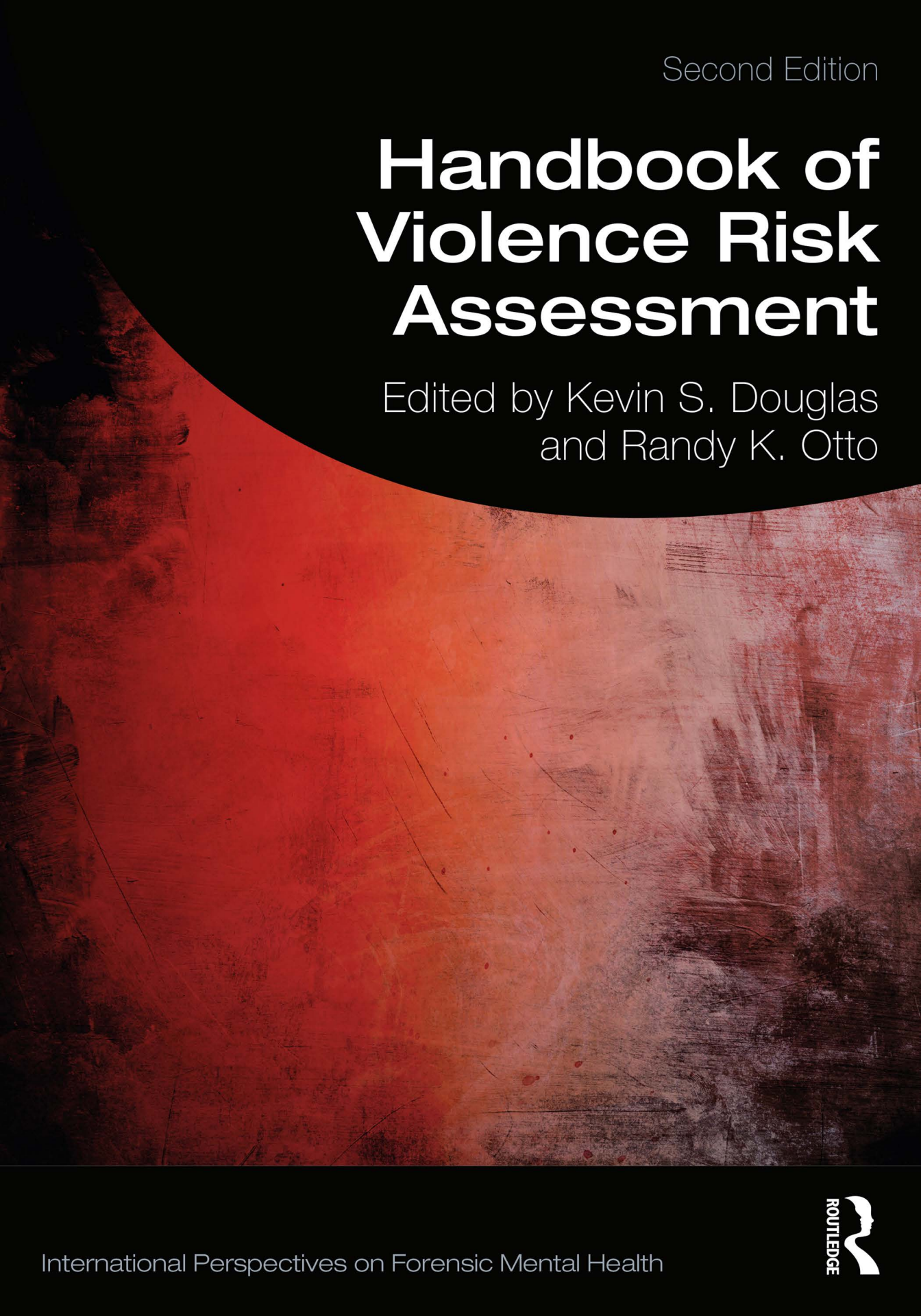
Second Edition
Handbook of Violence Risk Assessment
Edited by Kevin S. Douglas and Randy K. Otto
International Perspectives on Forensic Mental Health
ROUTLEDGE

Handbook of Violence Risk Assessment

The *Handbook of Violence Risk Assessment, Second Edition*, builds on the first edition's comprehensive discussion of violence risk assessment instruments with an update of research on established tools and the addition of new chapters devoted to recently developed risk assessment tools.

Featuring chapters written by the instrument developers themselves, this handbook reviews the most frequently used violence risk assessment instruments—both actuarial and structured professional judgment—that professionals use to inform and structure their judgments about violence risk. Also included are broader chapters that address matters such as the consideration of psychopathy and how the law shapes violence risk assessment.

Already the primary reference for practitioners, researchers, and legal professionals in this area, this second edition's easy-to-access, comprehensive, and current information will make it an indispensable reference for those in the field.

Kevin S. Douglas, LLB, PhD, is Professor of Clinical-Forensic Psychology at Simon Fraser University; Researcher at Helse Bergen HF Competence Centre in Forensic Psychiatry; and Senior Research Advisor at the Oslo University Hospital Competence Centre in Forensic Psychiatry.

Randy K. Otto, PhD, ABPP, has been a faculty member at the University of South Florida since 1989. His primary appointment is in the Department of Mental Health Law and Policy, and he has adjunct appointments in the Departments of Psychology and Criminology.

International Perspectives on Forensic Mental Health
A Routledge Book Series

Edited by Patricia Zapf
Palo Alto University

The goal of this series is to improve the quality of health care services in forensic and correctional settings by providing a forum for discussing issues and disseminating resources related to policy, administration, clinical practice, and research. The series addresses topics such as mental health law; the organization and administration of forensic and/or correctional services for persons with mental disorders; the development, implementation and evaluation of treatment programs and interventions for individuals in civil and criminal justice settings; the assessment and management of violence risk, including risk for sexual violence and family violence; and staff selection, training, and development in forensic and/or correctional systems. The book series will consider proposals for both monographs and edited works on these and related topics, with special consideration given to proposals that promote evidence-based best practices and that are relevant to international audiences. Workbooks and manuals targeted toward practitioners and reflecting evidence-based practice and intervention will also be considered.

Published Titles:

Handbook of Violence Risk Assessment, Second Edition
Edited by Kevin S. Douglas and Randy K. Otto

Sexual Predators: Society, Risk and the Law
Robert A. Prentky, Howard E. Barbaree & Eric S. Janus

Learning Forensic Assessment, Second Edition
Edited by Rebecca Jackson and Ronald Roesch

Evaluating Juvenile Transfer and Disposition Law, Science, and Practice
Kirk Heilbrun, David DeMatteo, Christopher King, Sarah Filone

Handbook of Forensic Mental Health Services
Ronald Roesch, Alana N. Cook

A Treatment Manual for Justice Involved Persons with Mental Illness Changing Lives and Changing Outcomes
Robert D. Morgan, Daryl Kroner, Jeremy F. Mills

Forthcoming Titles:

Safeguarding Forensic Violence Risk Assessment A Review Across Western Nations
Michiel van der Wolf

Handbook of Violence Risk Assessment

Second Edition

Edited by Kevin S. Douglas and Randy K. Otto

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Second edition published 2021
by Routledge
52 Vanderbilt Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

First edition published by Routledge 2015

Library of Congress Cataloging-in-Publication Data
Names: Douglas, Kevin S., editor. | Otto, Randy K., editor.
Title: Handbook of violence risk assessment / edited by Kevin S. Douglas and Randy K. Otto.
Description: Second edition. | New York, NY : Routledge, 2021. | Series: International perspectives on forensic mental health | Includes bibliographical references and index.
Identifiers: LCCN 2020035958 (print) | LCCN 2020035959 (ebook) | ISBN 9781138698703 (hbk) | ISBN 9781138698697 (pbk) | ISBN 9781315518374 (ebk)
Subjects: LCSH: Violence. | Violence—Forecasting. | Risk assessment. | Forensic psychiatry.
Classification: LCC HM1116 .H363 2021 (print) | LCC HM1116 (ebook) | DDC 303.6—dc23
LC record available at https://lccn.loc.gov/2020035958
LC ebook record available at https://lccn.loc.gov/2020035959

ISBN: 978-1-138-69870-3 (hbk)
ISBN: 978-1-138-69869-7 (pbk)
ISBN: 978-1-315-51837-4 (ebk)

