

STRUCTURER LE JUGEMENT PROFESSIONNEL : LES VÉRITABLES ENJEUX

par **Martine H.-Evans**

Professeur à l'université de Reims

James Bonta

Docteur en psychologie clinique, consultant, Corrections and criminal behavior

Marije Keulen de Vos

Docteur en psychologie forensique, chercheur senior au forensic psychiatric center de Rooyse Wissel, Venray, Pays-Bas

Olivier Vanderstukken

Psychologue clinicien, Ph.D, Service médico-psychologique (SMPR) de Lille-Annoeullin, Unité régionale de soins aux auteurs de violence sexuelle (URSAVS), Nord Pas-de-Calais, CHRU de Lille

Massil Benbouriche

Psychologue clinicien, Ph.D, Département de Psychologie, Wayne State University (5057 Woodward, Detroit, MI 48202, USA)

Cinq ans après les premiers articles parus relatifs à l'évaluation structurée¹, celle-ci fait l'objet de certaines critiques² sur lesquelles nous souhaiterions revenir. Nous n'accorderons pas d'attention aux arguments avançant « la société/la gouvernance du risque », tant il est évident que la réalité objective, comme la diversité des pratiques professionnelles³, est hautement plus complexe, très variable d'un État à un autre et fluctuante au cours de l'histoire⁴, notamment en Europe⁵. Nous ne relèverons pas plus l'incompréhension parfois observée de ce que constituent les méta-analyses (i. e. des revues complètes reprenant et totalisant l'ensemble des calculs de recherches sélectionnées sur la base de leurs qualités méthodologiques et faisant le point sur un sujet donné) et ce qu'elles représentent⁶. Enfin, rappelons que l'évaluation n'est que l'étape préliminaire du traitement (nous utiliserons ci-après le terme de traitement de manière générique, pour décrire toutes les formes d'intervention, suivi, probation, programmes, soins, etc.) et que derrière les modèles d'évaluation structurée, il existe aussi et surtout des modèles de traitement eux-mêmes structurés et validés⁷. Nous nous limiterons dès lors à analyser brièvement un certain nombre d'arguments avancés par les détracteurs de l'empirisme.

■ « C'est de la pseudoscience »

Les opposants aux outils et à la structuration du jugement professionnel rejettent les résultats des méta-analyses comme relevant de la « pseudoscience ». Pourtant, la définition même de la pseudoscience est de ne pas s'appuyer sur des données empiriques et des théories solides, mais de se parer des vertus de l'expertise personnelle et/ou de l'idéologie. En l'espèce, la pseudoscience consiste même à nier que la science puisse exister dans le domaine de l'humain ; il s'agit de faire fi des montagnes de méta-analyses qui ont été réalisées dans le champ de la psychologie, de la psychologie sociale, de la criminologie et du travail social, précisément depuis que ces sciences humaines se sont résolument éloignées de la pseudoscience et de l'idéologie, et ce notamment sous l'impulsion de pionniers comme Campbell.

■ « On ne peut faire entrer les humains dans des statistiques »

Un argument souvent invoqué par ceux qui ne comprennent pas la démarche scientifique de l'évaluation des risques de récidive, consiste à dire que, chaque cas étant unique, il est impossible

de les faire tous « entrer dans des statistiques ». Chaque cas est certes très différent et tout l'objet d'une évaluation adaptée (surtout lorsqu'elle porte sur les besoins criminogènes, voire intègre des facteurs de protection) est de s'approcher au plus près de la situation exacte plutôt que de l'enfermer dans une vision globale ou qui soit le fruit des heuristiques propres au praticien. Affirmer toutefois que les causes de la délinquance sont totalement uniques pour chaque personne contredit l'aptitude à penser scientifiquement. Pour utiliser une métaphore médicale – ce qui ne veut naturellement pas dire que nous avons une lecture médicale de la délinquance – c'est comme affirmer que la tuberculose n'est pas identifiable parce que chaque patient est unique. Certes, il va de soi qu'une personne chétive, sans accès aux vitamines, vivant dans la saleté et l'humidité aura un moins bon pronostic quant à sa tuberculose. Cela

(1) M. H.-Evans, Outils d'évaluation : sortir des fantasmes et de l'aveuglement idéologique, *AJ pénal* 2012. 75 ; M. Benbouriche, A. Vantéjoux, M. Lebourgault et A. Hirschelmann, L'évaluation du risque de récidive en France : expérience et attitudes des conseillers pénitentiaires d'insertion et de probation, *RI crim. et pol. techn.* 2012. 305 ; M. Benbouriche et J.-P. Guay, La prévention de la récidive en France : les principes d'une évaluation et d'une intervention, *Pratiques Psychologiques* 2015. 213.

