrelacements entre les faits essentiels de l'événement, et les circonstances extérieures et situationnelles". Ces circonstances sont faites "des faits quotidiens, des habitudes de vie, de la vie familiale, du cercle d'amis, des relations avec les voisins...". L'enchâssement contextuel consiste à associer un événement à, par exemple, un anniversaire de naissance, à une visite quelconque, ou à tout autre moment ou endroit qui rehausserait la crédibilité de la déclaration.

Il est à noter que ce ne sont pas toutes les déclarations crédibles qui s'inscriront sous ce critère, car un événement aurait pu se produire à un endroit ou en un temps qui ne retient pas l'attention.

5. Descriptions d'interactions:

L'abus sexuel est un événement qui met en scène deux acteurs: l'abuseur et l'abusé. Ainsi plus de détails une victime sera en mesure de livrer concernant les gestes posés, plus il sera aisé de comprendre ce qui s'est passé exactement. "Ce critère fait référence aux interactions entre le témoin et la personne sur qui porte le témoignage, et, particulièrement, aux liens qui retiennent leurs gestes et leurs conversations (Arntzen, 1983, p.34)". Cette citation met donc l'emphase sur l'enchaînement séquentiel action-réaction. Tout dévoilement d'abus sexuel contiendra probablement des descriptions d'interactions. Toutefois, afin de répondre à ce critère, la description doit être suffisamment élaborée pour décrire la séquence des faits et gestes lors de l'événement.

[Raskin et Esplin (1991) précisent qu'un minimum de trois éléments d'interaction doivent être présents pour que le critère soit rempli.]

6. Rappel de conversation:

On rencontre ce critère quand l'enfant indique le mot à mot d'une conversation lors de l'événement. Par exemple, si l'enfant rapporte cette phrase provenant de l'abuseur: "Enlève tous tes vêtements, maintenant!", cela répondrait au critère. Par ailleurs si l'enfant rapporte la phrase suivante: "Il m'a dit d'enlever mon linge.", cela répondrait plutôt au critère #5.

Ce critère est particulièrement appuyé si l'enfant rapporte des mots qui ne font pas partie de son vocabulaire habituel.

7. Références à des complications inattendues:

On retrouve ce critère quand l'enfant rapporte des interruptions inattendues lors de l'incident (Undeutsch, 1967, p.153; Arntzen, 1983, p.33; Szewczyk & Frohlich, 1984, p.313). Par exemple, on retrouverait ce critère si l'enfant rapporte que lors de l'incident quelqu'un a frappé à la porte, ou que le téléphone a sonné, ou que l'abuseur est tombé en bas du lit, ou encore que l'incident s'est interrompu brusquement etc. La quantité de détails concernant une complication inattendue donne force à ce critère (tout en contribuant au critère #3).

8. Détails inusités:

On retrouve ce critère quand l'enfant inclut des détails inusités ou singuliers dans sa déclaration, par exemple, si l'enfant décrit un drôle de vêtement ou un objet connu mais utilisé de manière inhabituelle. Par ailleurs, ces détails inusités ne devraient pas manquer

de réalisme. Les enfants qui rapportent un témoignage fabriqué d'avance sont peu susceptibles d'être créateur au point d'inclure des détails réalistes mais inusités. La crédibilité est augmentée quand des détails inusités sont apportés de telle façon qu'ils n'auraient pu être inventé par un témoin ou encore par quelqu'un qui voulait influencer l'enfant (Undeutsch, 1989, p.112).

Plusieurs publications parlent de ce critère (Undeutsch, 1967, p.138; Trankell, 1971, p.226; Arntzen, 1983, p.39; Szewcyk & Frohlich, 1984, p.313), et on le définit comme suit: "ce sont des détails étranges qui sont tellement inhabituels qu'il est alors peu probable qu'on les retrouve dans les fausses accusations".

9. Détails périphériques: [Détails superflus]

Tout comme dans le cas du critère 8. Il est peu probable de retrouver des détails périphériques dans les déclarations inventées. Ce critère réfère à l'inclusion de détails extérieurs à l'événement comme tel. Ces détails sont concrets et percutants, donnant à croire que la déclaration est le rappel d'un événement réel. Contrairement au critère 8, ces détails périphériques n'ont pas un caractère inusité.

[Raskin et Esplin (1991) précisent que si un élément satisfait un autre des critères spécifiques (c'est-à-dire #4 à #18) il n'est probablement pas superflu.]

10. Détails non compris mais rapportés de façon exacte:

Ce critère est un excellent indicateur quant à la validité de la déclaration. Si la déclaration contient un récit précis de l'événement tout en montrant que l'enfant a mal compris ou mal interprété l'événement, cela peut particulièrement valider la déclaration. Undeutsch croit que ce critère fait: "référence à des détails dont l'invention serait impossible et au-delà des capacités et de la compréhension des témoins". (Undeutsch, 1989, p.114). (Trankell, 1971, p.25). Par exemple, l'enfant pourrait dire que l'abuseur était en souffrance car il gémissait. Un enfant, selon sa propre interprétation, pourrait décrire un orgasme, par exemple, comme un éternuement.

11. Références à des incidents extérieurs:

Lors de la déclaration, un enfant pourrait faire des références à des événements extérieurs aux incidents en question. Par exemple, une fille pourrait dire que son père (le suspect) lui a demandé si elle avait déjà fait la "même chose" avec quelqu'un d'autre. Ce type d'information est associé au contexte général de l'abus sexuel mais ne fait pas partie

intégrante de l'événement comme tel. Pour que ce critère puisse être coté, la référence extérieure doit de quelque façon être associée à l'abus.

12. Références à ses propres états psychologiques: [...états subjectifs]

Ce critère fait référence à l'état émotionnel et/ou cognitif de la victime. Si l'enfant dit qu'il a ressenti de la peur, du dégoût, de l'embarras ou une autre émotion subjective, ceci est un indice de validité (Undeutsch, 1989, p.114; Arntzen, 1983, p.29; Frohlich, 1984, p.314).

Aussi, le critère peut être coté si l'enfant raconte avoir eu des idées ou des images autour d'une possible fuite ou résistance.

[Raskin et Esplin (1991) incluent toute émotion ou pensée vécue au moment de l'événement.]

13. Attribution d'un état psychologique à l'abuseur: [...d'états subjectifs...]

Une autre indication de la validité de la déclaration est quand l'enfant rapporte des perceptions concernant l'état psychologique du suspect lors de l'incident (Szewczyk & Frohlich, 1984, p.314). De tels propos incluent "les réactions émotives ou l'état mental de l'abuseur".

CONTENUS RELATIFS AUX MOTIVATIONS DE LA DÉCLARATION

14. Corrections spontanées:

La crédibilité d'un enfant est augmentée quand il fait des corrections spontanées lors du rappel de l'événement. Les corrections font en sorte de diminuer la probabilité d'une déclaration fictive. De plus, si les corrections permettent d'améliorer la qualité de la déclaration (selon les autres critères) les corrections spontanées peuvent permettre une meilleure évaluation de la validité de la déclaration. Undeutsch (1967, p.152) a accordé une importance considérable aux corrections spontanées. Un enfant qui raconte une fausse déclaration tente habituellement de donner bonne impression et ne risquera pas de se remettre en question en modifiant son récit.

