

L'utilisation de mesures indirectes et directes du comportement dans l'évaluation des
interventions ciblant les enfants agressifs

François Poulin

Annick Moisan

Université du Québec à Montréal

Stéphane Cantin

Université de Montréal

Personne contact : François Poulin; Département de Psychologie, UQAM, Case postale 8888,
Succursale Centre-ville, H3C 3P8, Montréal, Canada. Courriel: poulin.francois@uqam.ca.
Tél: (514) 987-3000 (poste 7766).

L'utilisation de mesures indirectes et directes du comportement dans l'évaluation des interventions ciblant les enfants agressifs

L'efficacité des interventions auxquelles sont exposés les enfants agressifs doit être évaluée avec rigueur. Pour ce faire, les chercheurs et praticiens doivent disposer de mesures fiables, valides, exemptes de biais et qui soient suffisamment sensibles pour permettre de détecter tout changement dans le comportement des enfants. Le premier objectif de cet article consiste à mettre en évidence les limites associées à l'utilisation de mesures indirectes pour évaluer le comportement des enfants exposés à une intervention. Le second objectif vise à illustrer les enjeux associés à l'utilisation de mesures directes basées sur l'observation du comportement. Une analyse des différentes possibilités qui s'offrent aux chercheurs et aux praticiens, notamment en ce qui a trait au choix des comportements ciblés par l'observation et au contexte où se déroule l'observation, est proposée.

Mots-clés : mesure et évaluation, observation, intervention, méthodologie

Behavioral outcomes of intervention targeting young aggressive children have to be carefully assessed. Researchers and practioners need reliable, valid, and bias-free instruments that are also sensitive enough to detect any change in children's behavior. The first goal of this article is to highlight limitations imposed by indirect measures of behavioral assessment with children exposed to an intervention. The second goal is to illustrate issues related to direct behavioral assessment using observation procedure. The different options for researchers and practioners regarding the choice of behavioral targets and observation settings are discussed.

Keywords: assessment, observation, intervention, methodology

L'utilisation de mesures indirectes et directes du comportement dans l'évaluation des interventions ciblant les enfants agressifs

Un nombre important de jeunes enfants manifeste fréquemment des comportements agressifs dès la maternelle (Breton et al., 1999; ELNEJ, 1999). Certains d'entre eux vont continuer à manifester ces comportements jusqu'à l'adolescence et vont rencontrer une panoplie de problèmes psychosociaux (Tremblay, 2011). Il est donc avisé d'intervenir de façon préventive auprès d'eux, et ce, le plus tôt possible en début de scolarisation. Ces enfants retiennent d'ailleurs grandement l'attention des intervenants en santé mentale. En effet, plus de 50% des demandes d'aide psychologique reçues de la part des professionnels de la santé oeuvrant auprès de la clientèle infantile concernent une problématique reliée à l'agressivité (Loeber, Burke, Lahey, Winters, & Zera, 2000). Les comportements d'agressivité mobilisent d'ailleurs une part substantielle des ressources consacrées à la jeunesse dans les réseaux de la santé mentale (hôpitaux et CSSS), dans le milieu scolaire (services spécialisés) et dans les organismes sociaux (centres jeunesse, centres de réadaptation), en plus des démarches personnelles de consultations effectuées par les familles en clinique privée. Dans ce contexte, la poursuite de travaux de recherche portant sur la conception et l'expérimentation de nouveaux outils d'intervention doit donc être priorisée.

La conception des interventions doit être basée sur des modèles théoriques validés permettant de rendre compte de l'émergence et du maintien des comportements agressifs (Vitaro & Tremblay, 2008). Dans cette perspective, les interventions ont pour objectif de cibler les facteurs de risque associés à l'agressivité de façon à réduire leur action nocive et les facteurs de protection de façon à maximiser leur effet bénéfique (Kellam & Van Horn, 1997). Différentes dimensions des relations que l'enfant entretient avec ses parents, ses pairs et son enseignant, de même que ses caractéristiques socio-cognitives sont des exemples de facteurs

souvent ciblés par les interventions. En intervenant sur ces facteurs, il est souhaité que l'enfant en vienne à manifester de moins en moins fréquemment certains comportements agressifs et de plus en plus fréquemment certains comportements positifs, comme les habiletés sociales. Pour ce faire, les interventions visent à modifier le comportement de l'enfant par l'application de contingences (renforcement, punition) et/ou par l'apprentissage de nouveaux comportements (modelage, enseignement ou autre). Une intervention réussie sur le comportement de l'enfant (variable dite proximale) pourrait en retour entraîner des gains significatifs sur la qualité de l'intégration sociale de l'enfant ou sa réussite scolaire (variables dites distales).

Par ailleurs, une évaluation rigoureuse et continue de l'efficacité des services d'intervention offerts aux enfants agressifs d'âge scolaire (de la maternelle jusqu'à la fin du primaire) est essentielle. Pour être concluante, une telle évaluation doit respecter certains critères méthodologiques (pour une discussion approfondie des critères d'évaluation, voir Brown et al., 2008; Chambless & Hollon, 1998; Vitaro, 2001). Ces critères incluent notamment l'utilisation d'instruments de mesure fiables et valides qui soient suffisamment sensibles pour bien détecter les changements (même les plus minimes) dans le comportement de l'enfant suite à l'exposition à l'intervention, et ce, dans différents contextes de socialisation. Le présent article s'inscrit dans cette perspective. La première partie de l'article présente les limites associées à l'utilisation de mesures indirectes pour évaluer le comportement des enfants exposés à une intervention. La deuxième partie propose une discussion des enjeux associés à l'utilisation de mesures directes basées sur l'observation du comportement à des fins d'évaluation d'une intervention. Une analyse des différentes possibilités qui s'offrent aux chercheurs et aux praticiens, notamment en ce qui a trait au

choix des comportements ciblés par l'observation et au contexte où se déroule l'observation (p. ex. milieu naturel versus laboratoire; contexte structuré versus non structuré) est proposée.