(2) S. Raoult, Récidive : trois ans après la conférence de consensus, pourquoi il n'y a pas de consensus. *AJ pénal* 2016. 25 ; CGT Insertion et Probation, Les dangers de l'évaluation, *Tract.* Juin 2017 ; CGT, L'évaluation pas à n'importe quel prix, *Fonction publique*, juin 2017, p. 155.

(3) R.C. Mawby et A. Worrall, *Doing probation work. Identity in a criminal justice occupation*, Abingdon, Routledge, 2013 ; G. Graham, *Rehabilitation work. Supporting desistance and recovery*, Abingdon, Routledge, 2017 ; P. Tomczak, *The penal voluntary sector*, Abingdon, Routledge, 2017.

(4) G. Robinson et F. McNeill (dir.), *Community Punishment. European Perspective*, Abingdon, Routledge et COST UE, 2015.

(5) S. Snacken et S. Dumortier (eds), *Resisting Punitiveness in Europe ? Welfare, Human Rights and Democracy*, London, Routledge, 2012.

(6) V. not., à propos du champ social, H. Soydan et L.A. Palinkas, *Evidence-Based Practice, in Social Work. Development of a New Professional Culture*, Abingdon, Routledge, 2014.

(7) J. Bonta et D.A. Andrews, *The Psychology of Criminal Conduct*, London, Routledge, 6^e éd., 2017 ; P.M. Yates, D. Prescott et T. Ward, *Applying the Good lives and Self-Regulation Models to Sex-Offender Treatment : A practical Guide for Clinicians*, Brandon, VM, The Safer Society, 2010.

n'empêche pas d'analyser sa pathologie en tant que tuberculose, ni ne doit d'ailleurs conduire à affirmer que la cause de celle-ci est l'humidité.

Il en va de même pour la délinquance. En dépit de lectures différentes relatives à leurs grandes causes, la plupart des auteurs intègrent tout à la fois des facteurs psychologiques, sociaux, environnementaux, familiaux, développementaux, cérébraux, épigénétiques et même évolutionnistes⁸. La vraie complexité est là et est certes loin d'être entièrement traitée par les programmes structurés actuels, mais ceux-ci tendent au moins à s'en saisir et s'en approcher.

Il peut encore exister une confusion entre évaluer et mesurer : évaluer renvoie au processus permettant de bien comprendre le fonctionnement d'une personne sur un ou plusieurs traits, alors que mesurer est le processus par lequel on assigne un nombre à des choses selon des règles déterminées. En ce sens, les outils ne constituent pas une finalité, mais une aide précieuse à l'appréciation des situations et à la détermination des cibles du suivi et du traitement. Enfin, s'appuyer sur la complexité individuelle des situations pour refuser de structurer le jugement – mais aussi la prise en charge – est tristement risible lorsqu'on entend sur quoi s'appuient en réalité les avis classiquement formulés par exemple en commission de l'application des peines (CAP)⁹.

■ « Il y a plein de faux positifs »

Notons au préalable que ce troisième argument se réfère quasiment toujours à un versant du problème et non à son double, celui des faux négatifs. Si l'on

passé sur cet inégal intérêt pour une erreur plutôt qu'une autre, il est pour le moins étonnant que ceux qui évoquent cet argument oublient que l'évaluation a bel et bien lieu de manière quotidienne dans la justice pénale et, mais la liste n'est pas exhaustive, par exemple dans les expertises psychiatriques non structurées et hélas parfois hors la moindre référence aux catégorisations internationalement reconnues, réalisées dans toutes les phases du continuum pénal ou encore dans les rapports transmis par les CPIP au JAP et, surtout, dans les avis donnés en CAP dans des conditions expéditives et sans cadre théorique ou empirique¹⁰. Nous dira-t-on que « avis défavorable, risque de récidive » formulé en quelques secondes présente moins de risques de faux positifs – et de faux négatifs – que l'évaluation actuarielle ou le jugement professionnel structuré (JPS) ? ! Faut-il rappeler, parmi de multiples études sur le jugement professionnel non structuré, qu'il n'est pas plus performant que tirer à pile ou face¹¹ ? Comme le disaient il y a huit ans déjà Andrews et Bonta¹², « s'il y a plus d'une centaine d'études concernant des centaines de professionnels et qui montrent qu'ils ne prédisent pas aussi bien que les outils actuariels, qu'est-ce qui peut bien vous faire penser que vous êtes supérieur aux autres ? ».