Szewczyk & Frohlich (1984, p.313) ont noté que ce critère ne peut pas être pris en considération si la correction n'est pas spontanée mais plutôt la réaction à des questions précises. Ainsi, si l'interrogateur fait ressortir des contradictions dans la déclaration de l'enfant, une correction qui s'en suivrait ne rencontrerait pas ce critère.

[Raskin et Esplin (1991) considèrent aussi les informations ajoutées à du matériel présenté précédemment dans la déclaration.]

15. Aveu de "blancs" de mémoire: [Aveu d'absence de souvenir ou de connaissance]

Que ce soit spontanément ou en réponse à des questions spécifiques, l'enfant pourrait avouer son ignorance quant à certains détails. Des "blancs de mémoire" concernant certains aspects d'un événement sont normaux dans une déclaration valide, mais il est peu probable qu'un enfant qui fait une déclaration fabriquée admette spontanément de tels blancs de mémoire. Toutefois, il arrive que des rappels fabriqués se caractérisent par la pauvreté des détails (Critère #3). Dans ce cas, le présent critère ne s'applique pas. Il ne s'applique que lorsque l'enfant a produit un rapport valide et qu'il admet occasionnellement des "blancs de mémoire".

[Raskin et Esplin (1991) incluent aussi les aveu d'absence de connaissance à propos d'un aspect de l'incident.]

16. Doutes à propos de sa propre déclaration:

Undeutsch (1967, p.153) fait remarquer qu'une personne produisant une fausse déclaration tentera d'avoir l'air crédible, et ne laissera planer aucun doute concernant sa propre déclaration. Donc, le fait d'avoir des doutes concernant sa propre déclaration est un indice de crédibilité.

[Raskin et Esplin (1991) spécifient que les doutes peuvent autant concerner la véracité que la crédibilité de la déclaration.]

17. Désapprobation de sa propre participation:

L'insertion de détails incriminants ou défavorables rend crédible la déclaration. Selon Undeutsch (1967, p.153) et Szewcyk & Prohlich (1984, p.314), l'auto-désapprobation de la part d'un témoin concernant "des gestes supposément mauvais qu'il aurait posés envers l'abuseur" est un indice de validité, car de telles "confessions" ne se retrouvent pas dans un témoignage fabriqué (là où le but est d'incriminer le suspect).

Remarque: Normalement, une victime ne désapprouve point sa propre conduite. Ce critère est inclus pour donner du poids à une telle remarque, là où elle apparaîtrait.

18. Le fait d'excuser l'abuseur:

Un autre indice de validité réside dans le cas où l'enfant tente d'expliquer ou d'excuser le comportement du suspect (Undeutsch, 1967, p.153). Il n'y a aucune raison pour laquelle un enfant excuserait l'abuseur, mais ici, comme au critère # 17, l'exonération donne du poids à la crédibilité si elle est apparue spontanément.

Il est important de noter que l'absence d'exonérations n'est pas un signe que la déclaration est fausse.

ÉLÉMENTS SPÉCIFIQUES CONCERNANT LE DÉLIT

19. Caractéristiques spécifiques du délit:

Il existe différentes croyances populaires concernant l'abus sexuel qui sont pourtant contraires au déroulement habituel de ce type de délit. L'analyse de la déclaration exige une connaissance du patron habituel de l'abus sexuel, mais aussi des patrons atypiques. Par exemple, l'inceste débute habituellement par un abus "bénin" (ex: toucher), et progresse graduellement vers des gestes plus graves qui sont accompagnés de changement dans l'attitude de l'abusé envers l'abuseur. L'évaluation peut être facilitée quand on retrouve dans une déclaration de telles références à des patrons connus par l'expert.

ANNEXE C
Formulaires de consentement

FORMULAIRE DE CONSENTEMENT POUR PARTICIPATION DANS UNE RECHERCHE
 Laboratoire d'hypnose de l'Université Concordia, 848-2213
 Professeur responsable: Dr Jean-Roch Laurence, 848-7555

J'accepte de participer à un programme de recherche réalisé par Dominic Beaulieu-Prévost de l'École de Psychologie à l'Université Laval pour son mémoire de maîtrise sous la direction de Dr Robert Rousseau et la co-direction de Dr Jean-Roch Laurence.

A. BUT

J'ai été informé que le but de la recherche est d'étudier l'utilité et l'efficacité des techniques d'analyse de contenu dans l'évaluation de la crédibilité de témoignages en contexte légal. Je comprends que la procédure exacte ne peut m'être révélé avant la fin de l'expérience.

B. PROCÉDURE

J'ai été informé que l'expérience est exécutée en une rencontre d'environ quarante minutes et qu'elle implique la procédure suivante: Je serai interrogé sur deux événements ou deux rêves, réels ou inventés, que j'ai eu à préparer auparavant. Avant chacun des deux interrogatoires, j'aurai environ 5min pour me préparer et me rappeler les détails du récit. La séance sera enregistrée pour assurer une transcription fidèle du narratif. À la fin de l'expérience, un dédommagement de cinq dollars me sera attribué, même si je ne termine pas la procédure. De plus, à moins que, durant un des deux interrogatoires, l'interviewer soit totalement convaincu que l'événement ou le rêve raconté ne soit jamais arrivé, je recevrai une compensation additionnelle de cinq dollars (soit dix dollars en tout). Si l'interviewer devient totalement convaincu que l'événement ou le rêve raconté n'est jamais arrivé, l'interrogatoire à propos de ce récit sera interrompu et seul le dédommagement de 5\$ me sera donné. J'ai donc comme instruction additionnelle de présenter mes deux récits comme étant réels, peu importe qu'ils soient réels ou inventés.

Je n'aurai aucune autre tâche à faire que celle ci-haut mentionnée. Il n'y a aucun bénéfice particulier relié à ma participation autre que d'avoir aidé à la recherche et il n'y a pas d'effet négatif connu sauf peut-être une légère fatigue à la fin de la rencontre. J'accepte que les informations relatives à ma participation soient utilisées dans cette recherche. J'ai été informé que mon nom ne sera pas associé à mes résultats et ne sera pas publié. Je comprends que ma participation dans cette recherche et le contenu des rencontres et des enregistrements sont strictement confidentiels et seront utilisés exclusivement dans le cadre de la recherche. Ces données ne seront accessibles qu'aux chercheurs et assistants de recherche directement impliqués dans cette recherche. De plus, les enregistrements seront détruits après une période maximale d'un an. Si les résultats de cette étude sont publiés, seuls les résultats de groupe, et non les résultats individuels, seront rapportés pour préserver la confidentialité de ma participation dans la recherche.