Typeset in Baskerville
by Apex CoVantage, LLC

Contents

List of Contributors viii
Foreword xiv
CHRISTOPHER D. WEBSTER
Introduction and Overview xvii
KEVIN S. DOUGLAS AND RANDY K. OTTO

PART I
Core Themes and Critical Analysis of the Field 1

1 Approaches to Violence Risk Assessment: Overview, Critical Analysis, and Future Directions 3
KIRK HEILBRUN, KENTO YASUHARA, SANJAY SHAH, AND BENJAMIN LOCKLAIR

2 The Use of Measures of Psychopathy in Violence Risk Assessment 28
DAVID DEMATTEO, EMILY HANEY-CARON, ELYSE MOWLE, AND JOHN F. EDENS

3 Brief and Emerging Violence Risk Assessment Measures 52
CATHERINE S. SHAFFER, ERIN K. FULLER, AND LAURA S. GUY

4 Constitutional and Evidentiary Issues Concerning Risk Assessment 70
CHRISTOPHER SLOBOGIN

PART II
Empirical-Actuarial Measures 91

5 The Classification of Violence Risk (COVR) 93
JOHN MONAHAN

6 Static-99R: An Empirical-Actuarial Risk Tool for Adult Males With a History of Sexual Offending 106
R. KARL HANSON AND DANA ANDERSON

7 Actuarial Guides for Appraising the Risk of Violent Reoffending Among General Offenders, Sex Offenders, and Domestic Assaulters 131
N. ZOE HILTON, MARNIE E. RICE, GRANT T. HARRIS, BRIAN JUDD, AND VERNON L. QUINSEY

PART III
Rational-Actuarial Measures 157

8 Risk/Need Assessment for Adults and Older Adolescents: The Level of Service (LS) Instruments 159
J. STEPHEN WORMITH AND JAMES BONTA

9 The Youth Level of Service/Case Management Inventory 191
ROBERT D. HOGE

10 Two Treatment and Change-Oriented Risk Assessment Tools: The Violence Risk Scale (VRS) and Violence Risk Scale—Sexual Offense Version (VRS-SO) 206
MARK E. OLVER AND STEPHEN C. P. WONG

PART IV
Structured Professional Judgment Measures 225

11 The Early Assessment Risk Lists for Boys (EARL-20B) and Girls (EARL-21G) 227
LEENA K. AUGIMERI, MARGARET WALSH, PIA ENEBRINK, DEPENG JIANG, ANDREA BLACKMAN, AND ARETI SMARAGDI

12 The Science of and Practice with the HCR-20 V3 (Historical-Clinical-Risk Management-20, Version 3) 253
KEVIN S. DOUGLAS AND CATHERINE S. SHAFFER

13 The Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II) 294
ROBERT A. PRENTKY AND SUE RIGHTHAND

14 Structured Professional Judgment Guidelines for Sexual Violence Risk Assessment: The Sexual Violence Risk-20 (SVR-20) Versions 1 and 2 and Risk for Sexual Violence Protocol (RSVP) 322
STEPHEN D. HART AND DOUGLAS P. BOER

15 Short-Term Assessment of Risk and Treatability (START): Rationale, Application, and Empirical Overview 359
TONIA L. NICHOLLS, KAREN PETERSEN, MADISON ALMOND, AND CAMERON GEDDES

16 The Spousal Assault Risk Assessment Guide (SARA) 389
P. RANDALL KROPP AND ANDREA GIBAS

17 The Structured Assessment of Protective Factors for Violence Risk (SAPROF) 410
MICHIEL DE VRIES ROBBÉ, VIVIENNE DE VOGEL, AND AGNES VELDHUIZEN

18 The Structured Assessment of Violence Risk in Youth (SAVRY) 438
RANDY BORUM, HENNY P.B. LODEWIJKS, PATRICK A. BARTEL, AND ADELLE E. FORTH

Index 462

Contributors

Madison Almond, BA (Hons)
Department of Psychology
Simon Fraser University
Burnaby, BC, Canada

Dana Anderson, PhD
Simcoe Psychology
Barrie, Ontario, Canada

Leena K. Augimeri, PhD
SNAP International Headquarters
Child Development Institute
Toronto, Canada

Patrick A. Bartel, PhD
Youth Forensic Psychiatric Services
Vancouver, British Columbia, Canada

Andrea Blackman, BA
SNAP International Headquarters
Child Development Institute
Toronto, Ontario, Canada

Douglas P. Boer, PhD
Centre for Applied Psychology
University of Canberra
Canberra, Australia

James Bonta, PhD
Consultant, Corrections and Criminal Behaviour
Ottawa, Ontario, Canada

Randy Borum, PsyD
School of Information
University of South Florida
Tampa, Florida, USA

David DeMatteo, JD, PhD
Department of Psychology & Thomas R. Kline School of Law
Drexel University
Philadelphia, Pennsylvania, USA

Kevin S. Douglas, LLB, PhD
Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada
&
Competence Centre in Forensic Psychiatry
Helse Bergen HF
Bergen, Norway
&
Competence Centre in Forensic Psychiatry
Oslo University Hospital
University of Oslo
Oslo, Norway

John F. Edens, PhD
Department of Psychological & Brain Sciences
Texas A&M University
College Station, Texas, USA

Pia Enebrink, PhD
Department of Clinical Neuroscience
Karolinska Institute
Stockholm, Sweden

Adelle E. Forth, PhD
Department of Psychology
Carleton University
Ottawa, Ontario, Canada

Erin K. Fuller, MA
Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada

Cameron Geddes, BSc
Psychiatry, University of British Columbia
BC Mental Health & Substance Use Services
Vancouver, British Columbia, Canada

Andrea Gibas, PhD
Ontario Shores Centre for Mental Health Sciences
Whitby, Ontario, Canada

Laura S. Guy, PhD
Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada

Emily Haney-Caron, JD, PhD
John Jay College of Criminal Justice and the Graduate Center
City University of New York
New York, New York, USA

R. Karl Hanson, PhD
Carleton University
Ottawa, Ontario, Canada

Grant T. Harris, PhD
Department of Psychology
Queen's University
Kingston, Ontario, Canada
&
Department of Psychiatry
University of Toronto
Toronto, Ontario, Canada

Stephen D. Hart, PhD
Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada
&
Faculty of Psychology
University of Bergen
Bergen, Norway

Kirk Heilbrun, PhD
Department of Psychology
Drexel University
Philadelphia, Pennsylvania, USA

N. Zoe Hilton, PhD
Department of Psychiatry,
University of Toronto
Toronto, Ontario, Canada
&
Waypoint Research Institute
Penetanguishene, Ontario, Canada

Robert D. Hoge, PhD, C.Psych
Department of Psychology
Carleton University
Ottawa, Ontario, Canada

Depeng Jiang, PhD
Department of Community Health Sciences
University of Manitoba
Winnipeg, Manitoba, Canada

Brian Judd, PhD
Private Practice
Olympia, Washington, USA

P. Randall Kropp, PhD
British Columbia Forensic Psychiatric Services Commission
Vancouver, British Columbia, Canada
&

Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada

Benjamin Locklair, JD, PhD
Philadelphia Department of Behavioral Health and Intellectual disAbility Services
Philadelphia, Pennsylvania, USA

Henny P. B. Lodewijks, PhD
Lodewijks Advies
Zutphen, The Netherlands

John Monahan, PhD
School of Law
University of Virginia
Charlottesville, Virginia, USA

Elyse Mowle, PhD
ICF, Inc.
Washington, District of Columbia, USA
&
Department of Psychological & Brain Sciences
Texas A&M University
College Station, Texas, USA

Tonia L. Nicholls, PhD
Psychiatry, University of British Columbia
BC Mental Health & Substance Use Services
Vancouver, British Columbia, Canada

Mark E. Olver, PhD
Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

Randy K. Otto, PhD
Department of Mental Health Law & Policy
University of South Florida
Tampa, Florida, USA

Karen Petersen, PhD
Psychiatry, University of British Columbia
Vancouver, British Columbia, Canada

Robert A. Prentky, PhD
Department of Psychology
Farleigh Dickinson University
Teaneck, New Jersey, USA

Vernon L. Quinsey, PhD
Departments of Psychology, Biology, and Psychiatry,
Queen's University
Kingston, Ontario, Canada