■ « Le flair des professionnels est plus sûr »

Il est assez stupéfiant que l'on puisse encore à l'heure actuelle, au vu des quantités industrielles de recherches ayant porté sur le sujet, affirmer que le flair professionnel non structuré est supérieur à une analyse professionnelle structurée. L'on sait en effet depuis les travaux du prix Nobel Kahneman¹³ que les hommes sont particulièrement incompétents à apprécier des situations, des risques ou des personnes. Ils prennent le plus souvent leurs décisions de manière irrationnelle, biaisée, préjugée et faussée. Cela est hélas aussi vrai des professionnels expérimentés, qui vont utiliser leur cerveau dit réflexif, donc puiser dans leurs heuristiques. Les recherches ont ainsi montré, depuis longtemps, que les praticiens jugent comme tout le monde sur la base de biais culturels, relatifs à leur éducation, à leur personnalité ou expérience¹⁴. C'est la raison pour laquelle, depuis des décennies, on prône la structuration du jugement, que ce soit en criminologie ou en psychologie (v. les travaux fondateurs de Meehl, 1954¹⁵).

Parmi les heuristiques les plus courantes¹⁶, on trouve : la *disponibilité*, soit la tendance à attribuer plus de poids à ce que l'on peut facilement retrouver dans sa mémoire ; l'*ancrage*, soit prendre un point de repère qui peut parfaitement être absurde¹⁷, et est en pratique souvent l'avis des autres praticiens (par ex. le parquet) ; la *stimulation*, soit la capacité à imaginer un scénario prédictif (et plus ce scénario surgit aisément et plus la personne tend à croire à ce qu'il suggère) ; l'*excès de confiance* dans son propre jugement et dans la supériorité de son point de vue ; la *représentativité*, soit le fait d'apprécier la probabilité que telle ou telle chose se produise selon le degré de sa similitude à des propriétés de sa population ou domaine d'appartenance. Garb a, par exemple, montré que les psychologues et psychiatres, lorsqu'ils émettent un jugement sur la pathologie mentale, se basent d'abord et avant tout sur leur perception de ce qu'est un patient prototypique de telle ou telle catégorie plutôt que sur des critères validés scientifiquement¹⁸. Ce qui est crucial est que ces heuristiques agissent hors du champ de la conscience des individus : les praticiens ne savent pas qu'elles jouent dans leurs analyses et raisonnements.

Or, parmi ces heuristiques, se trouvent hélas des représentations, biais et préjugés sur d'autres ethnies ou cultures. On sait ainsi que les diagnostics et traitements offerts aux minorités ne sont pas optimaux¹⁹. Par exemple, les personnes noires et latinos américaines

(8) R. Durant et T. Ward, *Evolutionary Criminology. Towards a Comprehensive Explanation of Crime*, Cambridge, MA : Academic Press, Elsevier, 2015.

(9) Ex. réel : « On n'a pas reçu le Kbis donc avis défavorable ».

(10) M. H.-Evans (à paraître), La mise en œuvre de la libération sous contrainte dans le Nord-Est de la France, Rapp. à la Mission droit et justice.

(11) B. J. Ennis et T. R. Litwack, *Psychiatry and the presumption of expertise : Flipping coins in the courtroom*, California Law Review, 62, 1974, 694.

(12) D. Andrews et J. Bonta, *The Psychology of Criminal Conduct. New Providence, NJ, LexisNexis*, 5^e éd., 2010 – traduction en français : *Le comportement délinquant. Analyse et modalités d'intervention*, Agen, Les Presses de l'Enap, 2015, p. 360.

(13) A. Tversky et D. Kahneman, *Judgment under uncertainty : Heuristics and Biases*, Science, 185, 1974, 1124 ; D. Kahneman, *Thinking fast and slow*, London, Penguin, 2012 (version traduite en français : *Système 1/système 2 : les deux faces de la pensée*, Paris, Flammarion).

(14) W. Williams et K. Miller, *The role of personal characteristics in perceptions of dangerousness. Criminal Justice and Behaviour*, 4, 1977, p. 241.

(15) P. Meehl, *Clinical versus Statistical Prediction : A Theoretical Analysis and a Review of the Evidence*, Minneapolis, Minneapolis University of Minnesota Press, 1954. Republié en 1963 et 1996.

(16) V. par ex., pour les juges, G. Gigerenzer et C. Engel (dir.), *Heuristics and the Law*, London, MIT Press, 2004.

(17) B. Englich, T. Mussweiler et F. Strack, *Playing Dice with Criminal Sentences : The Influence of Irrelevant Anchors on Experts'Judicial Decision Making*, Personality & Soc. Psychol. Bull., 32(2), 2006, p. 188.

(18) H.N. Garb, *Cognitive Heuristics and Biases in Personality Assessment*, in L. Heath, R.S. Tindale, E.J. Posavac, F.B. Bryant, E. Henderson-King, Y. Suarez-Balcazar et J. Myers, *Applications of Heuristics and Biases to Social Issues*, New York, Plenum Press, 1994, p. 73.