C. CONDITIONS DE PARTICIPATION

- Je comprends que je suis libre de refuser de participer à la recherche sans conséquences négatives.
- Je comprends que je suis libre de retirer mon consentement et de discontinuer ma participation en tout temps sans conséquences négatives.
- Je comprends que ma participation dans cette recherche est confidentielle.
- Je comprends que les résultats de cette recherche peuvent être publiés.
- Je comprends le but de cette recherche.

J'AI LU ATTENTIVEMENT LES INFORMATIONS CI-HAUT ET COMPRENDS CET ENGAGEMENT. JE CONSENS LIBREMENT ET ACCEPTE DE PARTICIPER DANS CETTE RECHERCHE.

NOM _____

SIGNATURE _____

DATE _____

FORMULAIRE DE CONSENTEMENT
 Laboratoire d'hypnose de l'Université Concordia, 848-2213
 Professeur responsable: Dr Jean-Roch Laurence, 848-7555

J'accepte de participer au programme de recherche réalisé par Dominic Beaulieu-Prévost de l'École de Psychologie à l'Université Laval pour son mémoire de maîtrise sous la direction de Dr Robert Rousseau et la co-direction de Dr Jean-Roch Laurence.

Le but de la recherche à laquelle j'ai participé est d'étudier l'utilité et l'efficacité des techniques d'analyse de contenu dans l'évaluation de la crédibilité de témoignages en contexte légal.

L'expérience fut exécutée en une rencontre d'environ quarante minutes et elle impliqua la procédure suivante: Je fus interrogé sur deux événements ou deux rêves, un réel et un inventé, que j'ai eu à préparer auparavant. Avant chacun des deux interrogatoires, j'ai eu environ 5min pour me préparer et me rappeler les détails de l'événement. La séance fut enregistrée pour assurer une transcription fidèle du narratif. À la fin de l'expérience, un dédommagement de dix dollars me fut attribué, même si je n'ai pas terminé la procédure. J'avais comme instruction additionnelle de présenter mes deux événements ou rêves comme étant réels, peu importe qu'ils soient réels ou inventés. Je n'ai eu aucune autre tâche à faire que celle ci-haut mentionnée. Je comprends que la condition pour recevoir le cinq dollars additionnel me fut donnée pour assurer une certaine motivation de ma part à être crédible mais qu'elle n'a affecté en aucune façon le dédommagement que j'ai reçu.

J'accepte que les informations relatives à ma participation soient utilisées dans cette recherche. J'ai été informé que mon nom ne sera pas associé à mes résultats et ne sera pas publié. Je comprends que ma participation dans cette recherche et le contenu des rencontres et des enregistrements sont strictement confidentiels et seront utilisés exclusivement dans le cadre de la recherche. Ces données ne seront accessibles qu'aux chercheurs et assistants de recherche directement impliqués dans cette recherche. De plus, les enregistrements seront détruits après une période maximale d'un an. Si les résultats de cette étude sont publiés, seuls les résultats de groupe, et non les résultats individuels, seront rapportés pour préserver la confidentialité de ma participation dans la recherche.

C. CONDITIONS DE PARTICIPATION

- Je comprends que je suis libre de retirer mon consentement et de discontinuer ma participation en tout temps sans conséquences négatives.
- Je comprends que ma participation dans cette recherche est confidentielle.
- Je comprends que les résultats de cette recherche peuvent être publiés.
- Je comprends le but de cette recherche.

J'AI LU ATTENTIVEMENT LES INFORMATIONS CI-HAUT ET COMPRENDS CET ENGAGEMENT. JE CONSENS LIBREMENT ET ACCEPTE DE PARTICIPER DANS CETTE RECHERCHE.

NOM

SIGNATURE

DATE

CONSENT FORM TO PARTICIPATE IN A RESEARCH PROJECT

Concordia University's hypnosis laboratory, 848-2213

Professor in charge: Dr Jean-Roch Laurence, 848-7555

I agree to participate in a research programme conducted by Dominic Beaulieu-Prévost from Laval University's psychology school for his master's thesis under the direction of Dr Robert Rousseau and the co-direction of Dr Jean-Roch Laurence.

A. PURPOSE

The purpose of the research is to study the usefulness and the efficacy of content analysis techniques to evaluate the credibility of self-reports in the legal context. I understand that the exact procedure cannot be explained before the end of the experiment.

B. PROCEDURE

I was informed that the experiment takes place in one session lasting approximately forty minutes and that it involves the following procedure: I will be interrogated about two events or two dreams, real or invented, that I had to prepare beforehand and write on a form for that purpose. I agree to bring this form with me to the session. Before each of the two interrogations, I will have about 5min to prepare myself and to remember the details of the story. The session will be tape-recorded to insure a reliable transcription of the narrative. At the end of the experiment, a five dollars compensation will be given to me, even if I don't finish the procedure. Furthermore, I will receive an additional five dollars compensation (for a total of ten dollars) unless, during one of the two interrogations, the interviewer becomes convinced that the event or dream I reported never happened. If the interviewer becomes convinced that the event or dream never happened, the interrogation about this story will be interrupted and I will only receive a five dollars compensation for the experiment. I must therefore report my two stories as if they are real, whether they are real or invented.

I won't have any other task to do than the one mentioned above. There are no specific benefits related to my participation other than helping research and there are no known negative effects or risks in this research. I agree that the information gained from my participation be used in this research. I understand that my name will remain confidential and will not be linked to my stories. I understand that my participation in this research and the content of the sessions and the tape recording are kept strictly confidential and will be used exclusively for the purposes of this research. Only the researchers and research assistants directly implicated in this research will have access to the data. Furthermore, the tape-recording will be destroyed after a maximum period of one year. The results of this study may be published. If so, only group results and not individual results will be reported to preserve the confidentiality of my participation in the research.

C. CONDITIONS OF PARTICIPATION

- I understand that I am free to refuse to participate to the research without negative consequences.
- I understand that I am free to remove my consent and discontinue my participation at anytime without negative consequences.
- I understand that my participation in this research is confidential.
- I understand that the results of this research may be published.
- I understand the purpose of this research.

I READ ATTENTIVELY THE INFORMATION ABOVE AND I UNDERSTAND THIS AGREEMENT. I FREELY CONSENT AND AGREE TO PARTICIPATE IN THIS RESEARCH.

NAME

SIGNATURE

DATE

CONSENT FORM

Concordia University's hypnosis laboratory, 848-2213
Professor in charge: Dr Jean-Roch Laurence, 848-7555

I agree to participate in a research programme conducted by Dominic Beaulieu-Prévost from Laval University's psychology school for his master's thesis under the direction of Dr Robert Rousseau and the co-direction of Dr Jean-Roch Laurence.

The purpose of the research to which I participated is to study the usefulness and the efficacy of content analysis techniques to evaluate the credibility of self-reports in the legal context.