Marnie E. Rice, Ph.D.
Department of Psychiatry and Behavioural Neurosciences
McMaster University
Hamilton, Ontario, Canada
&
Department of Psychology
Queen's University
Kingston, Ontario, Canada
&
Department of Psychiatry
University of Toronto
Toronto, Ontario, Canada

Sue Righthand, PhD
Department of Psychology
University of Maine
Orono, Maine, USA

Catherine S. Shaffer, PhD
Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada

Sanjay Shah, JD, PhD
Atlas Psychological Services
Atlanta, Georgia, USA

Christopher Slobogin, JD, LLM
College of Law
Vanderbilt University
Nashville, Tennessee, USA

Areti Smaragdi, PhD
SNAP International Headquarters
Child Development Institute
Toronto, Ontario, Canada

Agnes Veldhuizen, MSc
Center for Forensic Outpatient Treatment
The Forensic Care Specialists
Utrecht, The Netherlands

Vivienne de Vogel, PhD
Center for Social Innovation
HU University of Applied Sciences
Utrecht, The Netherlands
&
Department of Research
The Forensic Care Specialists
Utrecht, The Netherlands

Michiel de Vries Robbé, PhD
Department of Child and Adolescent Psychiatry

Amsterdam UMC, VU University Medical Center
Amsterdam, The Netherlands
&
Department of Psychiatry and Behavioural Neurosciences
McMaster University
Hamilton, Ontario, Canada

Margaret Walsh, BA
SNAP International Headquarters
Child Development Institute
Toronto, Ontario, Canada

Christopher D. Webster, PhD
Department of Psychiatry
University of Toronto
Toronto, Ontario, Canada
&
Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada
&
Department of Child and Youth Care
University of Victoria
Victoria, British Columbia, Canada
&
Child Development Institute
Toronto, Ontario, Canada

Stephen C. P. Wong, PhD
Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
&
Centre for Forensic Behavioural Science
Swinburne University of Technology
Melbourne, Australia

J. Stephen Wormith, PhD
Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

Kento Yasuhara, PhD
Psychology Department
University of New Haven
New Haven, CT, USA

Foreword

Christopher D. Webster

As a member of the Ontario Review Board and the British Columbia Review Board, I read, each year, dozens of formal reports from mental health professionals in the various disciplines. Insofar as it is the Boards which make the actual decisions as to whether to detain, to release into the community under conditions, or to grant absolute discharge, we, as members, are heavily reliant on the risk assessments provided by our colleagues on the other side of the table. Oftentimes, these reports are first rate and we do not have to torture the clinicians too greatly during questioning. At other times the reports are sub-standard by any reasonable criterion. Even if the clinicians claim to have followed some recognized and published scheme, it seems as if they are not using the scheme to full advantage. For sure, too, at times, we are left wondering if the framework adopted is even the right one (e.g., an HCR-20 V3 [see Chapter 12] might have been used in the assessment of a 19-year-old when a START:AV [see Chapter 15] would have been more fitting and appropriate; a VRAG-R [see Chapter 7] score is used as the sole basis to underpin a recommendation on behalf of an Indigenous person). As well, while sitting on a Board—one held at a well-resourced hospital or community service—it is impossible not to wonder why no attempts have been made to find out if the various assessment schemes are "working" in the particular context. It sometimes seems as though these devices are "plucked off the shelf" and that users are content to assume the guides in question have been validated elsewhere. There can be an absence of realization that follow-along studies are required to find out how these devices are performing in the specific circumstances in which they are being employed. This is in spite of the fact that our colleagues in the Netherlands have long demonstrated the value of attaching, whenever possible, a researcher to each clinical team (de Vogel, van den Broek, & de Vries Robbé, 2014; de Vries Robbé, personal communication, August 3, 2020).

Our colleagues on the "front lines" deserve sympathy. Caseloads tend to be high. Beds are filled as soon as they are vacated. Every day our colleagues are hounded for reports by Courts and Boards. As well, patients have to be interviewed and treated. Yet in all of this, risk assessment and risk management stand out as key issues. It is not a matter of "packing in" the requisite information during a residency or internship in the vain expectation that this pool of knowledge will last for a career. This is very much an evolving field. We have learned a great deal over the course of the 43 years I have been involved. At that early time we had, aside from the work of Häfner and Böker (1973, 1982) in Germany, not much beyond the inspired findings of "Operation Baxstrom" (Steadman & Halfon, 1971; Steadman & Keveles, 1972). This "opportunistic" study taught us that as clinicians we are apt to over-predict violence on a statistical basis. We also had at that time little besides the thoughtful commentaries from the front lines by Saleem Shah (1975) and Peter Scott (1977). Then we got the 1981 book by John Monahan. This was an organizational masterpiece; he told us, in effect, that we needed to "smarten up" in terms of defining prediction and outcome measures. So there were two main emerging threads in the late 1970s/early 1980s—Clinical and Organizational. The third thread—devices—began to evolve in the

late 1980s/early 1990s. Since then we have seen remarkable developments in the third aspect—actuarial/statistical and Structured Professional Judgment schemes—which is, of course, the focus of this book.

In this way we have a three-component model: (1) a large array of assessment schemes which have been subject to varying degrees of scientific test; (2) a pool of internationally connected clinicians, many of whom are highly able educators; and (3) organizational support—specialized journals, commendable publishers, worldwide distributors, professional organizations, and even, thanks to the late Dr. Derek Eaves, the International Association of Forensic Mental Health Services. This third arm is key. When we started out with the HCR-20 V1 in 1995 (Webster, Eaves, Douglas, & Wintrup, 1995), we had no way to get it into the hands of would-be users. All we had was a part-time student in a basement room struggling to keep up with inquiries and orders. It took time to set up a distribution service and to arrange for the delivery of educational events.

The area of risk evaluation is in constant flux. How could it be otherwise? It is not as if these days we are assessing on the basis of the kinds of crimes most commonly carried out in 1975. We'd not even heard of the sorts of offences now committed via the internet. Drug abuse back then was not as prevalent as it is today. Diagnostic frameworks have been amended. As well, new realizations arise about how particular sub-populations have been discriminated against. Women are more apt than before to express dissatisfactions concerning their roles and opportunities in Western society. Attitudes toward law enforcement shift. One senior colleague remarked to me that any risk evaluation scheme that has been in existence and unmodified for more than ten years or so needs to be approached with suspicion.

Our editors have served, and will continue to serve, in all three aspects of the "triad." This book, like its forerunner, fits into the organizational/administrative/policy slot. Bearing in mind the rate at which papers and even books on risk assessment and management appear, it is *essential* there be periodic *consolidations*. Busy administrators/policy wonks, clinicians, and researchers need to be able to reach for a single source, one which integrates information and points, as does Professor Heilbrun (see Chapter 1), to where we should now be headed. The field needs a text of this style every decade. Just as a submarine needs to surface periodically, so, too, does our field need to come up for air and a good look around.

By this stage I hope readers are clear that in referring to the "triad" I'm not harking back to fire-setting, cruelty to animals, and bed-wetting (cf. the still much-cited Hellman & Blackman, 1966). Recently I saw the movie *Ford versus Ferrari*. The movie, starring Matt Damon, centers on the idea that to win the Le Mans, three things are needed: (1) the right car loaded with the latest engineering advances (cf., the properly chosen assessment device or devices); (2) a driver who not only knows how to turn the wheel and apply the brakes, but who comprehends the internal workings, capabilities, and limits of the machine (cf., a clinician or team of professionals possessing much-needed acumen); and (3) an organization to back the enterprise (i.e., to finance it, to get the research done, to arrange publicity, and so on). Our area is no different. No matter how "wonderful" the scheme might be, it is of little use if clinicians do not take the trouble to learn how to "drive" it as the originators intended. As well, to prove itself, the car and driver must be subject to the actualities of the "race" (i.e., submit themselves to the rigors of trials as set by journal editors, audiences at scientific and professional meetings, etc.).

Let me end with a couple of comments on the structure of the book. I counted the number of contributors to the various chapters. I discovered I know approximately half of them. What encourages me is that I don't know the other half. That's good. It means we have plenty of upcoming researchers, clinicians, and policy-inclined folks entering the fray. My second comment is based on my "research." As noted previously, Douglas/Otto books are needed every ten years or so. On a whim, I took the first three chapters to pop up on my screen. I then counted all references cited as being published in 2010 or earlier and all those published in 2011 or later. The

scores were 52/52, 83/40, and 18/16 (i.e., 153/108). To me, this is evidence that the contributors did what the editors asked them to do: namely, bring the book up to date. It also helps reinforce my point that this field is continually "on the march."