(19) L. Heath, R.S. Tindale, E.J. Posavac, F.B. Bryant, E. Henderson-King, Y. Suarez-Balcazar et J. Myers, *Applications of Heuristics and Biases to Social Issues*, New York, Plenum Press.

sont bien plus souvent diagnostiqués comme psychotiques²⁰. S'agissant du travail social, une étude expérimentale²¹ a testé sur des étudiants en travail social une vignette qui décrivait le cas d'un jeune de treize ans avec des antécédents de troubles du comportement. Pour l'un des groupes d'étudiants, l'enfant était présenté comme étant blanc ; pour le second, comme étant noir. 59 % des étudiants pensant que l'enfant était noir ont estimé que ses problèmes provenaient de son environnement familial contre 30 % de ceux détenant la vignette présentant l'enfant blanc.

Ni la bonne volonté ni l'expérience professionnelle ne prémunissent contre ces biais et les services de probation ne sont pas plus immunisés que les professionnels de la justice pénale²². Il est bien de l'analyser quantitativement²³, mais on ne les combattra réellement qu'en limitant leur expression au stade du jugement et des modalités de suivi.

Si les outils actuariels ou structurés ne sont certes pas une garantie absolue contre les biais – qui peuvent se loger dans les items, et bien plus fréquemment dans leur cotation, d'où la recommandation fréquente de ne pas être le seul à coter, de former à l'identification des représentations sociales, et d'où notre insistance sur l'application du principe contradictoire (v. *infra*) – ils nous semblent de loin préférables au jugement biaisé et imbibé d'heuristiques « depuis son chapeau » qui en constitue l'alternative.

■ « Les Français sont différents des Canadiens »

Le caractère fondamentalement unique de l'humain expliquerait encore pourquoi les outils structurés ne sauraient être transposables aux Français. Notons cependant que ce curieux « complexe d'Astérix » fait fi de ce qui rassemble fondamentalement la communauté humaine. Il est vrai que la culture est une dimension essentielle, largement sous-estimée et qu'il serait pertinent de disposer d'une validation des outils sur les populations françaises. Notons toutefois qu'il y a déjà eu une évaluation du LS-CMI (Level of Service Inventory-Case Management Inventory) sur les habitants du Québec, qui sont les cousins francophones de la France²⁴. Dans l'ensemble, il y a

L'on sait en effet depuis les travaux du prix Nobel Kahneman que les hommes sont particulièrement incompétents à apprécier des situations, des risques ou des personnes. Ils prennent le plus souvent leurs décisions de manière irrationnelle, biaisée, préjugée et faussée.

d'ailleurs une bonne transférabilité des outils²⁵ aux minorités ethno-culturelles avec certes des aires sous la courbe (*areas under the curve* - AUC²⁶) limitées par rapport à la majorité²⁷, mais surtout en raison de l'insuffisante légitimité des praticiens qui les mettaient ici en œuvre comme des procédures se traduisant par des problèmes classiques d'attrition²⁸, soit un problème de réceptivité. La transférabilité est également bonne chez les femmes²⁹. Les facteurs de risque sont sensiblement les mêmes les concernant³⁰, on a plutôt affaire chez elles à des « chemins » (*pathways*) légèrement différents des hommes.

Ce ne sont donc pas les niveaux de risque ou les besoins criminogènes de la diversité qui sont distincts, mais leur réceptivité. Il conviendrait ainsi d'inviter à réfléchir à une meilleure représentativité des praticiens de la probation³¹, ou à s'appuyer sur la ressource essentielle que représentent les praticiens biculturels, véritables « traducteurs » culturels³². Il faudrait encore proposer des programmes adaptés pour tenir compte de la culture, et développer chez les praticiens la conscience individuelle et institutionnelle de leurs biais personnels ou, tout simplement, de leur « lunettes de blancs occidentaux »³³ afin d'en mieux protéger leurs avis, rapports

et prises de décisions. De même, et nous y reviendrons, conviendrait-il de travailler sur les garde-fous procéduraux qui pourraient ajouter à cette protection.

■ « La délinquance est uniquement un fait social »

La résistance aux outils structurés s'appuie aussi sur une lecture essentiellement sociale du phénomène délinquant et de la récidive qui est présentée avec l'évidence pseudoscientifique du bon sens. Il est alors sous-entendu que tout ira sans doute mieux lors du grand soir révolutionnaire et qu'en attendant, il convient de traiter par du travail social. Loin de nous de dire que le travail social ne participe pas du traitement. Bien au contraire, la disparition du travail social, à la fois du champ de compétences et des normes juridiques, constitue une erreur institutionnelle fondamentale³⁴, pour résumer un propos

(20) R.C. Schwartz et D. M. Blankenship, *Racial disparities in psychotic disorder diagnosis : A review of empirical literature*, World Journal of Psychiatry, 4(4), 2014, p. 133.