The experiment took place in one session of approximately forty minutes and it involved the following procedure: I was interrogated about two events or two dreams, real or invented, that I had to prepare beforehand and write on a form for that purpose. I agreed to bring this form with me to the session. Before each of the two interrogations, I had about 5min to prepare myself and to remember the details of the story. The session was tape-recorded to insure a reliable transcription of the narrative. At the end of the experiment, a ten dollars compensation was given to me, even if I did not finish the procedure. I had to report my two stories as if they were real, whether they were real or invented. I had no other task to do than the ones mentioned above. I understand that the condition to receive the additional five dollars was told to me to insure that I was motivated enough to sound credible but that it never affected the way I was compensated for my participation.

I agree that the informations concerning my participation be used in this research. I understand that my name will remain confidential and will not be linked to my stories. I understand that my participation in this research and the content of the sessions and the tape recording are kept strictly confidential and will be used exclusively for the purposes of this research. Only the researchers and research assistants directly implicated in this research will have access to the data. Furthermore, the tape-recording will be destroyed after a maximum period of one year. The results of this study may be published. If so, only group results and not individual results will be reported to preserve the confidentiality of my participation in the research.

C. CONDITIONS OF PARTICIPATION

- I understand that I am free to remove my consent and discontinue my participation at anytime without negative consequences.
- I understand that my participation in this research is confidential.
- I understand that the results of this research may be published.
- I understand the purpose of this research.

I READ ATTENTIVELY THE INFORMATION ABOVE AND I UNDERSTAND THIS AGREEMENT. I FREELY CONSENT AND AGREE TO PARTICIPATE IN THIS RESEARCH.

NAME _____
SIGNATURE _____
DATE _____

ANNEXE D
Protocoles (anglais et français)

PROTOCOLE FRANÇAIS

Appel téléphonique

Est-ce que je pourrais parler à [nom] s.v.p.

Bonjour, mon nom est Dominic et je travaille au laboratoire d'hypnose de l'Université Concordia. J'ai été dans votre classe de _____ pour présenter une recherche en psychologie légale et vous avez signé la liste que j'ai fait circuler. J'aimerais savoir si vous êtes toujours intéressés à participer ou en savoir plus sur la recherche.

Pour vous aider, je peux vous résumer l'expérience.

L'expérience va se dérouler à l'Université Concordia en une séance d'environ quarante minutes. Elle sera principalement constituée d'une entrevue où il vous sera demandé de raconter deux événements ou deux rêves, soit réels ou inventés. Il vous sera aussi demandé d'être le plus crédible possible lorsque vous raconterez les deux récits. Cinq dollars de dédommagement vous seront attribués pour votre participation et un cinq dollars additionnel vous sera donné si, durant l'entrevue, l'interviewer trouve vos histoires crédibles.

Acceptez-vous toujours de participer à l'expérience?

Avant de se rencontrer, Dr Jean-Roch Laurence va vous envoyer une lettre avec les instructions à suivre pour choisir vos deux récits. Pouvez-vous me dire l'adresse à laquelle cette lettre doit être envoyée?

La lettre sera envoyée d'ici demain. Vous devrez donc la recevoir dans 3-4 jours. Il ne nous reste qu'à fixer le rendez-vous.

Voulez-vous mon numéro de téléphone au cas où vous vouliez me rappeler d'ici là?

Mon nom est Dominic et mon numéro à la maison est le 482-0353. Je vous remercie et je vais vous appeler la veille pour vous rappeler le rendez-vous et m'assurer que vous avez reçu et lu la lettre, d'accord?

(Envoi de la lettre avec la signature de Jean-Roch Laurence, voir annexe E pour exemple)

Première rencontre

1. Présentation

2. Lecture du premier formulaire de consentement et explication de ce que cela implique.

-informer que je ne sais pas si les événements racontés sont réels ou inventés.

3. Signature du formulaire et donner les numéros de téléphone:

-du laboratoire d'hypnose (848-2213)

-du bureau de Jean-Roch (848-7555)

-et mon numéro personnel (482-0353).

4. Entrevue.

-S'assurer que le participant a en sa possession la lettre qui lui a été envoyée. Sinon, lui en donner une copie pour qu'il puisse la consulter durant son 5min de préparation.

-Explication des instructions à suivre pour l'entrevue:

(Voir lettre envoyée aux participants) + Avez-vous des questions?

Je vais maintenant vous laisser 5min de préparation avant que l'interrogatoire commence officiellement. Avez-vous des questions avant que je parte?

-Partir 5min et donner copie de lettre envoyée-

Êtes-vous prêts?

J'aimerais maintenant que vous me décriviez de la façon la plus complète possible votre premier événement/rêve, en commençant au début.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.

Les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose comme "Vous avez dit X, qu'est-il arrivé après?" ou "Et après, qu'est-il arrivé?".)

Vous rappelez-vous d'autres choses par rapport à cet événement/rêve?

Est-ce que c'est tout?

J'aimerais maintenant que vous me décriviez à nouveau de la façon la plus complète possible l'événement/rêve dont nous venons de parler, en essayant de n'oublier aucun détail, même s'ils paraissent insignifiants.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.
les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose
comme "Vous avez dit X, qu'est-il arrivé après?" ou "Et après, qu'est-il arrivé?".)

Vous rappelez-vous d'autres choses par rapport à cet événement?

Est-ce que c'est tout?

OK.

Très bien. Nous allons maintenant passer au deuxième événement/rêve.

J'aimerais que vous me décriviez de la façon la plus complète possible votre deuxième
événement/rêve, en commençant au début.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.
les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose
comme "Vous avez dit X, qu'est-il arrivé après?" ou "Et après, qu'est-il arrivé?".)

Vous rappelez-vous d'autres choses par rapport à cet événement?

Est-ce que c'est tout?

J'aimerais maintenant que vous me décriviez à nouveau de la façon la plus complète
possible l'événement/rêve dont nous venons de parler, en essayant de n'oublier aucun détail,
même s'ils paraissent insignifiants.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.
les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose
comme "Vous avez dit X, qu'est-il arrivé après?" ou "Et après, qu'est-il arrivé?".)

Vous rappelez-vous d'autres choses par rapport à cet événement?

Est-ce que c'est tout?

OK.

Très bien. J'aimerais maintenant vous poser quelques questions par rapport aux deux
événements/rêves, en commençant par le premier que vous avez raconté. Êtes-vous prêts?

(les questions posées sont ouvertes et non-directives. Poser 1-4 questions par rapport aux éléments mentionnés s'ils sont peu ou pas décrits.)

ex de questions

-quel âge aviez-vous à ce moment?

-que portiez-vous?

-Vous avez parlé de X, pouvez-vous me le décrire un peu plus?

Avez-vous d'autre chose à rajouter à propos de cet événement/rêve?

OK.

J'aimerais maintenant vous poser quelques questions par rapport au deuxième événement/rêve? Êtes-vous prêts?

(les questions posées sont ouvertes et non-directives. Poser 1-4 questions par rapport aux événements mentionnés s'ils sont peu ou pas décrits.)

ex de questions

-quel âge aviez-vous à ce moment?