I wish the editors, the contributors, and the publisher all due success for the new edition. It is not a book that needs to be shelved; it needs to be read. My hope, too, is that that some evaluees, and perhaps their counsel, will benefit from gaining a close understanding of how the many formats are designed to work and how the ensuing results should be interpreted. The work we do has far-reaching implications for making decisions on behalf of civilly detained patients, forensic psychiatric patients, and persons held in prisons while suffering from mental and personality disorders. The book is a "call to arms."

References

de Vogel, V., van den Broek, E., & de Vries Robbé, M. (2014). The use of the HCR-20V3 in Dutch forensic psychiatric practice. *International Journal of Forensic Mental Health*, *13*, 109–121.

Häfner, H., & Böker, W. (1973). Mentally disordered violent offenders. *Social Psychiatry*, *8*, 220–229.

Häfner, H., & Böker, W. (1982). *Crimes of violence by mentally disordered offenders* (H. Marshall, trans.). Cambridge: Cambridge University Press (Original work published 1973).

Hellman, D. S., & Blackman, N. (1966). Enuresis, firesetting, and cruelty to animals: A triad predictive of adult crime. *American Journal of Psychiatry*, *122*, 1431–1435.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.

Scott, P. D. (1977). Assessing dangerousness in criminals. *British Journal of Psychiatry*, *131*, 127–142.

Shah, S. A. (1975). Dangerousness and civil commitment of the mentally ill: Some public policy considerations. *American Journal of Psychiatry*, *132*, 501–505.

Steadman, H. J., & Halfon, A. (1971). The Baxstrom patients: Backgrounds and outcomes. *Seminars in Psychiatry*, *3*, 376–386.

Steadman, H. J., & Keveles, G. (1972). The community adjustment and criminal activity of the Baxstrom patients: 1966–1970. *American Journal of Psychiatry*, *129*, 80–86.

Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Burnaby, BC: Mental Health, Law, and Policy Institute, and Forensic Psychiatric Services Commission of British Columbia.

Introduction and Overview

Kevin S. Douglas and Randy K. Otto

In the Introduction and Overview to the first edition of this book, we wrote that,

> [because risk] for violence directed towards others was relevant to a variety of legal (e.g., civil commitment, criminal sentencing, parole decisions, disposition in delinquency proceedings) and clinical matters (e.g., case management of and decision making regarding clients who pose a risk of harm to others), its accurate assessment is of concern to legal, correctional, and mental health professionals.

This remains true today.

In the 11 years since publication of the first edition of this book, how has the field changed? How has it remained the same? There have been important developments in the law that have shaped risk assessment (for instance, whether the entire risk assessment enterprise is fair to certain subgroups), and important conceptual developments in violence risk assessment (for instance, greater emphasis on risk formulation; scenario planning; integration with risk reduction, treatment, and management; and examinees' strengths and protective factors). New measures are available, as well. Other measures—long in use and well-established at the time of the 2010 edition—have either been revised substantially or have been subjected to much greater evaluation. Much more is known about the "actuarial versus structured professional judgment" (SPJ) debate, and on the performance and role of complementary instruments often used within violence risk assessment (for instance, measures of psychopathy).

Hence, a second edition of the *Handbook of Violence Risk Assessment* was warranted. As with the first edition, we sought—where possible—to have instrument authors describe their tools and the research examining their utility. Where this was not possible, we invited scholars who had a great deal of knowledge and experience using and evaluating the instruments to author these chapters. We have continued the focus on the most commonly used and researched violence risk assessment instruments.

The second edition does include some new additions. We expanded Part I with two new chapters, in addition to revisions of the original two chapters. As with the first edition, this second edition includes a critical overview of key themes, findings, and developments in the risk assessment field, including an evaluation of the empirical support for both actuarial and SPJ measures (Heilbrun, Yasuhara, Shah, & Locklair). It also contains a chapter on the appropriate use, and limitations, of measures of psychopathy within risk assessment (DeMatteo, Haney-Caron, Mowle, & Edens).

As previously mentioned, there are two new chapters in Part I—one of these reviews what we have called "brief and emerging measures" (Shaffer, Fuller, & Guy). This chapter includes a review of ten measures that are either intended as screening or triaging devices, with some empirical evaluation, or more comprehensive measures that, while showing promise, do not yet

have substantial empirical evaluation. The second addition, authored by Christopher Slobogin, addresses key legal issues and developments with respect to violence risk assessment.

As with the first edition, the remainder of the second edition focuses on risk assessment instruments. As we did in the first edition, we asked authors to follow a template: (a) an overview of the instrument(s); (b) a description of instrument development, and populations with which it can be used; (c) a discussion of the research regarding its reliability and validity; (d) a review of limitations and future directions; and (e) a case example that provides guidance to instrument users around issues of report-writing and risk communication.

The careful reader will realize that we changed the organizational structure for the second edition. Rather than grouping instruments according to the population with which they are used, we organized them based on the approach they employ. We grouped instruments according to whether they are actuarial or SPJ in nature, and further divided actuarial instruments into whether they used "empirical" item selection (that is, derived from one or more specific samples), or "rational" item selection (that is, derived from a broad canvassing of the empirical and professional literatures). Because all SPJ instruments use rational item selection, no such distinction was necessary. Instruments are otherwise presented in alphabetical order.

Part II, Empirical-Actuarial Measures, includes the following instruments: Classification of Violence Risk (COVR; Monahan); the Static-99 Revised and related measures such as the STABLE-2007 (Hanson & Anderson); and the Violence Risk Appraisal Guide (VRAG) and related instruments (its revision, the VRAG-Revised; the Sex Offender Risk Appraisal Guide, or SORAG; the Ontario Domestic Assault Risk Assessment, or ODARA; and the Domestic Violence Risk Appraisal Guide, or DVRAG), by Hilton, Rice, Harris, Judd, and Quinsey.

Part III, Rational-Actuarial Measures, includes the following instruments: the Level of Service (LS) instruments for adults (Wormith & Bonta) and youth (Hoge; the Youth Level of Service/Case Management Inventory, or YLS/CMI); and the Violence Risk Scale, or VRS, and its Sex Offender Version, VRS-SO (Olver & Wong).

In Part IV, the following SPJ instruments are reviewed: the Early Assessment Risk List measures, for Boys (EARL-20B) or Girls (EARL-21G), by Augimeri, Walsh, Enebrink, Jiang, Blackman, and Smaragdi; the Historical-Clinical-Risk Management-20 Version 3, with brief review of its predecessor, Version 2 (Douglas & Shaffer); the Juvenile Sex Offender Assessment Protocol-II, or J-SOAP-II (Prentky & Righthand);[1] the Sexual Violence Risk-20 (SVR-20) Versions 1 and 2 and Risk for Sexual Violence Protocol (RSVP), by Hart and Boer; the Short-Term Assessment of Risk and Treatability, or START (Nicholls, Petersen, Almond, & Geddes); the Spousal Assault Risk Assessment Guide, or SARA (Versions 1–3), by Kropp & Gibas; the Structured Assessment of Protective Factors, or SAPROF, and its derivatives (de Vries Robbé, de Vogel, & Veldhuizen); and the Structured Assessment of Violence Risk in Youth, or SAVRY (Borum, Lodewijks, Bartel, & Forth).

As such, 14 chapters focus on a single instrument or group of related instruments. Together, more than 25 instruments are covered in these chapters. Adding the chapter on measures of psychopathy and the chapter on brief and emerging measures, over 35 of the most commonly used and researched instruments for violence risk assessment are described and reviewed in this second edition.

Authors were free to represent their opinions on the appropriate uses and the degree of validation of their measures. Of course, as readers may surmise, some authors likely disagree with others on the relative support for the various measures. We invite readers to read the chapters critically and, with guidance from the chapter authored by Heilbrun and his colleagues, which considers the degree of support for various approaches to risk assessment, decide for themselves.

We are very grateful for the time, effort, and expertise put into all of the chapters by their respective authors and hope that our readers will be similarly appreciative.