(21) J. Owusu-Bempah, *Race, self-identify and social work*, British Journal of Social Work, n° 24(20), 1994, p. 123.

(22) S. Lewis, P. Raynor, D. Smith et A. Wardak, *Race and Probation*, Collumpton, Willan Publishing, 2006.

(23) *Ibid.* ; F. Jobard et S. Névanen, La couleur du jugement. Discriminations dans les décisions judiciaires en matière d'infractions à agents de la force publique (1965-2005), RF sociol., 2007, p. 48, 2, 243. ; F. Jobard, R. Lévy et I. Goris, *Police et minorités visibles : les contrôles d'identité à Paris*, New York Open society Institute, Justice Initiative, 2009.

(24) J.-P. Guay, L'évaluation du risque et des besoins criminogènes à la lumière des données probantes : une étude de validation de la version française de l'inventaire de niveau de service et de gestion des cas – LS/CMI, Revue européenne de Psychologie Appliquée, 2016, p. 66, 199.

(25) J. Bonta, et D.A. Andrews, *The Psychology of Criminal Conduct*, London, Routledge, 6^e éd., 2016, p. 209 s.

(26) AUC est une classification métrique appliquée aux classifications binaires ; c'est l'une des plus fréquemment utilisée. Elle permet de limiter les faux positifs et négatifs. Si un instrument de mesure a un AUC de 1.0, il réalise une prédiction parfaite, ce qui est naturellement impossible ; s'il réalise un AUC de 50, il ne sert à rien car il n'est pas plus efficace que la chance ou tirer à pile ou face.

(27) 67 pour J.A. Wilson et L. Guttierrez, *Does one size fit all ? A meta-analysis examining the predictive ability of the level of service inventory (LSI) with aboriginal offenders*. *Criminal Justice and Behavior*, 41(2), 2014, p. 196 ; 57 pour M.E. Olver, K.C. Stockdale et J.S. Worthmuth, *Thirty Years of Research on the Level of Service Scales : A Meta-Analytic Examination of Predictive Accuracy and Sources of Variability*. *Psychological Assessment*, 26(1), p. 156.

(28) J.S. Wormith et M.E. Olver, *Offender Treatment Attrition and its Relationship with Risk, Responsivity, and Recidivism*, *Criminal Justice and Behavior*, 20(4), 2002, p. 447.

(29) P. Smith, F. T. Cullen et E. J. Latessa, *Can 14,373, women be wrong ? A meta-analysis of the LSI-R and recidivism for female offenders*. *Criminology & Public Policy*, 8, 2009, p. 183.

(30) V. le point récent sur la recherche par C. Kruttschnitt, *2015 Presidential Address to the American Society of Criminology : The politics, and place, of gender in research on crime*. *Criminology*, 54(1), 2016, p. 8.

(31) Lewis et a., préc. note 22.

(32) M.Y. Brannen et F. Lee, *Bridging cultural divides : Traversing organizational and psychological perspectives on multiculturalism*, in V. Benet-Martinez et Y.-Y. Hong (dir.), *The Oxford Handbook of Multicultural Identity*, New York, Oxford University Press, 2014, p. 417-437, spec. p. 424.

(33) J.H. Katz et A. Ivey, *White awareness : the frontier of racism awareness training*, *The Personnel and Guidance Journal*, 55, 1997, p. 485 ; J.-J. Bonniol et M. Vial, *Les modèles de l'évaluation : textes fondateurs avec commentaires*, Bruxelles, De Boeck, 2009.

(34) M. H.-Evans, *What's in a name : Penological and institutional connotations of probation officers'labelling in Europe*, *Eurovista*, 2(3), 2013, p. 121.

complexe, d'une part, parce que le travail social relève des droits de l'homme dus à toutes les populations³⁵ et, d'autre part, parce que c'est ainsi que l'on obtient l'engagement et la réceptivité des condamnés.

Il est exact que sur le chemin de la délinquance, la situation économique et sociale joue un rôle potentialisant. Par exemple, le chômage et l'absence de perspectives conduisent à ce que nombre de personnes se « perdent dans la transition » entre l'adolescence et l'âge adulte³⁶ et ralentissent leur désistance³⁷. Il a cependant été établi depuis des décennies que ce n'est pas la pauvreté en tant que telle qui cause directement la délinquance. Les recherches et méta-analyses en ce domaine sont si nombreuses que nous nous bornerons à renvoyer à Andrews et Bonta³⁸. En outre, on sait aussi depuis des décennies que se limiter au travail social ne réduit pas la délinquance³⁹ – sans compter qu'il existe une énorme variabilité dans ce que représente le travail social – et qu'en certains cas, il aggrave la récidive et aggrave aussi d'autres facteurs qualitatifs comme, justement, l'autonomie et l'accès à l'emploi, au logement, à la santé et à la santé mentale, etc.⁴⁰.