-que portiez-vous?

-Vous avez parlé de X, pouvez-vous me le décrire un peu plus?

Avez-vous d'autre chose à rajouter à propos de cet événement?

Très bien.

7.Debriefing

Le debriefing est fait à la fin du rendez-vous avec le participant. À la fin, le deuxième formulaire de consentement est lu au participant et toute question qu'il peut avoir concernant l'expérience est répondue. Finalement, dix dollars sont donnés au participant. Si le participant décide de ne pas signer le deuxième formulaire de consentement, l'information et les enregistrements reliés à sa participation sont immédiatement détruits.

Script de debriefing

Le but de la recherche à laquelle vous venez de participer est d'étudier la validité et l'efficacité des techniques d'analyse de contenu utilisées dans le domaine légal pour évaluer la crédibilité de témoignages. Plus particulièrement, la recherche évalue la capacité d'une technique appelée Statement Validity Analysis à différencier les récits réels des récits inventés.

(lire le deuxième formulaire de consentement avec le participant et réexpliquer les éléments importants si nécessaire)

Si vous voulez discuter de l'expérience avec l'équipe de recherche, soit maintenant ou plus tard, vos commentaires et vos questions sont les bienvenus. Si vous le désirez, vous pouvez contacter l'expérimentateur plus tard pour discuter des buts de l'expérience plus en profondeur.

Les informations que vous nous avez fournies et les enregistrements faits de votre participation dans cette étude sont confidentiels. Ils seront utilisés pour des fins de recherche exclusivement. L'accès à ce matériel est restreint aux chercheurs et assistants de recherche directement impliqués dans cette recherche. De plus, les enregistrements seront détruits après une période maximale d'un an.

La recherche va potentiellement bénéficier aux professionnels œuvrant dans les domaines clinique et légal. Entre autres, elle va apporter des indications sur l'efficacité et les contraintes reliées à l'utilisation des techniques d'analyse de contenu pour évaluer la crédibilité d'une déclaration.

Si vous désirez connaître les résultats de l'étude une fois complétée, avisez simplement l'expérimentateur principal. Si vous avez des questions que vous aimeriez poser maintenant ou plus tard, sentez-vous libres de contacter soit Dominic Beaulieu-Prévost, B.Sc., ou Dr. Jean-Roch Laurence, au 848-2213 ou 848-7555, ou écrivez à: Université Concordia, Laboratoire d'hypnose, H-531, 1455 boul. de Maisonneuve Ouest, Montréal, Québec, H3G 1M8. Nous vous remercions encore pour votre participation à cette recherche, votre aide fut d'une très grande importance.

PROTOCOLE ANGLAIS

Phone call

Could I talk to [nom] please?

Hi, my name is Dominic and I work at Concordia University in the hypnosis lab. I went into your _____ class to recruit participants for a forensic psychology study and you signed the list. I would like to know if you are still interested in participating and if you would like to have more information about the study.

(si accepte...)

The experiment will be done at Concordia University. The session should take about 40min. It would consist of an interview in which you will be asked to report either two events or two dreams. The stories would be either real or invented. You would also be asked to report the two stories so that they sound as credible as possible. You would be compensated with five dollars and an additional five dollars would be given to you if, during the interview, the interviewer believes your two stories real.

Would you still like to participate?

Before you come to the lab for the session, Dr Jean-Roch Laurence will send you a letter with instructions on how to choose your two stories. Could you give me the address to which the letter should be mailed?

The letter will be mailed either today or tomorrow, so you should receive it in 3-4 days. Could we schedule our appointment now?

Let me give you my phone number in case you have any questions before our appointment.

My name is Dominic and my phone number at home is 482-0353. Thank you very much and I will call you the day before to remind you of the appointment and to make sure that you received and read the letter, OK?

(Envoi de la lettre avec la signature de Jean-Roch Laurence, voir annexe E pour exemple)

Première rencontre

1. Présentation

2. Lecture du premier formulaire de consentement et explication de ce que cela implique.

-informer que je ne sais pas si les événements racontés sont réels ou inventés.

3. Signature du formulaire et donner les numéros de téléphone:

-du laboratoire d'hypnose (848-2213)

-du bureau de Jean-Roch (848-7555)

-et mon numéro personnel (482-0353).

4. Entrevue.

-S'assurer que le participant a en sa possession la lettre qui lui a été envoyée. Sinon, lui en donner une copie pour qu'il puisse la consulter durant ses 5min de préparation.

-Explication des instructions à suivre pour l'entrevue:

(Voir lettre envoyée aux participants) + Do you have any questions?

I will now give you 5min to prepare yourself before the interrogation officially begins. Do you have any questions before I leave?

-Partir 5min et donner copie de lettre envoyée-

Are you ready?

Now, I would like you to describe to me in as much details as possible your first event/dream, starting from the beginning.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie. les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose comme "You said X, what happened after that?" ou "And after that, what happened?" .)

Do you remember something else about this event/dream?

Anything else?

Now, I would like you to describe to me again, in as much details as possible, the event/dream we just talked about. Please try not to forget any details, even if they seem trivial.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.
les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose
comme "You said X, what happened after that?" ou "And after that, what happened?".)

Do you remember something else about this event/dream?

Anything else?

OK.

Very good. We will now go to the second event/dream.

Now, I would like you to describe to me in as much details as possible your first
event/dream, starting from the beginning.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.
les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose
comme "You said X, what happened after that?" ou "And after that, what happened?".)

Do you remember something else about this event/dream?

Anything else?

Now, I would like you to describe to me again, in as much details as possible, the
event/dream we just talked about. Please try not to forget any details, even if they seem
trivial.

(laisser le participant parler et ne pas poser d'autres questions avant la fin de cette partie.
les pauses sont tolérées, mais lorsqu'elles sont trop longues, l'interviewer dit quelque chose
comme "You said X, what happened after that?" ou "And after that, what happened?".)

Do you remember something else about this event/dream?

Anything else?

OK.

Very good. Now, I would like to ask you some questions about the two events/dreams,
starting with the first one you reported. Are you ready?

(les questions posées sont ouvertes et non-directives. Poser 1-4 questions par rapport aux éléments mentionnés s'ils sont peu ou pas décrits.)

ex de questions

-how old were you at that time?

-what were you wearing?

-you talked about X, could you describe it to me a bit more?

Is there anything else you would like to add about this event/dream?

OK.

Now I would like to ask you some questions about the second event/dream. Are you ready?

(les questions posées sont ouvertes et non-directives. Poser 1-4 questions par rapport aux événements mentionnés s'ils sont peu ou pas décrits.)

ex de questions

-how old were you at that time?

-what were you wearing?

-you talked about X, could you describe it to me a bit more?

Is there anything else you would like to add about this event/dream?

Very Good.

7. Debriefing

Le debriefing est fait à la fin du rendez-vous avec le participant. À la fin, le deuxième formulaire de consentement est lu au participant et toute question qu'il peut avoir concernant l'expérience est répondue. Finalement, dix dollars sont donnés au participant. Si le participant décide de ne pas signer le deuxième formulaire de consentement, l'information et les enregistrements reliés à sa participation sont immédiatement détruits.