Les mesures indirectes

Une stratégie couramment employée pour évaluer la fréquence d'émission de certains comportements par l'enfant consiste à consulter les individus qui interagissent avec lui sur une base régulière. Il peut s'agir d'adultes (parents, enseignants, éducateurs) ou encore de pairs (camarades de classe). L'observation est dite indirecte parce qu'elle se déroule par le truchement d'une tierce personne. Dans le cas des adultes, ils sont invités à compléter un questionnaire en ayant en tête le comportement de l'enfant ciblé par l'évaluation. À titre illustratif, le *Child Behavior Checklist* (Achenbach, 1991), le *Behavioral and Emotional Rating Scale* (Epstein & Sharma, 1998), le *School Social Behavior Scales* (Merrell, 1993), le *Social Skills Rating System* (Gresham & Elliot, 1990) et le *Walker-McConnell Scales of Social Competence and School Adjustment* (Walker & McConnell, 1995) sont des exemples d'instruments souvent utilisés dans les écrits. Typiquement, ces instruments sont composés de plusieurs énoncés qui décrivent différentes formes de comportements extériorisés, intériorisés ou prosociaux. Les évaluateurs sont alors invités à indiquer à quelle fréquence l'enfant ciblé manifeste chacun des comportements. Lorsqu'il est utilisé dans le cadre de l'évaluation d'une intervention, ce type de mesure présente certaines limites qui concernent à la fois la source d'information, le contexte d'évaluation, de même que la méthode utilisée.

Les parents sont une source d'information fort intéressante parce qu'ils côtoient beaucoup leur enfant et qu'ils connaissent son évolution. Cependant, les parents savent que leur enfant prend part à une intervention et, dans bien des cas, ils sont eux-mêmes activement impliqués dans l'intervention. Ils ont tendance à surestimer les effets des interventions et à rapporter des changements positifs dans le comportement de l'enfant, que ces changements

soient réels ou non (Stoolmiller, Eddy, & Reid, 2000). Les enseignants (ou éducateurs) sont également de bons juges du comportement de l'enfant. Ils côtoient plusieurs enfants de même âge et sont donc plus sensibles aux différences individuelles. Ils connaissent le développement des enfants et sont en mesure d'observer l'enfant dans différents contextes propices à l'émission de comportements agressifs tels que les grands groupes de pairs (Saudino, Ronald, & Plomin, 2005). Cependant, tout comme pour les parents, leur évaluation devient moins fiable lorsqu'ils savent que l'enfant est exposé à une intervention. Par ailleurs, les évaluations produites par ces répondants sont basées sur une appréciation des comportements de l'enfant dans un contexte précis: à la maison dans le cas des parents et à l'école dans le cas des enseignants. L'évaluation d'intervention à partir d'une seule de ces deux sources ne permet donc pas de rendre compte du transfert et de la généralisation des acquis d'un contexte à l'autre.

Les perceptions des parents et des enseignants sont généralement recueillies par questionnaire, ce qui constitue également une limite importante. Les parents et les enseignants sont invités à indiquer à quelle fréquence l'enfant ciblé manifeste différents comportements (p. ex. « frappe les autres enfants ») au cours d'une période donnée (p. ex. dernier mois) à l'aide d'une échelle de réponse en trois, quatre ou cinq points d'ancrage. À titre d'exemple, le très populaire *Child Behavior Checklist* (CBCL; Achenbach, 1991) propose aux évaluateurs un choix de réponse en trois points (jamais, quelques fois, souvent). Ce type de mesure présente une limite importante quant à sa capacité à détecter des changements plus minimes que ceux correspondant aux choix de réponse. Suite à sa participation à une intervention, un enfant pourrait émettre un comportement moins fréquemment, mais ce changement pourrait ne pas être suffisamment substantiel pour que l'évaluateur passe d'un point d'ancrage à un autre (p. ex. de «souvent» au prétest à «quelques fois» au post-test). De plus, la gamme de

comportements évalués par ces questionnaires est souvent très restreinte ou trop générale, faisant ainsi en sorte que les comportements que l'intervention vise à modifier ne sont pas toujours mesurés adéquatement.

Le groupe de pairs constitue également une source d'évaluation pertinente lorsqu'il s'agit de rendre compte de l'impact d'une intervention. La perception des pairs est habituellement évaluée à partir d'une procédure de nominations (p. ex. Masten, Morison, & Pellegrini, 1985). Typiquement, les élèves d'un groupe classe doivent désigner les camarades de classe qui correspondent le mieux à différentes descriptions comportementales ou énoncés (p. ex. « quels sont les élèves qui frappent et bousculent souvent les autres ? »). Le nombre de nominations reçu à chacun des énoncés pour chaque élève est ensuite compilé. Un nombre élevé de nominations signifie que le comportement est perçu par les pairs comme étant fortement caractéristique de l'enfant. Bien qu'elles soient très valables lorsqu'il s'agit de rendre compte de la perception des pairs, cette procédure présente des limites importantes lorsqu'il s'agit d'évaluer l'impact d'une intervention. Ces limites concernent la source d'information et la méthode utilisée. Il est maintenant bien établi que l'évaluation que font les enfants du comportement de leurs pairs est en partie entachée par des biais de réputation qui circulent dans le groupe et qui dépassent souvent le comportement actuel de l'enfant (Hymel, 1986). Des chercheurs ont d'ailleurs déjà constaté que même lorsqu'une intervention parvenait avec succès à diminuer les comportements perturbateurs des enfants ciblés (tels qu'observés par des observateurs indépendants), la perception négative des pairs à l'endroit des enfants ciblés avait tendance à demeurer inchangée (Bierman & Furman, 1984; Cillessen, Bukowski, & Haselager, 2000). Ce biais de réputation au sein du groupe de pairs viendrait en retour compromettre le maintien des acquis chez les enfants ciblés par l'intervention.

Sur le plan de la méthode, un score élevé sur une mesure de nominations par les pairs ne signifie pas qu'un enfant émet nécessairement le comportement à une fréquence élevée de façon absolue, mais plutôt qu'il émet ce comportement plus fréquemment que ses camarades de classe. Un enfant qui manifeste moins de conduites agressives à la suite d'une intervention est tout de même susceptible de demeurer celui qui s'avère le plus agressif au sein de sa classe (malgré la diminution des conduites agressives, il occupe la même position relative au sein du groupe de pairs). Dans ce scénario, bien que son comportement ait changé, l'enfant est susceptible d'obtenir un nombre élevé de nominations qui sera fortement similaire au prétest et au post-test. Hansen et ses collègues (1996) soulignent d'ailleurs que la forte stabilité temporelle des évaluations issues des nominations par les pairs rend cette procédure peu propice à l'examen du changement pouvant résulter d'une intervention. Une stratégie alternative proposée par Nangle et ses collègues (1993) consiste à aller au-delà du nombre de nominations reçues et d'examiner en détail l'origine des nominations. Ainsi, il est possible qu'un enfant reçoive le même nombre de nominations avant et après l'intervention, mais que ces nominations n'aient pas été émises par les mêmes pairs. Une autre option intéressante, mais peu utilisée, consiste à remplacer les *nominations* par les pairs par une procédure d'*évaluation* par les pairs dans laquelle les participants sont appelés à se prononcer sur le comportement de chacun des élèves de leur classe (incluant l'enfant ciblé) à l'aide d'une échelle de fréquence (« jamais », « quelques fois », « souvent ») (Card, Hodges, Little, & Hawley, 2005).