Note

1. The Juvenile Sex Offender Assessment Protocol-II, or J-SOAP-II, is best described as neither actuarial nor SPJ, because it does not offer estimates about risk, either categorically or numerically. It does quantify the proportion of total possible risk factors present on its various scales, which is intended to guide treatment efforts, and ultimately inform a clinician's assessment of risk. In principle, methods could be developed to use this instrument in either an actuarial or SPJ format, but that would be inconsistent with its developers' intentions. We have categorized this instrument, for purposes of organization, along with the SPJ instruments.

Part I

Core Themes and Critical Analysis of the Field

Chapter 1

Approaches to Violence Risk Assessment

Overview, Critical Analysis, and Future Directions

Kirk Heilbrun, Kento Yasuhara, Sanjay Shah, and Benjamin Locklair

The problems presented by violence in contemporary society have been an important consideration for many decades. As the law has turned to the behavioral and medical sciences to improve the prospects for accurately appraising and managing the risk of violent behavior, the past thirty years have witnessed the development and validation of specialized tools for the prediction and management of certain kinds of serious violence and criminal offending. The second edition of this book offers a description of some of these tools, developed for specific purposes in legal contexts, and the relevant research associated with them.

Risk Assessment Considerations

This chapter provides a context in which the specific tools described in subsequent chapters can be considered. There are six important considerations applicable to risk assessment tools that will be addressed in the first section: context, purpose, population, parameters, approach, and applicability. Subsequent sections of the chapter focus on approaches to risk assessment, the major features of each approach, the strengths and weaknesses of each approach, and their respective states of validation. The present chapter does not provide an in-depth review of the validation data relevant to specific tools. That is addressed in the subsequent chapters describing the various risk assessment tools.

Context

The first consideration in conducting a violence risk assessment involves the broad context in which this appraisal is being conducted. Such context influences the nature of the decision to be made, who is responsible for this decision, and the consequences that can result from differing appraisals. There are four important (and different) contexts in which a risk assessment is likely to be conducted: legal, clinical, school/workplace, and threats to protectees. *Legal contexts* are those in which a legal (or quasi-legal) decision-maker must render a decision in the course of litigation or administrative tribunal. Such contexts may be in the domain of criminal or civil proceedings (see, e.g., Melton et al., 2018). Decisions in this context may involve initial civil commitment or criminal sentencing, release from incarceration or secure hospitalization, or steps that may be associated with a planned release (such as community notification or post-sentence civil commitment for sexual offenders).

Clinical contexts involve circumstances under which interventions are delivered to reduce the risk of violence or offending in the broader context of a treatment relationship. This context is sufficiently broad to include interventions in secure settings (e.g., jails, prisons, forensic hospitals) as well as those delivered using some leverage (e.g., to individuals on parole, probation, diversion from prosecution, or juvenile home-based placement) and others using no leverage at all (e.g.,

mental health treatment delivered on a voluntary basis to individuals in the community, in jurisdictions in which there is a settled *Tarasoff*-type obligation to warn or protect).

School/workplace contexts encompass threats of harm to others that are typically not yet a part of formal legal proceedings, in which both the seriousness of the threat and the nature of the indicated risk-reduction strategies are important parts of the overall appraisal. Assessment in these settings is typically conducted in response to the concerns raised by the actions of specific individuals such as students or employees, rather than in a broader process involving all individuals in the setting.

Finally, *threats to protectees* also involve the dual issues of threat seriousness and risk management. However, they differ because they are directed toward specific individuals (those under the protection of the Secret Service or federal marshals, for example), or the kind of broader targets involved in domestic or international terrorism. Included as well in this category are threats involving potential harm to a domestic partner, where there is typically a specified victim and the initial appraisal is made by police who are called to the scene.

Most of the chapters in this book are devoted to different specialized risk assessment tools that are almost entirely applicable in legal contexts. Accordingly, that will also be the focus of discussion in this chapter. It should be noted that a number of specialized measures referenced in this chapter and described further in this book would be appropriate for use in *clinical* and *threats to protectees* contexts, as these have just been described.

Purpose

One of the important influences on contemporary conceptions of risk assessment is the Risk-Need-Responsivity (RNR) model described by Canadian researchers (see Chapter 8, this volume; Andrews, Bonta, & Hoge, 1990; Andrews, Bonta, & Wormith, 2006; Bonta & Andrews, 2017). This involves the appraisal of three related domains. *Risk* refers to the probability that the examinee will engage in a certain kind of behavior in the future, typically either violence/violent offending, or criminal offending of any kind, with higher-risk individuals receiving more intensive intervention and management services under the risk principle. This kind of risk classification has most frequently been measured by using static risk factors, which do not change through planned intervention, although some tools (e.g., the LS measures; see Andrews & Bonta, 2001; Andrews, Bonta, & Wormith, 2004; Chapters 8 and 9; and all structured professional judgment tools, for example, Douglas, Hart, Webster, & Belfrage, 2013; Kropp & Hart, 2015; see also Hart, Douglas, & Guy, 2017; Chapters 11-18 in this volume) use both static risk factors and risk-relevant needs. *Needs* are variables describing deficits which are related to the probability of such targeted outcomes; they are composed of dynamic risk factors (called *criminogenic* needs in the RNR model) or protective factors that have the potential to change through such planned intervention. Targeting interventions to address criminogenic needs is the goal described by the need principle. *Responsivity* refers to the extent to which an individual is likely to respond to intervention(s) enacted to reduce the probability of the targeted outcome behavior.

A comparable distinction involving risk assessment in legal contexts has been made between prediction and risk management (Heilbrun, 1997), in which it was observed that some legal decisions are best informed by a prediction of whether the individual will reoffend or otherwise behave violently. Perhaps the most frequent example of such a decision involves civil commitment, in which the legal decision-maker must decide whether the likelihood that the individual will harm others (or self) is sufficient to justify involuntary hospitalization. There is a limited risk management component to this decision, although future decisions regarding such individuals (such as release from hospitalization) must consider risk management to a much greater extent. If

the risk is sufficiently high, the court will presumably grant the petition for civil commitment. If not, the petition is likely to be denied (or granted on grounds other than risk to harm others). But under neither circumstance would the court be particularly concerned with a specific approach to lowering the violence risk. In contrast to civil commitment, when the court *does* maintain jurisdiction over the individual following the decision—for example, when the defendant is diverted to mental health court or committed as Not Guilty by Reason of Insanity—then both level of risk and the nature of the risk-relevant needs can help inform the court's decision.[1]

Specialized risk assessment tools are designed to provide information that either is particular to the question of prediction or that addresses both risk and needs (with associated implications for intervention). This will be discussed in greater detail subsequently in this chapter. To our knowledge, there are no specialized tools that focus *only* on risk reduction (although a number of tools include risk reduction intervention-planning as a primary goal). However, there is a technique (anamnestic assessment) derived from applied behavior analysis that promotes the informed selection of risk-relevant needs based on the individual's history. Such an approach can also be used with risk-need tools, in a manner to be discussed later in this chapter.

Populations

One of the important considerations in risk assessment involves the population to which the individual being assessed actually belongs. There are important differences in base rates of violence, risk factors and protective factors,[2] and risk-relevant interventions for differing populations. In addition, specialized tools are typically derived and validated to apply to a single population (e.g., juvenile offenders) or related populations (adult offenders or insanity acquittees, sex offenders), but not across widely discrepant populations.

There are four considerations in delineating a population for risk assessment purposes. These are age, gender, behavioral health status, and location. *Age* typically refers to three distinct groups: preadolescent children, adolescents, and adults (and possibly older adults). *Gender* is important because a specialized tool may not be validated for females or may use different norms. The *behavioral health status* of a population refers to whether it is selected through assessment, intervention, or legal action as having a number of individuals with behavioral health problems. Since justice-involved and behavioral health are among the populations to which risk assessment is most often applied, it is particularly important to distinguish behavioral health populations without formal justice involvement at the relevant time (e.g., those who are civilly committed) from general offender populations in which behavioral health problems are not part of the selection criteria. This can become more complex when considering populations such as criminal defendants who have been adjudicated Not Guilty by Reason of Insanity which, by definition, include both behavioral health and offending among the selection criteria.

Finally, the variable of *location* refers to the setting from which the population is drawn. Offending populations can be drawn from the community (e.g., those on probation or parole) or from an institutional setting. School and work are examples of settings with populations that are typically not offenders. There can again be increased complexity when the risk appraisal involves assessment of individuals in one setting but considers their risk for violence or offending in another—such as when convicted persons or insanity acquittees are appraised for their risk of violence in the community following release from prison or the hospital, respectively. The consideration of these four variables—age, gender, behavioral health status, and setting—allows the identification of a population with sufficient specificity to determine whether a given risk assessment tool should be used. This choice is straightforward when the examinee is part of a specific population for which a particular tool was developed (e.g., use of a sexual violence risk tool with an individual convicted of a sex offense).