Script de debriefing

The goal of the research to which you participated is to study the validity and the efficacy of content-analysis techniques used in the forensic and clinical contexts to assess the credibility of a statement. More specifically, the research studied the capacity of a procedure called Statement Validity Analysis to differentiate real stories and invented stories.

(lire le deuxième formulaire de consentement avec le participant et réexpliquer les éléments importants si nécessaire)

If you want to discuss about the experiment with the research team, either now or later, your comments and questions are welcomed. If you want, you can call back the experimenter to discuss more about the goals of the experiment.

The informations you provided and the recordings related to your participation in this study are confidential. They will exclusively be used for research purposes. Access to this material is limited to the researchers and research assistants directly implicated in this research. Furthermore, the recordings will be destroyed after a maximum period of one year.

The research will potentially benefit the professionals working in the fields of forensic and clinical psychology. For example, it will provide information on the efficacy and the limits of content-analysis techniques in the context of credibility assessment. If you would like to know about the results of this study, you just have to inform the main experimenter. Si you have questions that you would like to ask now or later, feel free to contact either Dominic Beaulieu-Prévost, B.Sc., or Dr. Jean-Roch Laurence, at 848-2213 or 848-7555, or write at: Concordia University, Hypnosis laboratory, H-531, 1455 boul. de Maisonneuve Ouest, Montréal, Québec, H3G 1M8. We thank you for your participation to this research, your help was important and appreciated.

ANNEXE E

Lettre envoyée aux participants (versions française et anglaise)

Dr Jean-Roch Laurence
Laboratoire d'hypnose
Concordia University

Bonjour cher participant(e),

comme convenu voici les instructions reliées à la recherche sur l'évaluation de la crédibilité de témoignages en contexte légal. Il vous est demandé de lire attentivement ces instructions avant votre rendez-vous au laboratoire et d'amener cette lettre avec vous au rendez-vous. Comme il fut expliqué au téléphone, vous avez à préparer deux récits d'événements ou de rêves sur lesquels vous serez interrogés durant l'entrevue. Les événements ou rêves choisis doivent être soit réels ou inventés, selon les indications mentionnées dans cette lettre. Un événement (ou rêve) réel est un événement (ou rêve) qui vous est personnellement arrivé tandis qu'un événement (ou rêve) inventé est un événement(ou rêve) qui ne vous est jamais arrivé et que vous devrez inventer pour l'entrevue. N'oubliez pas que la personne qui va vous interviewer ne sait pas quel type de récit vous devez raconter (réel ou inventé) ni l'ordre dans lequel vous devez le faire. Il est important de ne jamais lui dire et de lui présenter vos deux événements (ou rêves) comme étant réels, peu importe la vérité. Vous devrez aussi essayer d'être le plus crédible possible lorsque vous serez interrogés sur chacun des événements(ou rêves). Pour vous aider, une compensation de 5\$ sera ajoutée au dédommagement de base de 5\$ à moins que, durant un des deux interrogatoires, l'interviewer soit totalement convaincu que l'événement (ou le rêve) raconté ne soit jamais arrivé. Si l'interviewer arrive à cette conclusion, l'interrogatoire à propos de cet événement (ou rêve) sera interrompu et seul le dédommagement de 5\$ vous sera donné.

Veillez noter que malgré que l'interviewer ne sait pas si les récits sont réels ou inventés, il sait si vous avez deux rêves ou deux événements à raconter. Il n'y a donc aucun problème à lui dire le genre de récit (rêve ou événement) que vous raconterez.

Pour l'entrevue, vous aurez à raconter:

- Deux événements
- Deux rêves

Le premier sera... réel
 inventé

Le deuxième sera... réel
 inventé

Pour vous aider à choisir et/ou inventer vos deux événements(ou rêves) de façon adéquate, les règles suivantes ont été établies:

POUR LES ÉVÉNEMENTS

- 1) L'événement doit être arrivé après l'âge de 13ans.
- 2) Ce ne doit pas être un événement pour lequel vous avez été ou êtes présentement en psychothérapie.
- 3) Ce ne doit pas être un événement qui s'étend sur une longue période (p. e. "j'ai travaillé 1 an comme barman") car un tel événement est impossible à raconter en entier durant une entrevue.
- 4) Choisir (ou inventer) un événement anecdotique, c.-à-d. qui est différent de ce qui vous arrive ou ce que vous faites quotidiennement.

- 5) Choisir (ou inventer) un événement dont vous vous souvenez assez bien et qui a une certaine valeur émotionnelle, c.-à-d. par rapport auquel vous n'êtes pas totalement indifférent.
- 6) Choisir (ou inventer) un événement marquant, c.-à-d. qui a eu (ou aurait eu) un certain impact sur votre vie au quotidien ou qui est surprenant ou inusité.
- 7) Pour les événements inventés, essayez d'inventer un événement assez crédible pour passer pour un événement réel, l'objectif étant de tromper l'interviewer.

POUR LES RÊVES

- 1) Le rêve doit être arrivé après l'âge de 13ans.
- 2) Ce ne doit pas être un rêve pour lequel vous avez été ou êtes présentement en psychothérapie.
- 3) Ce ne doit pas être une série de rêves qui s'étend sur une longue période (p. e. "j'ai rêvé pendant un mois à...") car une telle série de rêves est impossible à raconter en entier durant une entrevue.
- 4) Si possible, choisir (ou inventer) un rêve comportant peu ou pas d'éléments irréels (voler dans les airs, créatures imaginaires,..).
- 5) Choisir (ou inventer) un rêve dont vous vous souvenez assez bien et qui a une certaine valeur émotionnelle, c.-à-d. par rapport auquel vous n'êtes pas totalement indifférent.
- 6) Choisir (ou inventer) un rêve marquant, c.-à-d. qui a eu (ou aurait eu) un certain impact sur votre vie au quotidien ou qui est surprenant ou inusité.
- 7) Pour les rêves inventés, essayez d'inventer un rêve assez crédible pour passer pour un vrai rêve, l'objectif étant de tromper l'interviewer.

Comme vous serez interrogés sur ces deux événements(ou rêves), il vous est demandé de prendre quelques minutes pour choisir ou inventer chacun d'eux. Vous pouvez utiliser d'autres feuilles pour travailler vos deux récits mais il est requis d'au moins décrire chaque événement(ou rêve) en quelques mots et d'en écrire les principaux éléments sur la feuille annexée à cette lettre. Vous devrez ensuite amener cette feuille à l'entrevue.

Durant la rencontre, vous aurez 4-5 minutes pour vous préparer avant chacun des deux interrogatoires. Ce temps vous sera donné pour vous rafraîchir la mémoire ou vous rappeler les détails de votre invention. Par contre, il est attendu que vous avez déjà choisi et écrit les deux événements (ou rêves) avant d'arriver au laboratoire(voir feuille annexée).