En résumé, les mesures du comportement de l'enfant recueillies auprès des adultes qui le côtoient ou de ses pairs sont fréquemment utilisées dans l'évaluation d'intervention, notamment parce qu'elles sont peu coûteuses, relativement faciles à obtenir, à compiler et à analyser. Elles présentent cependant certaines faiblesses qui peuvent limiter la portée des

conclusions des études d'évaluation qui les utilisent. Dans ce contexte, les mesures directes basées sur l'observation du comportement doivent être incluses dans le devis d'évaluation d'une intervention autant que possible.

Les mesures directes basées sur l'observation du comportement

Le recours à l'observation directe présente des avantages certains lorsqu'on souhaite vérifier si une intervention a entraîné une modification du comportement chez l'enfant. Le chercheur ou le praticien qui considère l'utilisation d'une mesure observationnelle est cependant confronté à une série de possibilités et de questionnements. Ces questionnements concernent le choix des cibles d'observation (c.-à-d. les comportements observés) et le choix du contexte d'observation (milieu naturel versus laboratoire).

Avantages des mesures d'observation directe pour évaluer le changement suite à une intervention

Plusieurs avantages des mesures d'observation directe militent en faveur de leur utilisation dans les devis d'évaluation des interventions qui visent la modification de certains comportements. Premièrement, cette méthode permet d'obtenir des informations plus détaillées comparativement aux méthodes indirectes qui reflètent surtout les impressions globales des répondants (Margolin et al., 1998). L'observation directe permet d'évaluer de façon précise la fréquence ou la durée d'un comportement dans différents contextes de socialisation bien ciblés (Dishion & Granic, 2003). Il est donc possible de gagner en précision dans l'évaluation du phénomène. L'observation directe fournit ainsi une évaluation moins biaisée et donc plus représentative de la réalité que les méthodes indirectes en évitant d'avoir recours aux perceptions des participants (Coutu, Provost, & Bowen, 2005).

Deuxièmement, l'observation directe est plus sensible au changement récent dans le comportement de l'enfant, ce qui constitue un avantage majeur lorsque l'on souhaite évaluer

l'impact d'une intervention (Eddy, Dishion, & Stoolmiller, 1998). Le jugement global des adultes qui côtoient l'enfant (parents et enseignants) semble peu affecté par les changements récents dans son comportement (Eddy et al., 1998). Une méta-analyse menée par Stage et Quiroz (1997) fait ressortir que les mesures observationnelles détectent davantage le changement suite à une intervention que le jugement des enseignants. Pourtant, les enseignants sont encore plus proches de l'opinion des observateurs externes que ne le seraient les parents qui tendent plutôt à présenter une image plus positive de leur enfant (Swindells & Stagnitti, 2006). Par ailleurs, tel que mentionné plus haut, l'observation directe permettrait d'éviter que les gens rapportent une amélioration lorsqu'ils sont impliqués dans l'intervention comme peuvent l'être les parents ou les enseignants.

Choix des comportements ciblés par l'observation

Un premier questionnaire auquel est confrontée une personne qui souhaite avoir recours à l'observation directe pour évaluer l'impact d'une intervention concerne le choix des comportements sociaux qui seront ciblés par l'observation. Ces comportements doivent être nécessairement en lien avec les objectifs poursuivis par l'intervention. Cette affirmation semble évidente à première vue, mais elle n'est pas toujours respectée. Les comportements choisis doivent également avoir une importance dans l'environnement social de l'enfant de sorte qu'une modification de ces comportements puisse amener des changements au plan social (Foster, Inderbitzen, & Nangle, 1993). Les interventions mises en place auprès des enfants agressifs visent typiquement la diminution de leurs comportements agressifs (Kazdin, 2005). Dans cette perspective, la manifestation de comportements agressifs directs ou indirects, qu'ils soient de type réactif ou proactif, constitue des cibles d'observation pertinentes. Les études portant sur l'évaluation d'impact de ces interventions incluent de plus en plus des mesures observationnelles qui ciblent ce type de comportement (p. ex. CPPRG,

1999; Leflot et al., 2010; Stoolmiller, Eddy, & Reid, 2000; Webster-Stratton, Reid, & Stoolmiller, 2008). Par ailleurs, plusieurs des interventions mises en place auprès des enfants agressifs visent également l'augmentation de comportements jugés souhaitables, notamment les habiletés sociales (Boisjoli & Vitaro, 2005). Conséquemment, ce type de comportements (p. ex. partager, coopérer, aider, donner ses idées, prendre contact, etc.) constitue également une cible d'observation fort pertinente. Cependant, ces comportements sont plus rarement l'objet d'observation directe dans les évaluations d'intervention, sauf dans le cas de protocole à cas unique (Gresham, Sugai, & Horner, 2001). Foster, Inderbitzen et Nangle (1993) proposent une discussion approfondie des enjeux liés à l'observation des habiletés sociales en contexte d'intervention. Enfin, certaines interventions intègrent également des habiletés liées au contrôle des émotions et à la résolution de problèmes (Lochman, Boxmeyer, Powell, Barry & Pardini, 2010). Bien qu'il soit plus difficile d'observer directement chacun de ces aspects, leurs manifestations indirectes constituent certainement des cibles d'observation à privilégier (p. ex. impulsivité, interactions conflictuelles ou affects négatifs lors des situations de frustration).

Contrairement aux méthodes d'observation indirectes, l'observation directe permet de considérer le comportement dans son contexte, ce qui ouvre plusieurs possibilités, notamment l'identification de séquences de comportements (c.-à-d. ce qui précède et ce qui suit) et l'examen des processus d'interactions sociales entre les enfants (Gardner, 2000). L'affiliation avec des pairs déviants et les processus d'entraînement à la déviance (p. ex. renforcement des comportements agressifs) constituent des cibles d'observation d'autant plus pertinentes que les changements observés sur ces variables sont susceptibles de rendre compte de l'efficacité des interventions (Snyder et al., 2008). L'observation contextualisée permet aussi de déterminer quels comportements surviennent dans quels contextes spécifiques (Coutu et al.,

2005). Dans cet ordre d'idée, il devient aussi possible de procéder à l'analyse fonctionnelle du comportement, bien que cette analyse soit plus typiquement utilisée dans les études de cas ou les interventions individualisées.