Parameters

The next important consideration in risk assessment involves what is being predicted, with what frequency of outcome, at what probability or category of risk, in what setting, over what period of time, and the nature of the risk factors and protective factors involved in the appraisal. It is important to specify, in the beginning, what outcome(s) are of concern. The broader the class of outcomes, the higher the base rate will be. The impact of base rates is discussed elsewhere in this chapter. It is also necessary to consider how this outcome will be measured. Both researchers and practitioners have employed approaches to appraising violent behavior or violent offending that include (1) self-report, (2) the report of collateral observers, and (3) official records (e.g., rearrest, rehospitalization). In the context of non-criminal community violence in the United States, the most sensitive of these measures is self-report, with collateral observation second and official records a distant third (Monahan et al., 2001; Steadman et al., 1998). But this may vary widely across jurisdictions, countries, and contexts, depending on whether the individual has a strong incentive to deny or minimize self-reported impairments or intentions, whether s/he is seen regularly by a collateral observer, and the extent to which official records are likely to reflect the occurrence of such behavior.

Next, we consider the specific parameters of the behavior itself. Are we targeting serious violence only, or is more minor aggression considered as well—or only a specific type of violence (e.g., sexual reoffending)? Are verbal threats included? What about behavior (such as arson) that may be directed at property but has the potential to harm other persons? This is a crucial question for researchers, but it is also important for practitioners to be as specific as possible. It is also important to indicate whether the appraisal would consider a single act of the behavior as a "yes," or whether the specified outcome encompasses possible multiple acts of the behavior (e.g., as part of stalking or child abuse). In addition, the appraisal must designate an outcome period of time. For some risk assessment, the relevant outcome period might be quite short—perhaps no longer than 24 to 48 hours, for instance, when an individual is being considered for involuntary hospitalization. Intermediate outcome periods often used in community and correctional outcome research in the last decade tend to range from 6 to 12 months. Much longer periods, up to 5 to 10 years, have been used in the development of some specialized tools and were more typical of violence research in the 1980s and 1990s. These research-employed outcome periods may be quite applicable to a given case in practice—or poorly applicable. This should be clarified whenever possible.

In both research and practice, location is a particularly important aspect of context in risk assessment. There is a substantial difference between the appraisal of violence risk in an institutional setting such as a juvenile placement, prison, or hospital versus risk in the community if and when the examinee is living in the community. The degree of structure associated with the former means greater monitoring by staff, clearer expectations and less tolerance for deviation from these expectations, lack of access to drugs and alcohol (presumably), and the provision in some instances of interventions to lower violence risk. There is also a much higher probability of detection and imposition of consequences in response to violent behavior that is displayed in institutional settings. In the community, in contrast, there is a wider range of circumstances encountered in daily living. An individual may be at work or in school, which itself may be the specific context in which violence risk is assessed. The individual may be living in a setting with others who may be targets for violence, particularly when domestic tensions are exacerbated by substance use and weapons are readily accessible. But location is a very important parameter in risk assessment; while it contributes to the situational influences on violent behavior, it has been investigated less extensively than have personal variables (Silver, 2001; Steadman, 1982). Fortunately, that is changing. Some investigators have pursued the measurement of situational

influences on violence in hospitals (Ogloff & Dafferns, 2006; Welsh, Bader, & Evans, 2013) and prisons (Abbiati, Palix, Gasser, & Moulin, 2019; Cooke, Wozniak, & Johnstone, 2008) through the development and review of specialized tools that may prove useful when validated and may influence the development of comparable tools for measuring community violence risk.

The next parameter involves the manner in which risk level is to be communicated. As discussed in more detail later in this chapter, there are currently differences in the structure of specialized tools that center on whether a risk assessment should provide a quantitative estimate of the probability of violence, or a categorical appraisal that seeks to distinguish lower- from higher-risk individuals without using a specific number. However, it is fair to say that one of the shared purposes of contemporary risk assessment tools involves providing some appraisal of relative risk. By contrast, a process can also be employed in risk appraisal that seeks only to identify risk factors and protective factors applicable to the individual being assessed. This process, sometimes termed "anamnestic" assessment (meaning recollection or account of one's history; Melton et al., 2018; Otto, 2000), will also be discussed in this chapter. Risk communication is a particularly important consideration in practice, as the entire risk assessment process may be misunderstood or discounted if risk communication is ineffective.

Finally, the important parameters in risk assessment include the identification of risk factors (associated with increased risk of violent outcome), promotive factors (associated with increased likelihood of a favorable outcome, such as responsible behavior), and protective factors (associated in a decrease in risk when risk factors are present). Specialized risk assessment tools have selected the risk and protective factors that are most relevant and empirically related to outcome for their populations of use. They vary in how they select these factors; some do so from a specific dataset or meta-analysis, while others review the broader empirical literature to identify important, recurring influences. This has been an area addressed by researchers on a group level—but it must be considered by practitioners seeking to include individualized considerations into their risk assessment in every evaluation they conduct.

Approach

For the purposes of this chapter, three approaches to risk assessment are described: actuarial,[3] structured professional judgment (SPJ), and anamnestic. Actuarial assessment is "a formal method" that "uses an equation, a formula, a graph, or an actuarial table to arrive at a probability, or expected value, of some outcome" (Grove & Meehl, 1996, p. 294). It uses predictor variables that can be quantified, either through classification (e.g., gender) or rating with a high degree of reliability. The predictors, and the weights assigned to them (if applicable), are validated against the outcome that is being predicted through research. Both the risk factors and, if applicable, their weights, are typically derived empirically as well (although neither empirical item selection nor weighting is a required element of the actuarial approach), from a single dataset or larger meta-analysis. So it is the quantified, mechanistic, reproducible combination of predictive factors, validated through empirical research against known outcomes and yielding a total score, that is the *sine qua non* of actuarial assessment.

SPJ involves the presentation of specified risk (and protective) factors, which are usually derived from a broad review of the literature rather than from analysis of a specific dataset. Risk factors are well-operationalized so, as with the actuarial approach, their applicability can be coded reliably (usually as no, possibly, or yes). Evaluators complete an SPJ tool by rating all the specified factors after examining the subject of the assessment, perhaps interviewing collateral informants, and reviewing records and other relevant data sources. Then, when rating overall risk, evaluators consider the presence and individual relevance of the identified risk (and protective) factors, develop a formulation of risk (why the individual acted in a violent manner), specify

future scenarios of concern (what the individual will do again that worries the evaluator), and the anticipated nature and intensity of management, treatment, or supervision needs in reaching an overall conclusion about risk. This is accomplished by having the evaluator draw this conclusion, however, rather than having the rated risk factors combined in a predetermined fashion using an established formula, as is done in actuarial assessment. This procedure makes it more feasible to use risk factors that are dynamic (potentially changeable through planned intervention) as contrasted with the static (unchanging through planned intervention), largely historical risk factors that constitute the large majority of the risk factors that make up actuarial instruments.

Actuarial and SPJ risk assessment approaches are similar in several respects. Both employ variables that have been selected for their demonstrated relationship to the outcome of violence (although actuarial measures typically rely upon variables drawn from a single dataset, while SPJ approaches have tended to use variables for which there appears to be broad empirical support in the literature). Both specify the variables to be considered, and both require that these variables be sufficiently well-operationalized so they can be coded or rated reliably.

These approaches differ in some important respects as well. Actuarial approaches combine their risk factors into a score, which should be cross-validated against known outcomes in the validation phase of developing an actuarial risk assessment tool. SPJ approaches do not combine risk factors numerically in a predetermined or mechanistic fashion, instead providing structured guidance to allow the user to make the judgment as to whether the individual is (for example) high, moderate, or low risk. The respective accuracy of both approaches is discussed later in this chapter.