Au nom des chercheurs reliés à cette recherche, je vous remercie du temps que vous y consacrez. Sans vous, cette recherche n'aurait pas pu exister.

Sincèrement vôtre,

Jean-Roch Laurence,
 Directeur du laboratoire d'hypnose,
 Concordia University

RECHERCHE SUR L'ÉVALUATION DE LA CRÉDIBILITÉ

Participant(e) # _____

- Deux rêves
 Deux événements

Premier récit raconté: Réel Inventé

Description générale: _____

Éléments principaux et notes: _____

Deuxième récit raconté: Réel Inventé

Description générale: _____

Éléments principaux et notes: _____

Dr Jean-Roch Laurence
Hypnosis laboratory
Concordia University

Dear participant,

below are the instructions related to the research on the credibility assessment of self-reports in a forensic context. You are asked to read these instructions attentively before coming for your appointment at the laboratory and to bring this letter with you to the appointment. As you were explained on the phone, you have to prepare two stories of events or dreams on which you will be interrogated during the interview. The events or dreams chosen must be either real or invented, depending on the indications mentioned in this letter. A real event (or dream) is an event (or dream) that happened to you personally while an invented event (or dream) is an event (or dream) that never happened to you and that you will have to create for the interview. Do not forget that the person who will do the interview doesn't know whether your story is real or invented nor the order in which you will report them. It is important to never tell him and to report your two events (or dreams) to him as if they are real, whether they are or not. You will also have to sound as credible as possible when you will be interrogated on each event (or dream). In fact, an extra 5\$ compensation will be added to the basic 5\$ compensation if you sound convincing. This extra 5\$ will not be given to you if, during one of the two interrogations, the interviewer becomes totally convinced that the event (or dream) you reported never happened. If the interviewer arrives at that conclusion, the interrogation about this event (or dream) will be interrupted and only the basic 5\$ compensation will be given to you.

Please take note that while the interviewer will not know whether your stories are real or invented, he will know whether you are reporting two dreams or two events. As such, there is no problem if you tell him whether you are reporting a dream or an event.

For the interview, you will to report:

- Two events
- Two dreams

The first will be... real
 invented

The second will be... real
 invented

To help you to choose and/or invent your two events (or dreams) correctly, the following rules have been established:

FOR THE EVENTS

- 1) The event must have happened after the age of 13.
- 2) It should not be an event for which you were or are seeing a psychotherapist.
- 3) It should not be an event that is spread across a long period of time (e.g. "I worked 1 year as a barman") because such an event is impossible to report entirely during an interview.
- 4) Choose (or invent) an unusual event, i.e. one that is different from what normally happens to you or what you do in your everyday life.
- 5) Choose (or invent) an event that you remember well and which has a certain emotional value, i.e. one where you felt an emotion such as anger or surprise.

- 6) Choose (or invent) a significant event, i.e. an event which had (or would have had) a certain impact on your everyday life or which was surprising or unusual.
- 7) For the invented events, try to create an event which is credible enough to sound like a real event, the objective being to fool the interviewer.

FOR THE DREAMS

- 1) The dream must have happened after the age of 13.
- 2) It should not be a dream for which you were or are seeing a psychotherapist.
- 3) It should not be a series of dreams spread across a long period of time (e.g. "I dreamed for a month about...") because such a series of dreams is impossible to report entirely during an interview.
- 4) If possible, choose (or invent) a dream with few or no imaginary elements (flying, mythical creatures,...).
- 5) Choose (or invent) a dream that you remember well and which has a certain emotional value, i.e. one where you felt an emotion such as anger or surprise.
- 6) Choose (or invent) a significant dream, i.e. a dream which had (or would have had) a certain impact on your everyday life or which was surprising or unusual.
- 7) For the invented dreams, try to create a dream which is credible enough to sound like a real dream, the objective being to fool the interviewer.

Since you will be interrogated about these two events (or dreams), it is important that you take a few minutes to choose or invent each of them. Once you have done so, please write it down on the form included in this letter. Describe each event (or dream) in a few words and write the main elements on the form. You can also use additional sheets of paper to work out your two stories. You must bring this form to the appointment but you will not be able to refer to it during your interrogation.

During the appointment, you will have 4-5 minutes to prepare yourself before each of the two interrogations. The purpose of this period of time is to help you refresh your memory concerning the details of your story.

On behalf of the researchers implicated in this research, I thank you for your time and help. Without you, this research could not take place.

Sincerely yours,

Jean-Roch Laurence,
 Director of the hypnosis laboratory,
 Concordia University

RESEARCH ON CREDIBILITY ASSESSMENT

Participant # _____

- Two dreams
- Two events

First story reported: Real Invented

General description: _____

Main elements and details: _____

Second story reported Real Invented

General description: _____

Main elements and details: _____

ANNEXE F

Statistiques descriptives

Tableau F-1

Moyennes (M) et écarts-types (ÉT) des événements et des rêves en fonction de chaque critère.

CRITÈRE	ÉVÉNEMENTS		RÊVES	
	<u>M</u>	<u>ÉT</u>	<u>M</u>	<u>ÉT</u>
<u>Critères principaux</u>				
1. Structure logique	1.78	0.48	1.81	0.47
2. Production non-structurée	0.47	0.61	0.16	0.37
3. Quantité de détails	84.22	43.98	56.19	21.46
4. Enchâssement contextuel	1.28	0.78	0.44	0.62
5. Interactions	0.94	1.22	0.50	0.72
6. Rappel de conversation	6.19	10.66	2.31	2.53
7. Complic. inattendues	0.19	0.40	0.00	0.00
8. Détails inusités	0.08	0.28	0.31	0.47
9. Détails superflus	0.25	0.44	0.09	0.30
11. Incidents extérieurs	0.06	0.23	0.06	0.25
12. États subjectifs	4.58	3.24	3.25	2.60
13. Attributions	1.36	1.57	0.66	0.79
14. Corr. spontanées (total)	2.14	1.97	1.13	1.13
15. Absence de souvenir	0.33	0.53	1.47	2.68
16. Doutes (total)	2.31	2.90	1.19	1.80
<u>Critères additionnels</u>				
Réminiscence	13.22	12.24	7.78	8.66
Interactions (2+éléments)	1.81	1.65	1.03	0.90
Citations indirectes	1.39	1.29	0.75	0.92
Corrections spontanées	1.42	1.78	0.53	0.84
Additions spontanées	0.72	0.78	0.59	0.98
Doutes (explicite)	0.47	1.13	0.38	0.87
Image de soi	1.25	0.69	1.03	0.59

Tableau F-2

Moyennes (M) et écarts-types (ÉT) des déclarations réelles et des déclarations inventées en fonction de chaque critère.