Ajoutons qu'il est bien difficile d'expliquer la délinquance routière, sexuelle ou familiale à l'aune de causes exclusivement sociales. Pour ne prendre que les violences faites aux femmes, il n'est pas possible de traiter par le travail social des cognitions violentes, des cognitions « entitled » et machistes et ou s'attaquer à la triade psychopathie-personnalité antisociale et personnalité borderline qui constitue une partie considérable des populations en cause⁴¹. Par ailleurs, l'emploi et le logement ne sont pas suffisants. Qui travaille dans le milieu du suivi et du traitement de la délinquance sait bien que la difficulté n'est pas seulement de trouver un emploi aux personnes placées sous main de justice (PPSMJ) ; il

est de les y fidéliser. Le parcours professionnel erratique, certes en partie souvent causé par un niveau d'éducation insuffisant, est aussi le résultat des vies difficiles, voire chaotiques des PPSMJ, et de leurs cognitions, attitudes et fréquentations.

Enfin, prétendre que les outils structurés ne prennent pas en compte les facteurs sociaux suggère de n'en avoir jamais manié aucun. Le logement, l'emploi, l'éducation, la famille, tout y est bien présent, du moins pour les outils généralistes ; les outils spécialisés relatifs à la délinquance sexuelle et violente, par exemple, reposent sur des items nécessairement différents.

■ « De toute façon les scientifiques ne sont pas d'accord entre eux »

Cet argument surprenant revient finalement à dire que puisqu'il existe, inévitablement, des divergences entre scientifiques, la solution logique serait de ne point s'appuyer sur la science. Pour le formuler à l'aide d'une nouvelle métaphore médicale, c'est comme affirmer que « les oncologues ne sont pas d'accord sur la manière d'administrer la chimiothérapie, sur les dosages, etc. » et d'en conclure que « donc » il faut retourner à la saignée – à moins bien entendu qu'il ne soit finalement suggéré qu'il convient de ne pas prendre de décision et, pour le cancer, de s'abstenir de soigner. Au demeurant, de quels scientifiques en désaccord parle-t-on ici ? Seuls méritent cette appellation, au regard de la question traitée, ceux qui produisent les essais randomisés, les méta-analyses, contribuent à l'élaboration de modèles, sur la base de constructions théoriques valides à la fois sur le plan interne et externe et qui émergent sur ce terrain, ainsi que ceux qui proposent des modèles de traitement. Si l'on s'en tient à cette définition, il est exact qu'ils divergent sur des questions comme le type d'évaluation [par ex. forensique et structuré pour le *Good lives model* [GLM] ou actuariel ou structuré pour le modèle Risque-Besoins-Réceptivité [RBR]] ou encore la nature ou les objectifs du traitement (forensique et collaboratif pour le GLM et psycho-social pour le RBR). En revanche, ils sont très largement d'accord sur l'essentiel et notamment sur la nécessité de structurer le traitement et d'évaluer en amont⁴². Pour ne citer que lui, le GLM, longtemps ennemi du RBR et réciproquement, intègre en réalité désormais résolument ses principes⁴³. Nous concédons toutefois que la traduction canadienne de l'expression internationale *Evidence-Based Practices* (EBP) en « données probantes » est fort peu heureuse, qui suggère que la connaissance est définitive et indépassable. La réalité est naturellement que la science évolue constamment. Une meilleure traduction de cette notion en français serait peut-être « conforme ou fondé sur les données acquises de la science ».

■ « Tout est dans tout et inversement »

Des confusions sont encore faites entre le jugement professionnel structuré et les outils actuariels.

Les outils JPS sont davantage tolérés au motif qu'ils laisseraient plus de place à l'interprétation. C'est oublier qu'il existe derrière la dichotomie JPS/outils actuariels une question cruciale de formation, de compétences et d'expérience professionnelle. Alors que les outils actuariels sont utilisables par des agents de probation n'ayant pas de formation ou d'expérience clinique approfondie, les JPS reposent largement sur l'expertise clinique et psychologique, laquelle nécessite au surplus un solide bagage en psychologie forensique⁴⁴, ce que ni les CPIP ni même les tous les psychologues ne possèdent⁴⁵. En outre, de nombreux JPS sont à usage spécialisé (violence, violence domestique, délinquance sexuelle...), là où les outils actuariels sont souvent à usage généraliste. La question de

(35) En ce sens : M. et T. Connolly, *Morals, rights and practice in the human services. Effective, and fair decision-making in health, social care, and criminal justice*, London, Jessica Kingsley, 2008.