The third risk assessment approach might better be characterized as a process than a specialized tool.[4] Anamnestic assessment (Melton et al., 2018) uses applied behavior analytic strategies to gather detailed information about the individual's history of violence. For each prior violent event, the individual is questioned in detail concerning the preceding and subsequent thoughts, feelings, and behaviors; the act itself; those involved; and other relevant details (e.g., whether drugs or alcohol were ingested, by whom, and at what level; whether weapons were involved, and their source; how victim[s] were targeted; where and when the event occurred). The goal of this process is not to derive an estimate of violence risk. Rather, the evaluator can use this procedure to identify risk factors and protective factors that recur across violent acts. Such information can help make more accurate ratings on such factors if a formal risk assessment tool is used. It can also facilitate identification of risk-relevant intervention targets—dynamic risk factors that are present during different violent acts that may, if improved through targeted intervention, reduce that individual's risk of future violence.

The use of unstructured clinical judgment will not be discussed for several reasons. First, more than 50 years of research has compared the accuracy of unstructured clinical judgment with actuarial approaches, with a consistent, modest advantage in predictive accuracy observed for the latter (see meta-analyses conducted by Ægosdóttir et al., 2006; Fazel, Singh, Doll, & Grann, 2012; Grove, Zald, Lebow, Snitz, & Nelson, 2000). Several meta-analyses are relevant. In the area of violence prediction, Mossman's (1994) meta-analysis indicated that clinical predictions were more accurate than chance (AUC = .67) but less accurate than the mean AUC for all studies (.78), for cross-validated discriminant function predictions (.71), or for behavior-based predictions (.78). Second, it seems clear that much of the progress that has been made during the last two decades in developing risk assessment approaches that are more accurate and scientifically supported has resulted in increasing their structure, both in the selection of risk factors and the way in which they are rated and combined. Unstructured clinical judgment by itself is no longer a useful or necessary approach to appraising violence risk. (When there are no applicable structured measures available, evaluators should at least "structure" their risk appraisals by focusing on known risk factors for the target outcome, derived from the scientific literature, the individual's history, or

both.) However, in the context of a risk assessment conducted as part of broader forensic assessment, clinical judgment does facilitate important aspects of data gathering such as interviewing of the individual and third parties, identifying and reviewing relevant records, interpreting all data, and reasoning toward conclusions (Heilbrun, 2009).

Another meta-analysis (Campbell, French, & Gendreau, 2009) incorporated 88 studies conducted between 1980 and 2006, focusing on prediction of violence (mostly nonsexual) in adults. The mean effect sizes of actuarial and SPJ measures were comparable. Newer instruments, as well as those developed using criminological-related theories or research, showed larger effects than did measures developed using less-relevant theory.

More recently, Fazel et al. (2012) performed a systematic review and meta-analysis using tools commonly used to predict or structure judgments about violence, sexual offending, and general offending. They used 73 samples (24,847 participants from 13 countries, with a reoffending rate of 23.7% over a mean of 49.6 months). Risk assessment tools produced low to moderate positive predictive values (median 41%) when predicting violent offending, but had higher negative predictive values (ranging from 81% to 95%). Instruments designed to predict or structure judgments about violent offending performed better than those used in the context of sexual or general offending. Their conclusion was that such specialized measures were very accurate in identifying low-risk individuals, but substantially less accurate in identifying high-risk individuals.

Applicability

The value of a specialized risk assessment tool depends upon its applicability. This section has outlined several important considerations in judging the potential utility of a particular tool. To what extent are the nature of the evaluation and the attributes of the specialized tool congruent? Such congruence should be judged according to context, purpose, populations, and parameters. The information needed to make this decision should be included in the manual of the specialized tool—one reason why it is important to have such a manual when using a specialized tool in practice (American Education Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Heilbrun, Rogers, & Otto, 2002). As this book offers chapters describing specialized tools focusing on children, adolescents, and adults in juvenile, criminal, and civil contexts, it will be clear that these respective tools are congruent with the kind of risk assessment described in those chapters.

Approaches to Risk Assessment

In this section, we offer a more detailed discussion of the functions of each of the major approaches to risk assessment described previously—actuarial, structured professional judgment, and anamnestic. It is useful to distinguish two major goals in risk assessment: prediction/classification, and risk reduction (Heilbrun, 1997). Some legal decisions call for expert evaluations that are strongly focused on the question of prediction. Examples include civil commitment, end of sentence sexual offender commitment, and criminal sentencing.

In the case of civil commitment and end of sentence sexual offender commitment, the legal decision-maker is concerned with whether the risk of violent behavior or future sexual offending is sufficiently high to justify a decision restricting individual liberty on public safety grounds—but the court will not retain jurisdiction and make subsequent decisions based on whether this risk has been decreased. By contrast, a legal decision to involuntarily hospitalize a defendant found Not Guilty by Reason of Insanity may indeed involve the retention of such jurisdiction—so the court might be informed by accurate information concerning both risk level and risk reduction potential. Regarding criminal sentencing, many states apply risk assessment in order to make

individualized decisions regarding supervision and services. Some states (e.g., Virginia, Utah) apply risk assessment as a front-end approach in determining appropriate sentences. For instance, higher-risk offenders may receive longer sentences and lower-risk offenders shorter sentences when there is discretion on the part of the court (e.g., Monahan & Skeem, 2016).

Table 1.1 summarizes three important questions in risk assessment and the extent to which each of these questions is typically addressed by actuarial, SPJ, and anamnestic assessment. The strength of actuarial assessment, presuming that the tool is appropriately derived and cross-validated, is in facilitating a quantitative estimate of risk level. However, this is a double-edged sword. Because of the inextricably quantitative nature of actuarial risk assessment, actuarial tools lend themselves to research that calculates accuracy using statistical measures such as sensitivity, specificity, and Area Under the Curve (in ROC analysis), as well as 95% confidence intervals of the risk categories. This is important in promoting empirically supported practice. However, actuarial tools that are not sound—derived on relatively small samples (to which most statistical procedures are highly vulnerable), not well-validated, and/or having wide confidence intervals that make it difficult to meaningfully distinguish between risk categories—may advertise more than they can deliver. Developers of actuarial tools differ on their stance of whether to solely use actuarial methods to derive risk estimates, or to use actuarial tools as one component of a broader assessment approach (e.g., compare the approaches recommended by Monahan in still Chapter 5 and Hilton et al. in still Chapter 7).

SPJ tools typically have one section that functions much like an actuarial measure of risk, providing several risk factors which can be scored and summed to yield an overall risk measure for research purposes. In practice, evaluators are encouraged to consider all risk factors, including the total number of risk factors, but are not required to link their judgments directly and proportionately to linear numerical indices, as most actuarial measures require. The juxtaposition of these two practices raises a philosophical question ("How much accuracy is sufficient?") and a practical question ("How much accuracy is feasible?") regarding the prediction of violence. In the next section, the evidence regarding the relative accuracy of each approach is discussed. But it should also be noted that the third approach, anamnestic assessment, is not well-suited at all for drawing a quantitative conclusion about the risk for future violence.

When an individual's risk is being considered over a period of time, this often reflects an interest in how that risk might have changed. Skeem and her colleagues (2006) have distinguished between risk status (relatively unchanging aspects of violence risk) and risk state (aspects of risk that do change, even over short periods of time). If there are interventions being delivered with the goal of decreasing the person's likelihood of violent behavior, it is important to appraise the impact of such interventions through considering how overall risk may have changed. Some actuarial tools (e.g., the Level of Service/Case Management Inventory—Chapter 8 of this volume—and the Violence Risk Scale—still Chapter 10) do this, as they incorporate dynamic risk factors and hence provide a different overall "risk score" as the dynamic variables change. Other actuarial tools (e.g., the Classification of Violence Risk—still Chapter 5—and the

Table 1.1 Approaches and Important Questions in Risk Assessment

Question / *Approach*	*Prediction/Classification*	*Reflects Change in Risk Status*	*Provides Information for Risk Reduction Intervention Planning*
Actuarial	Yes	Variable	Limited
Structured professional judgment	Yes	Yes	Yes
Anamnestic	No	Limited	Yes

Violence Risk Appraisal Guide-revised—still Chapter 7) do not; they are designed to provide a stable estimate of the individual's risk that focuses on risk status.

SPJ tools are designed to appraise risk needs as well as risk status, so they offer a way of estimating change in risk state. They do so in the same manner that they provide the risk estimate: the evaluator considers the applicable risk factors, including the dynamic factors, and whether there is a perceived change in the necessary intensity of intervention or management, or the level or configuration of risk factors to be addressed. The evaluator then judges whether the overall risk has changed as a result. Further, it is important for evaluators to appraise and convey whether risk factors, and the relevance of such risk factors, have changed. Research has demonstrated that changes (increases or decreases) in SPJ-tool risk factors, and their relevance, are predictive of changes (increases or decreases) in subsequent inpatient and community violent behavior (Hogan & Olver, 2016, 2019).