CRITÈRE	D. RÉELLES		D. INVENTÉES	
	<u>M</u>	<u>ÉT</u>	<u>M</u>	<u>ÉT</u>
<u>Critères principaux</u>				
1. Structure logique	1.62	0.60	1.97	0.17
2. Production non-structurée	0.35	0.49	0.29	0.58
3. Quantité de détails	68.47	33.21	73.59	42.08
4. Enchâssement contextuel	0.91	0.83	0.85	0.82
5. Interactions	0.79	1.04	0.68	1.04
6. Rappel de conversation	4.38	6.08	4.35	9.86
7. Complic. inattendues	0.18	0.39	0.03	0.17
8. Détails inusités	0.09	0.29	0.29	0.46
9. Détails superflus	0.18	0.39	0.18	0.39
11. Incidents extérieurs	0.09	0.29	0.03	0.17
12. États subjectifs	3.91	2.96	4.00	3.10
13. Attributions	0.91	1.26	1.15	1.35
14. Corr. spontanées (total)	1.85	1.94	1.47	1.42
15. Absence de souvenir	1.06	2.58	0.68	0.98
16. Doutes (total)	1.85	2.79	1.71	2.20
<u>Critères additionnels</u>				
Réminiscence	12.97	14.03	8.35	6.05
Interactions (2+éléments)	1.38	1.44	1.50	1.38
Citations indirectes	1.26	1.38	0.91	0.90
Corrections spontanées	1.18	1.78	0.82	1.09
Additions spontanées	0.68	0.94	0.65	0.81
Doutes (explicite)	0.38	0.70	0.47	1.26
Image de soi	1.18	0.67	1.12	0.64

Tableau F-3

Moyennes (M) et écarts-types (ÉT) des événements réels (ÉR), événements inventés (ÉI), rêves réels (RR) et rêves inventés (RI) chez les francophones en fonction de chaque critère.

CRITÈRE	ÉR		ÉI		RR		RI	
	M	ÉT	M	ÉT	M	ÉT	M	ÉT
<u>Critères principaux</u>								
1. Structure logique	1.71	0.49	2.00	0.00	1.83	0.41	2.00	0.00
2. Production non-structurée	0.29	0.49	0.43	0.79	0.50	0.55	0.17	0.41
3. Quantité de détails	75.29	45.13	69.71	43.13	60.17	31.01	62.33	26.39
4. Enchâssement contextuel	1.14	0.90	0.71	0.76	0.33	0.52	0.33	0.52
5. Interactions	0.86	0.69	0.57	0.79	0.67	0.82	0.67	0.52
6. Rappel de conversation	4.57	6.63	2.43	2.82	3.33	4.32	3.17	3.19
7. Complic. inattendues	0.14	0.38	0.14	0.38	0.00	0.00	0.00	0.00
8. Détails inusités	0.00	0.00	0.00	0.00	0.17	0.41	0.50	0.55
9. Détails superflus	0.29	0.49	0.14	0.38	0.00	0.00	0.00	0.00
11. Incidents extérieurs	0.00	0.00	0.14	0.38	0.33	0.52	0.00	0.00
12. États subjectifs	5.14	3.93	3.43	1.27	3.50	3.51	4.83	3.31
13. Attributions	1.29	1.11	0.86	0.90	0.50	0.84	0.83	0.75
14. Corr. spontanées (total)	1.43	1.62	1.86	1.68	2.17	1.94	1.00	0.63
15. Absence de souvenir	0.43	0.53	0.14	0.38	1.17	1.17	1.50	1.64
16. Doutes (total)	2.71	5.44	2.29	3.25	1.83	1.47	1.50	2.26
<u>Critères additionnels</u>								
Réminiscence	16.57	22.57	6.57	6.55	15.00	15.01	9.00	7.69
Interactions (2+éléments)	1.00	0.58	1.29	1.25	1.00	0.63	0.83	0.41
Citations indirectes	2.29	1.38	0.86	0.69	0.50	0.55	0.83	0.75
Corrections spontanées	0.86	1.46	1.43	1.27	1.17	1.60	0.33	0.52
Additions spontanées	0.57	0.79	0.43	0.53	1.00	2.00	0.67	0.82
Doutes (explicite)	0.71	0.95	1.29	2.21	0.83	0.98	0.83	1.60
Image de soi	1.57	0.79	0.86	0.38	1.33	0.52	0.67	0.52

Tableau F-4

Moyennes (M) et écarts-types (ÉT) chez les anglophones des événements réels (ÉR), événements inventés (ÉI), rêves réels (RR) et rêves inventés (RI) en fonction de chaque critère.

CRITÈRE	ÉR		ÉI		RR		RI	
	<u>M</u>	<u>ÉT</u>	<u>M</u>	<u>ÉT</u>	<u>M</u>	<u>ÉT</u>	<u>M</u>	<u>ÉT</u>
<u>Critères principaux</u>								
1. Structure logique	1.55	0.69	1.91	0.30	1.50	0.71	2.00	0.00
2. Production non-structurée	0.55	0.52	0.55	0.69	0.10	0.32	0.00	0.00
3. Quantité de détails	86.55	30.86	96.82	55.64	48.80	13.54	57.50	19.56
4. Enchâssement contextuel	1.45	0.69	1.55	0.69	0.50	0.71	0.50	0.71
5. Interactions	1.18	1.33	1.00	1.61	0.40	0.97	0.40	0.52
6. Rappel de conversation	7.00	8.37	8.88	16.62	2.00	1.63	1.50	1.18
7. Complic. inattendues	0.45	0.52	0.00	0.00	0.00	0.00	0.00	0.00
8. Détails inusités	0.09	0.30	0.18	0.40	0.10	0.32	0.50	0.53
9. Détails superflus	0.27	0.47	0.27	0.47	0.10	0.32	0.20	0.42
11. Incidents extérieurs	0.09	0.30	0.00	0.00	0.00	0.00	0.00	0.00
12. États subjectifs	4.18	2.48	5.36	4.27	3.00	2.40	2.40	1.43
13. Attributions	1.45	1.75	1.64	2.01	0.30	0.48	1.00	0.94
14. Corr. spontanées (total)	2.73	2.57	2.18	1.72	1.00	0.82	0.70	0.67
15. Absence de souvenir	0.36	0.67	0.36	0.50	2.20	4.54	0.90	0.88
16. Doutes (total)	2.45	2.02	1.91	0.94	0.60	0.52	1.20	2.49
<u>Critères additionnels</u>								
Réminiscence	16.73	10.66	11.82	4.96	5.10	5.65	5.40	4.27
Interactions (2+éléments)	2.18	1.94	2.27	1.90	1.00	1.33	1.20	0.79
Citations indirectes	1.45	1.69	1.09	0.83	0.80	0.92	0.80	1.23
Corrections spontanées	2.09	2.47	1.09	1.38	0.40	0.52	0.40	0.52
Additions spontanées	0.64	0.50	1.09	1.04	0.60	0.52	0.30	0.48
Doutes (explicite)	0.09	0.30	0.18	0.40	0.20	0.42	0.00	0.00
Image de soi	0.91	0.70	1.64	0.50	1.10	0.57	1.00	0.67