(36) F. Lösel, A. Bottoms et D. Farrington, *Young adult offenders. Lost in transition ?*, Willan, Cambridge criminal series, 2013.

(37) S. Farrall, B. Hunter, G. Sharpe et A. Calverley, *Criminal careers in transition. The social context of desistance from crime*, Oxford, Oxford University Press, 2014.

(38) Réf. préc.

(39) R. Martinson, *What Works ? Questions and Answers About Prison Reform. The Public Interest*, 35, 1974, p. 22 ; M.W. Klein, *Street Gangs and Street Workers*. Englewood Cliffs, NJ, Prentice Hall, 1971 ; E.M. Schur, *Radical Nonintervention : Rethinking the Delinquency Problem*. Englewood Cliffs, NJ, Prentice Hall, 1973 ; J. Fisher, *Is casework effective ? A review. Social work*, 18, 1973, p. 5.

(40) J. McCord, *A Thirty-Year Follow-Up of Treatment Effects. American Psychologist*, 33, 1978, p. 284 ; J. McCord, *Cures that Harm : Unanticipated Outcomes of Crime Prevention Programs. Annals of the American Academy of Political and Social Science*, 587, 2003, p. 16.

(41) A. Holtzworth-Munroe et G. Stuart, *Typologies of male batterers : Three subtypes and the differences among them. Psychological Bulletin*, 116(3), 1994, p. 476 ; M.P. Johnson, *Conflict and control : Gender symmetry and asymmetry in domestic violence. Violence Against Women*, 12(11), 2006, p. 1003.

(42) M. H.-Evans (à paraître), *RNR, Evidence Diversity and Integrative Theory*, in P. Ugwuodike, P. Raynor et J. Annison (eds.), *International Research on Supporting Rehabilitation and Desistance*, Bristol, Policy Press, 2006.

(43) V. par ex., P.M. Yates, D. Prescott et T. Ward, préc. note 7.

(44) T. A. Gannon et T. Ward, *Where has all the psychology gone ? A critical review of evidence-based psychological practice in correctional settings. Aggression and Violent Behavior*, 19, 2014, p. 435.

(45) C. Meyer, M. Borch-Jacobsen, J. Cottraux, D. Pleux et J. Van Rillaer (dir.), *Le livre noir de la psychanalyse. Vivre, penser et aller mieux sans Freud*, Paris, les Éditions des Arènes.

savoir lesquels, à public égal, sont les plus performants est particulièrement complexe, certaines recherches concluant à la supériorité de l'un, ou de l'autre, ou à une valeur similaire⁴⁶.

Des confusions sont faites, en second lieu, entre les outils statiques et dynamiques. Aux seconds sont attribuées les limites évidentes des premiers, lesquels ne sont pour l'heure capables que de répartir les populations en fonction de l'intensité du suivi, respectant certes en cela le principe « R » (pour « Risque » dans « Risque-besoins-réceptivité »), mais sans pouvoir ensuite orienter le traitement. En jouant sur les mots, on glisse rapidement vers la critique déjà évoquée de la « société du risque ». Hélas, n'est proposée comme garde-fou aux dangers bien réels de l'évaluation évoqués *infra*, que l'utilisation de tels outils

à la discrétion des professionnels, ce qui nous ramène à nouveau à la « gouvernance » des heuristiques. Il n'est pas inexact de souligner que l'administration pénitentiaire (AP) ne mesure sans doute pas tous les enjeux éthiques de ces outils ; toutefois, il ne faut pas se fourvoyer quant à la nature des problèmes éthiques en cause.

Si l'on ne peut que louer la volonté de l'AP de diffuser au maximum l'innovation criminologique, la recherche montre hélas que la méthode industrielle n'assure pas une diffusion optimale.

■ Un chiffon rouge ?

Toute forme d'évaluation, structurée ou non, pose de véritables problèmes⁴⁷. Songeons, par exemple, au cadre institutionnel dans lequel s'inscrit l'évaluation, qui est un contexte pénitentiaire monopolistique et centralisé d'État. Nous nous étonnons ici alors de ne pas voir mobilisées la sociologie des organisations, pas plus que les théories relatives au corporatisme ou à la bureaucratie.