Dans tous les cas, les comportements que l'on souhaite observer doivent être préalablement identifiés et définis de manière opérationnelle. La fidélité des données observationnelles de même que le calcul des différents indices d'accord interjuge sont intimement associés à l'opérationnalisation des comportements à observer (Aspland & Gardner, 2003; Cone, 1999; Gardner, 2000; Hintze, 2005; Margolin et al., 1998). Ainsi, une description précise et détaillée des comportements favorisera une plus grande fidélité des données recueillies. Par ailleurs, les cibles d'observation doivent également être définies en fonction de la technique d'échantillonnage retenue (pour une recension de ces techniques, voir Altmann & Altmann, 2003; Coutu et al., 2005; Yoder & Symons, 2010) et du niveau d'analyse privilégié (micro ou macroscopique). Un système d'observation microscopique permet de consigner la fréquence, l'intensité et la durée des comportements ciblés, l'émetteur et la cible de chacun des comportements manifestés ainsi que la réponse des pairs aux comportements émis par les enfants. Un tel système permet notamment l'analyse séquentielle des données observationnelles et l'identification des contingences sociales associées à la manifestation des comportements (Kerig, 2001; Snyder et al., 2006). En contrepartie, un système d'observation macroscopique permet d'évaluer de manière plus globale le comportement des enfants pour un intervalle de temps plus ou moins long (Dishion & Granic, 2003). À partir de critères préétablis, les observateurs font un travail de synthèse et d'interprétation afin de qualifier ce qu'ils ont vu. Ce niveau macroscopique peut être intéressant afin d'englober des interactions ou une situation se passant sur une longue période de temps. Il peut aussi servir d'indicateur général afin de cibler si une sphère du

fonctionnement est problématique, auquel cas il convient d'examiner de façon plus microscopique ce qui se passe dans cette sphère spécifique (Cone, 1999). Des évaluations récentes de programmes d'intervention mis en place auprès d'enfants agressifs combinent ces deux niveaux d'analyse (CPPRG, 1999; Webster-Stratton et al., 2008). Ceci démontre que les deux niveaux d'analyse peuvent fournir des informations pertinentes, voire complémentaires (Dishion & Granic, 2003).

Choix du contexte (et de la tâche) d'observation.

Puisque les comportements ciblés par l'évaluation (p. ex. agressivité ou habiletés sociales) se manifestent essentiellement lors d'interactions entre pairs, l'enfant doit être observé en compagnie d'autres enfants. Au moins trois contextes d'observation sont possibles. L'observation peut être menée: 1) en compagnie des pairs qui participent à l'intervention; 2) en milieu naturel ou 3) en laboratoire. Dans chacun de ces contextes, les enfants peuvent être observés en interaction libre ou dans le cadre d'une tâche structurée. Les avantages et inconvénients de chacune de ces possibilités lorsqu'elles sont utilisées pour évaluer l'impact d'une intervention sont discutés dans les prochains paragraphes.

Plusieurs modalités d'intervention offertes aux enfants agressifs se déroulent en groupes sur plusieurs séances. Les interventions axées sur l'entraînement aux habiletés sociales nécessitent que les enfants puissent pratiquer certaines habiletés entre eux ou encore prendre part à des jeux de rôles (Bierman & Greenberg, 1996; Nangle et al., 2002). De plus, en les regroupant avec des pairs prosociaux, les enfants agressifs peuvent adopter certains de leurs comportements par simple effet de modelage (Boisjoli, Vitaro, Lacourse, Barker & Tremblay, 2008). Il pourrait être tentant pour l'expérimentateur de choisir d'évaluer l'impact de son intervention en observant le comportement de l'enfant ciblé lorsqu'il interagit avec les pairs de ce groupe au début de l'intervention et de nouveau avec ces mêmes pairs quelques

semaines (ou mois) plus tard à la fin de l'intervention. Tout changement dans le comportement de l'enfant pourrait alors être interprété comme étant un effet positif de l'intervention.

Cette procédure d'observation présente l'avantage d'être facile à utiliser puisque les enfants sont sur place et que l'expérimentateur n'a pas à se déplacer pour les observer. Cependant, il serait risqué de conclure à l'efficacité d'une intervention sur la simple base des résultats de ces observations. D'une part, dans ces conditions, il n'est pas possible de constituer un groupe contrôle équivalent (il faudrait alors observer des enfants agressifs ayant participé avec la même intensité à un programme d'activités autre impliquant un groupe de pairs constant). D'autre part, des décennies de recherche évaluative dans le domaine de l'entraînement aux habiletés sociales ont mis en évidence les problèmes de généralisation et de maintien qui y sont associés (Bierman & Greenberg, 1996; Gresham et al., 2001; Maag, 2006). Les enfants qui participent à ces interventions semblent faire des apprentissages qui se limitent au contexte de ces groupes. Ils parviennent plus difficilement à généraliser ces acquis dans leur milieu de vie, c'est-à-dire dans leurs interactions avec leur groupe de pairs naturel. Il s'avère donc nécessaire d'observer l'enfant dans d'autres contextes en compagnie de d'autres pairs avant de tirer des conclusions sur les effets de l'intervention.

De façon générale, l'observation directe du comportement de l'enfant ne devrait pas se dérouler dans un endroit où il se retrouve uniquement à des fins d'intervention (p. ex. groupe d'intervention ou clinique). En plus du problème discuté précédemment, il est possible que l'enfant adapte son comportement à ce contexte, notamment en fonction des attentes des intervenants. D'ailleurs, peu importe l'endroit où se déroule l'observation, il est préférable que les intervenants ne soient pas présents lorsqu'on observe le comportement de l'enfant. Il est recommandé que l'observation des enfants dans le contexte de l'intervention soit réservée

à l'évaluation de la mise en oeuvre (implantation) des interventions (Snyder et al., 2006) et à l'examen des processus thérapeutiques en jeu (Dishion, Poulin, & Burraston, 2001).

Une deuxième possibilité est d'observer le comportement de l'enfant dans son milieu naturel en interaction avec des pairs qui lui sont familiers. La cour de récréation, les terrains de jeu ou encore le service de garde sont des exemples de contextes où l'enfant interagit librement avec ses pairs. La classe constitue également un contexte d'observation intéressant, bien que le niveau de structure y soit plus élevé et que les interactions entre pairs y soient davantage régulées. Le milieu naturel offre l'avantage de la représentativité du monde réel. Sur la cour d'école, par exemple, les enfants choisissent avec qui ils interagissent et sont en présence de leurs vrais amis (Pellegrini, 2001). Les observations menées dans ce contexte permettent donc de vérifier si l'enfant est parvenu à généraliser à son véritable univers social les compétences qu'il a acquises suite à l'intervention.

Par ailleurs, l'observation du comportement en milieu naturel offre également l'avantage de pouvoir être effectuée auprès d'enfants d'une condition contrôle, en plus de ceux ayant été exposés à l'intervention. De plus, il peut également être possible d'établir une norme pour chacun des comportements ciblés en se basant sur les comportements manifestés par l'ensemble des autres enfants (ou encore un sous-groupe d'enfants représentatifs). L'ampleur des changements observés auprès des participants suite à l'intervention pourrait ainsi être appréciée à la lumière des normes établies. À titre d'exemple, le *Systematic Screening for Behavior Disorders* (SSBD; Walker & Severson, 1990) est une procédure d'évaluation en entonnoir combinant à la fois la passation de questionnaires et l'observation in vivo des comportements lors de jeux libres et en classe afin de dépister les enfants à risque de problèmes intériorisés et extériorisés.

L'utilité de l'observation en milieu naturel pour déterminer l'efficacité d'une intervention qui vise à diminuer (ou à augmenter) la manifestation de certains comportements peut toutefois être compromise par le problème de *censure* identifié par Stoolmiller et ses collègues (2000). Ainsi, il est possible que la fréquence naturelle d'émission du comportement ciblé par l'observation soit très basse et que ce comportement ne survienne pas pendant la période d'observation. Par exemple, des comportements agressifs à haute intensité comme se bagarrer ou encore des habiletés sociales spécifiques comme partager ou coopérer ne sont pas émis à tout moment. Un comportement qui ne survient pas durant la période d'observation n'est pas équivalent à une absence de comportement; ceci peut simplement signifier que le comportement n'a pas eu le temps de se manifester pendant cette période (Dishion & Granic, 2003). Dans cette perspective, lorsque l'occurrence d'un comportement dans le milieu naturel est faible, il est nécessaire d'accroître le temps d'observation auprès des sujets afin de réduire l'effet de censure. Dans le contexte de l'évaluation d'une intervention, la censure peut mener à une réduction de la taille d'effet et peut donc faire en sorte qu'on choisisse d'abandonner une intervention qui pourrait en réalité être efficace. Dans le cadre de leur évaluation du programme de prévention *LIFT*, Stoolmiller et ses collègues (2000) ont observé les comportements agressifs des enfants sur la cour de récréation et ont été confrontés à un problème de censure. L'utilisation d'un modèle analytique qui tenait compte de ce problème de censure (approche TOBIT par équations structurelles) leur a permis d'obtenir des tailles d'effet nettement supérieures comparativement à un modèle analytique qui ne prenait pas en considération ce problème de censure.

Une autre façon de contourner le problème de censure consiste à placer l'enfant dans une situation d'observation structurée dans laquelle il sera plus susceptible de manifester le comportement ciblé. Bien que ces procédures d'observation structurée puissent être mises en

place dans le milieu naturel des enfants, elles sont souvent utilisées en laboratoire où il est possible de standardiser le contexte d'interaction entre les enfants. Des tâches structurées impliquant des interactions sociales sont conçues et mises en place par l'expérimentateur et ont pour objectif d'optimiser les conditions pouvant mener à l'émission du comportement. En restreignant certains de ses comportements et en plaçant certaines demandes à l'enfant, les tâches structurées peuvent faciliter l'émergence de comportements peu fréquents que l'on souhaite observer (Aspland & Gardner, 2003; Cone, 1999; Lindahl, 2001). En mettant en place une structure dans le déroulement de la tâche d'observation, l'expérimentateur souhaite diminuer l'étendue des influences dues aux différentes situations possibles (Gardner, 2000). De telles tâches permettent également une meilleure comparaison entre les participants puisqu'ils ont ainsi été évalués dans un contexte le plus uniforme possible (c.-à-d. procédure d'observation standardisée; Aspland & Gardner, 2003; Lindahl, 2001).

Le recours à une tâche d'observation structurée présente certains avantages. Une illustration provient du domaine conjugal. Les chercheurs ont constaté que le fait d'observer des couples durant leurs activités quotidiennes était beaucoup moins révélateur que de leur demander de discuter d'un sujet de désaccord. En effet, le fait d'orienter la discussion vers un sujet plus délicat permet de mettre à jour certaines dynamiques difficiles à observer hors de ce contexte (Dishion & Granic, 2003). Le même principe s'applique pour l'observation des comportements perturbateurs chez les enfants. Certaines tâches d'observation peuvent permettre de mieux discriminer les enfants hyperactifs de ceux qui ne le sont pas (Gardner, 2000). En effet, lorsque les dyades mère-enfant étaient observées dans un contexte de jeu libre, aucune différence n'était rapportée entre les deux groupes d'enfants. Cependant, lorsqu'on avait recours à des tâches plus structurées tendant vers un but, des différences émergeaient. Enfin, comme le souligne Gardner, les chercheurs font souvent la supposition

que leur tâche est en lien avec ce qui se passe dans le milieu naturel, mais peu de chercheurs comparent vraiment les deux situations. Cette précaution permettrait pourtant de gagner au plan de la représentativité réelle de la tâche. Le choix de la tâche structurée devrait se baser sur des preuves que la gestion harmonieuse de cette tâche peut avoir des impacts sur les relations interpersonnelles de l'enfant (Foster et al., 1993). Toujours selon Foster et ses collègues, les tâches de cet ordre seraient idéales pour vérifier si les enfants ont fait des acquis suite à un entraînement aux habiletés sociales.

Deux procédures peuvent être utilisées en laboratoire: 1) l'enfant et ses pairs sont observés en interaction dans des situations de jeu structuré et 2) une mise en situation est créée avec des pairs complices afin d'observer les réactions de l'enfant. Dans la première procédure, les enfants sont placés dans un contexte d'interaction susceptible de favoriser la manifestation des comportements souhaités. Le choix des jouets et l'arrangement du local d'observation peuvent être planifiés de façon à favoriser l'émission de comportements d'agressivité (p. ex. nombre de jouets limités, certains jouets plus attrayants; Dodge, Coie, Pettit, & Price, 1990). La seconde procédure implique l'exposition de l'enfant ciblé à une situation sociale minutieusement planifiée et étroitement contrôlée qui implique la participation de pairs complices. L'objectif est de créer un contexte standardisé pouvant favoriser au maximum la possibilité d'émission du comportement par l'enfant. Les pairs complices sont entraînés à se comporter d'une façon très précise par l'expérimentateur. Le comportement de l'enfant ciblé dans cette situation constitue alors l'objet de l'observation. Des situations de provocation (p. ex. le pair complice détruit une construction de l'enfant) sont utilisées dans le but de favoriser l'émission de réactions agressives (voir Vitaro, Pelletier, & Coutu; 1989). Des comportements plus positifs peuvent également être observés dans des situations où l'on demande à l'enfant de se joindre à une dyade d'enfants (complices) en train

de jouer et que l'on observe les stratégies qu'il adopte pour se faire accepter (p. ex. Putallaz & Wasserman, 1990).