Le focus sur le chiffon rouge de l'anti-empirisme cache une autre forêt institutionnelle qui conduit à la mise en œuvre industrielle et nationale d'une forme imparfaite de RBR (même si l'on a réalisé d'importants progrès depuis les programmes de prévention de la récidive [PPR]⁴⁸) alors que l'on sait, avec le malheureux exemple anglais à l'époque institutionnellement similaire, qu'une telle méthode de diffusion conduit à l'échec des programmes⁴⁹. Si l'on ne peut que louer la volonté de l'AP de diffuser au maximum l'innovation criminologique, la recherche montre hélas que la méthode industrielle n'assure pas une diffusion optimale⁵⁰.

Une longue série d'autres difficultés éthiques est posée par l'évaluation, touchant à la fois à l'autodétermination des personnes ainsi qu'aux protections procédurales nombreuses qui devraient accompagner toute forme d'évaluation, quelle qu'elle soit et notamment : principe du contradictoire ; droits de la défense ; discussion des preuves ; objectivisation des items ; droit à la contre-évaluation, etc⁵¹. La collaboration interinstitutionnelle⁵² permet encore de réduire les conséquences *nocebo* des facteurs institutionnels évoqués *supra* et d'augmenter l'efficacité. Là encore, aucune analyse critique n'est faite par ces *social scientists* sur le contexte institutionnel qui, en France, limite cette collaboration⁵³.

L'évaluation standardisée du risque n'est pas une finalité en soi, mais bien une étape nécessaire et essentielle afin d'adapter le traitement (fréquence et intensité) en fonction de chaque PPSMJ⁵⁴. Enfin, il est certain que les outils d'évaluation ne peuvent pas tout. Pour les dossiers les plus inquiétants et complexes, une analyse forensique structurée s'impose ; de même que l'évaluation des dimensions psychopathologiques (not. axes I – psychoses – et II – troubles de la personnalité) conforme aux données de la science, voire pour d'autres, une mesure de la sévérité de l'addiction.

La France a donc de nombreux obstacles à surmonter ; entretenir le passéisme pour des motifs idéologiques nous maintient dans un retard coupable sur la scène internationale, qu'elle soit académique ou professionnelle. Elle n'est, en outre, pas digne des besoins complexes des populations délinquantes comme des victimes que les opposants à l'empirisme oublient trop souvent. Quant aux académiques, l'urgence est qu'ils deviennent à la fois les évaluateurs authentiquement scientifiques des politiques et pratiques professionnelles dont la France a tant besoin, ainsi que des forces de proposition.

(46) J. Bonta et D.A. Andrews, *The Psychology of Criminal Conduct*, préc. note 7 ; R.K. Otto et K.S. Douglas (dir.), *Handbook of Violence Risk Assessment*, Abingdon, Routledge, 2010.

(47) Nous laisserons ici de côté la question de l'objet même de ces évaluations, déjà traitée dans M. H.-Evans, 2012, art. préc. note 1 ; ainsi que la question, à conséquences éthiques incontestables de la compétence clinique des évaluateurs praticiens déjà mentionnée *supra*.

(48) O. Vanderstukken et M. Benbouriche, Principes de prévention de la récidive et principe de réalité en France : les programmes de prévention de la récidive à la lumière du modèle « Risque-Besoins-Réceptivité », *AJ pénal* 2014. 522.

(49) J. Can, L. Falshaw, F. Nugent et C. Friendship, *Understanding what works : Accredited cognitive skills programmes for adult men and young offenders*, Home Office Research Findings, 226, London, Home Office, 2003 ; et récemment, mais sur la même période, à propos de la délinquance sexuelle : A. Mews, L. Di Bella et M. Purver, *Impact evaluation of the prison-based Core Sex offender Treatment Programme*, Ministry of Justice, Analytical Series, 2017.

(50) E. Rogers, *Diffusion of innovations*, New York, NY, The Free Press, 5^e éd., 2003.

(51) M. H.-Evans, *Law as an Extrinsic Responsibility Factor : What's Just is What Works !*, *European Journal of Probation*, 8(3), 2016, p. 146.

(52) M. H.-Evans, « All hands on deck » (re)mettre le travail en partenariat au centre de la probation, *AJ pénal* 2013. 139 ; et plus récemment : A. Giudicelli, Le champ de la prévention de la récidive dans sa dimension multi-partenariale et interinstitutionnelle en Charente-Maritime : étude locale et perspectives comparatistes, Rapp. 11.23, Université de la Rochelle, 2014 ; GIP Mission de recherche droit et justice.

(53) Pour un ex. inverse : G. Graham, *Rehabilitation work. Supporting desistance and recovery*, 2017, préc. note 3.

(54) M. Benbouriche, O. Vanderstukken et J.-P. Guay, Les principes d'une prévention de la récidive efficace : Le modèle Risque-Besoins-Réceptivité, *Pratiques psychologiques*, 21(3), 2015, p. 219.