Ces procédures d'observation en laboratoire sont utilisées par plusieurs équipes de chercheurs qui se consacrent à l'étude du développement social des enfants. Cependant, elles sont plus rarement utilisées à des fins d'évaluation d'une intervention. Plusieurs facteurs pourraient contribuer à expliquer cette situation. Premièrement, l'observation en laboratoire présente un haut niveau de complexité sur le plan logistique et peut entraîner des coûts élevés. Il faut obtenir l'accord des parents de tous les enfants impliqués (outre l'enfant ciblé), s'assurer du transport de chacun au laboratoire avant et après l'intervention, etc. Deuxièmement, malgré la qualité des contrôles exercée, la validité écologique des observations en laboratoire et la possibilité de généraliser les résultats obtenus sont encore débattues.

Lorsqu'une mesure observationnelle dans un contexte de tâche structurée est utilisée pour vérifier si une intervention a permis de modifier (diminuer ou augmenter) certains des comportements de l'enfant, une attention doit être portée au type de tâche choisi, particulièrement si l'on souhaite utiliser la même tâche avant et après l'intervention. Lorsqu'une tâche d'observation implique une quelconque forme d'apprentissage (p. ex. les enfants doivent résoudre un problème ou encore construire quelque chose), elle peut poser problème si elle est utilisée à deux reprises puisqu'elle peut induire naturellement des résultats à la hausse sans témoigner d'un effet de l'intervention. Le problème inverse peut également survenir lorsqu'une tâche peu intéressante est utilisée puisque les enfants risquent de s'ennuyer et de s'en désintéresser lorsqu'ils y participent une seconde fois. Toutefois, le problème lié à l'habituation à la mesure ne se pose pas si les participants ne connaissent pas les comportements qui sont évalués (Coutu et al., 2005).

Problèmes pouvant être rencontrés lors de l'observation et recommandations

Différents problèmes peuvent être rencontrés lors de la pratique de l'observation directe. Les principaux sont les biais, la réactivité et la variabilité des comportements. Harris et Lahey (1982) offrent une bonne recension des biais survenant lors de l'observation directe et des façons de les contrôler. Règle générale, un système d'observation pas trop complexe et bien opérationnalisé, un bon entraînement des observateurs ainsi que des séances de recalibration au besoin permettent d'obtenir un accord interjuge acceptable (Harris & Lahey, 1982; Hoyt & Kerns, 1999).

La réactivité réfère à l'influence possible de la procédure d'observation (ou de la présence de l'observateur) sur le comportement des participants (Aspland & Gardner, 2003). Les comportements observés risqueraient alors de ne pas être représentatifs de ce qui se passe en dehors des périodes d'observation. Bien qu'il soit possible d'évaluer la réactivité afin de mesurer son impact sur les comportements visés (Dishion & Granic, 2003; Gardner, 2000), il est davantage recommandé d'essayer de la diminuer de différentes façons. Il est possible d'accorder un temps d'habitation aux procédures d'observation, de faire en sorte que peu d'observateurs soient présents et qu'ils ne varient pas d'une séance à l'autre ou encore d'utiliser des techniques d'enregistrement peu intrusives comme celle développée par Pepler et ses collègues (Craig, Pepler, & Atlas, 2000; Hawkins, Pepler, & Craig, 2001; Pepler & Craig, 1995) pour l'observation sur la cour de récréation.

La grande variabilité des comportements peut également poser problème. En effet, comme le nombre d'observations est souvent limité à peu d'échantillons de temps avant et après une intervention (parfois un seul échantillon pour chaque temps de mesure), il devient difficile de tirer des conclusions et de constater l'effet d'une intervention sur le taux d'agressivité ou sur une autre cible d'intervention. Si le comportement est très variable d'un

moment d'observation à l'autre, la variabilité diminue la fidélité des observations en ajoutant une source d'erreur. Selon Stoolmiller et ses collègues (2000), cette situation peut mener à une erreur de type II, soit au rejet d'une intervention qui fonctionne.

Deux solutions sont proposées pour pallier la variabilité des mesures. Premièrement, Pellegrini (2001) propose d'observer l'enfant dans un contexte très précis, ce qui permet d'augmenter la précision des inférences, mais qui risque toutefois de ne pas être représentatif du comportement dans d'autres situations. Deuxièmement, plusieurs échantillons d'observations devraient être recueillis. En effet, Eddy et ses collègues (1998) soulignent que le changement dans le contexte d'une intervention n'est pas nécessairement linéaire et que la prise en compte de plusieurs moments de mesure pourrait aider à qualifier le changement. Selon ces chercheurs, un minimum de trois temps de mesure est nécessaire afin de préciser la direction du changement et le nombre de mesures augmenterait cette précision, en plus de fournir des informations sur les mécanismes de changement potentiels qui pourraient prendre place au cours de l'intervention. Toutefois, afin d'éviter de trop nombreuses mesures, Eddy et ses collègues recommandent de mesurer le changement à des temps théoriquement importants (p. ex. chaque semaine dans le cas où il y aurait une séance d'intervention par semaine ou encore selon les théories du développement pour le suivi post-intervention).

Conclusion

Les chercheurs ou praticiens intéressés à évaluer l'impact d'une intervention sur le comportement des enfants à partir d'une mesure directe basée sur l'observation peuvent choisir d'utiliser un instrument déjà existant ou encore préférer développer une mesure qui répondra spécifiquement à leurs propres besoins. Plusieurs auteurs proposent des recensions d'outils permettant l'observation des comportements sociaux entre enfants (p. ex. Brown, Odom, & Holcombe, 1996; Cone, 1999; Pellegrini, 2001). Il est toutefois nécessaire

d'identifier préalablement les centrations de l'observation et le contexte interpersonnel à l'intérieur duquel les enfants seront observés. Les cibles d'observation devraient idéalement être en lien avec le contenu et les objectifs spécifiques poursuivis par l'intervention. Par ailleurs, différents contextes d'observation peuvent être nécessaires afin d'évaluer l'impact d'une intervention sur différentes dimensions du comportement de l'enfant. Alors que l'observation en milieu naturel sur la cour de récréation et dans la classe se prête bien à l'évaluation des comportements perturbateurs, un contexte d'observation structuré peut s'avérer plus adéquat lorsqu'il s'agit d'évaluer l'impact d'une intervention sur la manifestation de comportements prosociaux. Dans tous les cas, il s'avère également pertinent de recueillir des données indirectes afin de voir si les changements observés dans ces différents contextes d'observation s'accompagnent d'un changement de perception chez les parents, les enseignants et le groupe de pairs.

Bien que l'observation directe soit plus complexe à mettre en place et plus coûteuse, elle apporte des avantages indéniables dans un contexte d'évaluation d'impact d'une intervention. L'observation directe est plus riche, plus précise et moins sujette aux biais que les mesures indirectes. Elle permet de plus d'avoir accès à certains comportements qui ne seraient pas nécessairement rapportés dans des questionnaires. Toutefois, l'argument principal justifiant le recours à l'observation directe des comportements sociaux reste sa grande sensibilité au changement. Les programmes d'intervention mis en place auprès des enfants nécessitent beaucoup de ressources. La précision lors de leur évaluation est donc d'une importance capitale. De plus en plus, la recherche dans le domaine de l'évaluation de programmes de prévention reconnaît l'apport de l'observation directe. Cette contribution de l'observation directe gagnerait à être reconnue non seulement en recherche, mais aussi dans le milieu clinique.

Références

- Achenbach, T.M. (1991) *Integrative guide for the 1991 CBCL/4-18, YSR and TRF profiles*. Burlington: University of Vermont.
- Altmann, S.A. & Altmann, J. (2003). The transformation of behaviour field studies. *Animal Behaviour*, 65, 413-423.
- Aspland, H., & Gardner, F. (2003). Observational measures of parent-child Interaction: An introductory review. *Child and Adolescent Mental Health*, 8(3), 136-143.
- Bierman, K. L. & Greenberg, M. T. (1996). Social skills training in the Fast Track program. In R. D. Peters & R. J. McMahon (Eds.), *Preventing Childhood Disorders, Substance Abuse, and Delinquency*, 65-89.
- Bierman, K. L., & Furman, W. (1984). The effects of social skills training and peer involvement on the social adjustment of preadolescents. *Child Development*, 55, 151-162.
- Boisjoli, R. & Vitaro, F. (2005). Quels sont les paramètres qui peuvent influencer l'efficacité des programmes d'entraînement aux habiletés sociales auprès d'enfants agressifs? *Revue de Psychoéducation*, 34 (2), 331-354.
- Boisjoli, R., Vitaro, F., Lacourse, É., Barker, E. D., & Tremblay, R. E. (2008). Impact and clinical significance of a preventive intervention for disruptive boys: 15 - year follow - up. *British Journal of Psychiatry*, 415-419.
- Breton, J.-J., Bergeron, L., Valla, J.-P., Berthiaume, C., Gaudet, N, Lambert, et al. (1999). Quebec Child Mental Health Survey: Prevalence of DSM-III--R mental health disorders. *Journal of Child Psychology and Psychiatry*. 40 (3), 375-384.
- Brown, C. H., Wang, W., Kellam, S.G., Muthen, B.O., Petras, H., Toyinbo, P., Poduska, J., et al. (2008). Methods for testing theory and evaluating impact in randomized field

trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug Alcohol Depend*, 95 (Suppl. 1), S74-S104.

- Brown, W.H., Odom, S.L. & Holcombe, A. (1996). Observational assessment of young children's social behavior with peers. *Early Childhood Research Quarterly*, 11(1), 19-40.
- CPPRG (1999). Initial impact of the Fast Track Prevention Trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology*, 67 (5), 631-647.
- Card, N.A., Hodges, E.V.E., Little, T.D. & Hawley, P.H. (2005). Gender effects in peer nominations for aggression and social status. *International Journal of Behavioral Development*, 29(2), 146-155.
- Chambless, D.L. & Hollon, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66 (1), 7-18.
- Cillessen, A. H. N., Bukowski, W. M., & Haselager, G. J. T. (2000). Stability of sociometric categories. *New Directions for Child and Adolescent Development*, 88, 75-93.
- Cone, J. D. (1999). Observational assessment: Measure development and research issues. In P. C. Kendall, J. N. Butcher & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology* (pp. 183-223). New York: John Wiley and Sons, Inc.
- Coutu, S., Provost, M.A. & Bowen, F. (2005). L'observation systématique des comportements: Une démarche structurée pour une évaluation valide. In S. Bouchard & C. Cyr (Eds). *Recherche psychosociale: Pour harmoniser recherche et pratique*. Sainte-Foy: Presses de l'Université du Québec.
- Craig, W.M., Pepler, D., & Atlas, R. (2000). Observations of bullying in the playground and in the classroom. *School Psychology International*, 21, 22-36.

- Dishion, T. J., & Granic, I. (2003). Naturalistic observation of relationship processes. In E. M. Heiby & S. N. Haynes (Eds.), *Comprehensive handbook of psychological assessment* (Vol. 3, pp. 143-161). New York: Wiley.
- Dishion, T. J., Poulin, F., & Burraston, B. (2001). Peer group dynamics associated with iatrogenic effects in group interventions with high-risk young adolescents. In D. W. Nangle & C. A. Erdley (Eds.). *New directions for child and adolescent development: Friendship and psychological adjustment* (pp. 79-92). San Francisco: Jossey-Bass.
- Dodge, K. A., Coie, J. D., Pettit, G. S., & Price, J. M. (1990). Peer status and aggression in boys' groups: Developmental and contextual analyses. *Child Development*, *61*, 1289-1309.
- Enquête longitudinale nationale sur les enfants et les jeunes. (1999). *Grandir au Canada, Enquête longitudinale nationale sur les enfants et les jeunes*. Ottawa : Développement des ressources humaines Canada. Statistique Canada.
- Eddy, J. M., Dishion, J. T., & Stoolmiller, M. (1998). The analysis of intervention change in children and families: Methodological and conceptual issues embedded in intervention studies. *Journal of Abnormal Child Psychology*, *26*, 53-69.
- Epstein, M. H., & Sharma, J. M. (1998). *Behavioral and Emotional Rating Scale (BERS): A strength-based approach to assessment - Examiner's manual*. Austin: PRO-ED.
- Foster, S.L., Inderbitzen, H.M. et Nangle, D.W. (1993). Assessing acceptance and social skills with peers in childhood: Current issues. *Behavior Modification*, *17*, 255- 286.
- Gardner, F. (2000). Methodological issues in the direct observation of parent-child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review*, *3*(3), 185-198.

- Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System test manual*. Circle Pines, MN American Guidance Service.
- Gresham, F.M., Sugai, G. & Horner, R.H. (2001). Interpreting outcomes of social skills training for students with high-incidence disabilities. *Exceptional Children*, 67 (3), 331-344.
- Harris, H. C., & Lahey, B. B. (1982). Recording system bias in direct observational methodology: A review and critical analysis of factors causing inaccurate coding behavior. *Clinical Psychology Review*, 2, 539-556.
- Hansen, D.J., Nangle, D.W., & Ellis, J.T. (1996). Reconsideration of the use of peer sociometrics for evaluating social-skills training: Implications of an idiographic assessment of temporal stability. *Behavior Modification*, 20, 281-299.
- Hawkins, D.L., Pepler, D.J., et Craig W.M. (2001). Naturalistic observations of peer interventions in bullying. *Social Development*, 10, 512-527.
- Hintze, J.M. (2005). Psychometrics of direct observation. *School Psychology Review*, 34, 507-519.
- Hoyt, W. T., & Kerns, M.-D. (1999). Magnitude and moderators of bias in observer ratings: A meta-analysis. *Psychological Methods*, 4(4), 403-424.
- Hymel, S. (1986). Interpretations of peer behavior: Affective bias in childhood and adolescence. *Child Development*, 57(2), 431-445.
- Kazdin, A. E. (2005). Child, parent, and family-based treatment of aggressive and antisocial child behavior. Dans E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders* (pp. 445-476), Washington, D.C.: American Psychological Association.

- Kellam, S.G. & Van Horn, Y.V. (1997). Life course development, community epidemiology, and preventive trials: A scientific structure for prevention research. *American Journal of Community Psychology*, 25 (2), 177-188.
- Kerig, P. K. (2001). Introduction and overview: Conceptual issues in family observational research. In D. C. R. Kerr & K. M. Lindahl (Eds.), *Family Observational Coding Systems* (pp. 1-22). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Leflot, G., van Lier, P.A.C., Onghena V. et Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of Abnormal Child Psychology*, 38, 869-882,
- Lindahl, K. M. (2001). Methodological issues in family observational research. In P. K. Kerig & K. M. Lindahl (Eds.), *Family Observational Coding Systems* (pp. 23-32). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Lochman, J. E., Boxmeyer, C. L., Powell, N. P., Barry, T. D. et Pardini, D. A. (2010). Anger control training for aggressive youths. Dans J. R. Weisz et A. E. Kazdin Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed.), (pp. 227-242). New York, NY, US: Guilford Press.
- Loeber, R., Burke, J.D., Lahey, B.B., Winters, A., & Zera. M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years, Part I. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 1468-1484.
- Maag, J. W. (2006). Social Skills Training for Students With Emotional and Behavioral Disorders: A Review of Reviews. *Behavioral Disorders*, 32(1), 5-17.
- Masten, A.S., Morison, P. & Pellegrini, D.S. (1985). A revised class play method of peer assessment. *Developmental Psychology*, 21(3), 523-533.

- Margolin, G., Oliver, P. M., Gordis, E. B., O'Hearn, H. G., Medina, A. M., Ghosh, C. M., et al. (1998). The nuts and bolts of behavioral observation of marital and family interaction. *Clinical Child and Family Psychology Review*, 1(4), 195-213.
- Merrell, K. W. (1993). *School Social Behavior Scales*. Iowa City: Assessment-Intervention Resources.
- Nangle, D.W., Ellis, J.T., Christopher, J.S., & Hansen, D.J. (1993). Sociometric peer nominations: An idiographic examination of temporal stability. *Education and Treatment of Children*, 16, 175-186.
- Nangle, D. W., Erdley, C. A., Carpenter, E. M., Newman, J. E. (2002). Social skills training as a treatment for aggressive children and adolescents: A developmental - clinical integration. *Aggression and Violent Behavior*, 7, 169-199.
- Pellegrini, A. D. (2001). Practitioner review: The role of direct observation in the assessment of young children. *Journal of Child Psychology and Psychiatry*, 42(7), 861-869.
- Pepler, D. J., & Craig, W. M. (1995). A peek behind the fence: Naturalistic observations of aggressive children with remote audiovisual recording. *Developmental Psychology*, 31(4), 548-553.
- Putallaz, M. et Wasserman, A. (1990). Children's entry behavior. Dans S.R. Asher et J.D. Coie (Eds.), *Peer rejection in childhood, Cambridge studies in social and emotional development* (pp. 60-89). New York, NY, US: Cambridge University Press
- Saudino, K. J., Ronald, A., & Plomin, R. (2005). The etiology of behavior problems in 7-year-old twins: Substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. *Journal of Abnormal Child Psychology*, 33(1), 113-130.

- Snyder, J., Reid, J., Stoolmiller, M., Howe, G., Brown, H., Dagne, G., et al. (2006). The role of behavior observation in measurement systems for randomized prevention trials. *Prevention Science*, 7(1), 43-56.
- Snyder, J., Schrepferman, L., McEachern, A., Barner, S., Johnson, K., & Provines, J. (2008). Peer deviancy training and peer coercion : Dual processes associated with early-onset conduct problems. *Child Development*, 79, 252-268.
- Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review*, 26(3).
- Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology*, 68(2), 296-306.
- Swindells, D., & Stagnitti, K. (2006). Pretend play and parents' view of social competence: The construct validity of the Child-Initiated Pretend Play Assessment. *Australian Occupational Therapy Journal*, 53, 314-324.
- Tremblay, R. E. (2011). Origins, Development, and Prevention of Aggressive Behavior. In Daniel P. Keating, *Nature and Nurture in Early Child Development* (pp. 169-187). New York, NY : Cambridge University Press.
- Vitaro, F. (2001). Évaluation des programmes de prévention: Principes et procédures. In F. Vitaro & C. Gagnon (Eds.), *Prévention des problèmes d'adaptation chez les enfants et les adolescents, Tome I* (pp. 67-100). Sainte-Foy: Presses de l'Université du Québec.
- Vitaro, F., Pelletier, D., & Coutu, S. (1989). Stratégies de résolution de problèmes sociaux hypothétiques: impact de provocations réelles sur les réponses d'enfants en début de scolarisation. *Revue Internationale de Psychologie Sociale*, 2, 513-529.
- Vitaro, F. & Tremblay, R. E. (2008). Clarifying and maximizing the usefulness of targeted

preventive interventions. Dans Michael Rutter et al. (Eds.) *Rutter's child and adolescent psychiatry* (5th ed.), (pp. 989-1008). Wiley-Blackwell, xv, 1230 pp.

Walker, H. M., & McConnell, S. (1995). *Walker-McConnell Scale of Social Competence and School Adjustment, Elementary Version*. San Diego: Singular.

Walker, H. M., & Severson, H. H. (1990). *Systematic Screening for Behavior Disorders*. Longmont: Sopris West.

Webster-Stratton, C., Reid, M.J. & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the Incredible Years Teacher and Child Training Programs in high-risk schools. *Journal of Child Psychology and Psychiatry*, 49 (5), 471-488

Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York: Springer Publishing Company.