Rather than assuming that a single risk estimate remains valid over time, it is important to provide periodic reassessments of risk factors (Douglas & Skeem, 2005). However, the question of how accurately any risk tools gauge the *change* in overall violence risk has seldom been researched (Heilbrun, Douglas, & Yasuhara, 2009). In one study, 30 male forensic psychiatric inpatients were administered the HCR-20, Version 2 (Webster, Douglas, Eaves, & Hart, 1997) and the Short-Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Desmarais, 2009; still Chapter 15) every three months for a year (Wilson, Desmarais, Nicholls, Hart, & Brink, 2013). The investigators reported that changes in dynamic risk factors significantly predicted institutional violence even after controlling for static factors. Dynamic risk factors were better predictors than historical risk factors of short-term institutional violence for periods of up to six months. The researchers suggested that changes in dynamic risk factors over time "are associated with future aggression beyond what can be predicted with historical variables alone" (Wilson et al., 2013, p. 385). In a study of short- (1 year) and long-term (11 years) risk level, improvements in dynamic risk and protective factors on the HCR-20^{V3} (still Chapter 12) and the Structured Assessment of Protective Factors for violence risk (SAPROF; de Vogel, de Ruiter, Bouman, & de Vries Robbé, 2009, 2012) during treatment were associated with lower violent recidivism rates (de Vries Robbé, de Vogel, Douglas, & Nijman, 2015). Findings indicated that non-recidivists made greater improvements during treatment than recidivists.

Several other studies have considered the relationship between changes in dynamic risk factors and violence risk. Following individuals released from civil and forensic hospitalization, investigators (Michel et al., 2013) reported that 6 of the 10 C and R items on the HCR-20 changed over time—and that this change was positively associated with changes in observed violence. This positive relationship (an increase in C and R item-defined risk associated with an increase in observed violence) has been reported by other investigators as well (Hogan & Olver, 2016; Penney, Marshall, & Simpson, 2016).

Anamnestic approaches can descriptively convey previous and current changes in risk-relevant needs through linkage with an intervention plan designed to address the specific risk factors identified in the initial assessment. When individualized risk factors are reduced in number and intensity through treatment, it is reasonable to assume that risk has been reduced. But how much? Over what period of time? These and related questions have not yet been answered through research.

Actuarial measures that accurately estimate risk do provide some information relevant to intervention planning by distinguishing between higher- and lower-risk individuals, applying the RNR risk principle that higher risk individuals are appropriate for more intensive interventions over a longer period of time (Andrews & Bonta, 2006). Unless dynamic factors are assessed, actuarial measures do not offer specific treatment targets and cannot measure changes in risk

after interventions are implemented. However, static risk factors can inform other aspects of risk-relevant intervention decision. For instance, Olver and Wong (2011) reported that Static-99 (Chapter 6 of this volume) predicted sexual recidivism well among offenders with a smaller treatment change versus those with greater treatment change. Static measures may also be used to design interventions based on both risk and needs, assuming risk-relevant needs can be measured in another way (see, e.g., Quinsey, Harris, Rice & Cormier, 1998; 2005). For example, assessing risk using an actuarial measure and risk-relevant needs using a separate measure, or by using anamnestic assessment, could identify at least four categories of risk needs (high risk/high needs, high risk/low need, low risk/high needs, and low risk/low needs) with associated needs for intervention and monitoring that differ considerably across categories.

Validation of Actuarial, Structured Professional Judgment, and Anamnestic Approaches to Risk Assessment

In this section, we highlight important trends in the empirical literature relevant to actuarial and structured professional judgment approaches to risk assessment. In addition, we discuss studies that offer an empirical comparison of the performance of tools of each type. There are no empirical studies on anamnestic risk assessment to be reviewed, but we offer brief comments on the validation status of this approach.

Actuarial Approaches to Violence Risk Assessment

The empirical literature in this area features a number of meta-analyses conducted since the development of the first specialized actuarial violence risk assessment instruments. These are in addition to the meta-analytic studies noted earlier citing a consistent, modest advantage in accuracy for actuarial relative to unstructured clinical prediction over 50 years of research (Ægosdóttir et al., 2006; Grove et al., 2000). The Mossman (1994) meta-analysis indicates that clinical predictions were more accurate than chance (AUC = .67) but modestly less accurate than cross-validated discriminant function predictions (AUC = .71)—a difference that was statistically significant ($p = .004$).

An important meta-analysis conducted more than two decades ago (Bonta, Law, & Hanson, 1998) considered studies of mentally disordered offenders and used outcomes of any kind of criminal recidivism (general offending) and crimes against persons (violent offending). As demonstrated in Table 1.2, there are a number of indications for support of predictors that could be incorporated into actuarial or SPJ prediction, including historical variables (criminal history, juvenile delinquency, hospital admissions, violence, escape), personality variables (antisocial personality), and substance abuse. The positive predictor with the largest recidivism effect size for violent recidivism was objective risk assessment (a measure that can be reliably scored, as contrasted with clinical judgment).

Two meta-analyses described earlier (Campbell et al., 2009; Fazel et al., 2012) included actuarial measures as well as SPJ measures. The mean effects of actuarial measures were comparable to SPJ measures (Campbell et al, 2009), while the positive predictive values (median = .41) of all measures were considerably lower than the negative predictive values. This means that actuarial (and other) risk assessment measures are quite accurate in identifying low-risk individuals, but less accurate in specifying who is high risk.

Consistent with the observed predictive strength of both historical variables and antisocial personality disorder, psychopathy is moderately associated with violent offending in the community. In one meta-analysis (Gendreau, Goggin & Smith, 2002), investigators compared the Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003) with the Level of Service Inventory-Revised

Table 1.2 Predictors of General and Violent Recidivism Among Mentally Disordered Offenders

	General		*Violent*	
	Recidivism Effect Size	*N*	*Recidivism Effect Size*	*N*
Positive Predictors				
Objective risk assessment	.39	1,295	.30	2,186
Adult criminal history	.23	4,312	.14	2,163
Juvenile delinquency	.22	4,312	.20	985
Antisocial personality	.18	1,736	.18	1,634
Nonviolent criminal history	.18	2,910	.13	1,108
Institutional adjustment	.13	627	.14	711
Hospital admissions	.12	1,874	.17	948
Poor living arrangements	.12	39	NR	
Gender (male)	.11	1,936	NR	
Substance abuse (any)	.11	234	.08	2,013
Family problems	.10	730	.19	1,481
Escape history	.10	646	NR	
Violent history	.10	2,240	.16	2,878
Drug abuse	.09	1,050	NR	
Marital status (single)	.07	987	.13	1,068
Negative Predictors				
Mentally disordered offender	−.19	3,009	−.10	2,866
Homicide index offense	−.17	1,147	NR	
Age	−.15	3,170	−.18	1,519
Violent index	−.14	905	−.04	2,241
Violent index (broadly defined)	−.10	3,240	.08	1,950
Sex offense	−.08	2,371	.04	1,636
Not guilty by reason of insanity	−.07	1,761	−.07	1,208
Psychosis	−.05	2,733	−.04	1,208
Mood disorder	−.04	1,856	.01	1,520
Treatment history	−.03	3,747	NR	
Offense seriousness	−.01	1,368	.06	1,879

Source: From Bonta et al. (1998)

(Andrews & Bonta, 1995; still Chapter 8) in their respective associations with both general and violent recidivism. Using over 50 studies, they found a slight difference in favor of the LSI-R for strength of association with violent recidivism.[5] Both specialized tools performed well, but this comparison is consistent with the tendency to use the PCL-R as a benchmark against which other tools are compared.[6]

Other meta-analyses have been conducted addressing the predictive validity of the PCL-R (Hare, 1991) for violent criminal recidivism. Walters (2003) described the capacities of the PCL-R and the Lifestyle Criminality Screening Form[7] (Walters, White, & Denney, 1991) to predict both general and violent criminal recidivism among offenders. The tools were equally effective in predicting criminal recidivism (AUC values were .67 for the PCL-R and .67 for the LCSF).

Another meta-analysis (Leistico, Salekin, DeCosta, & Rogers, 2008) considered the relationship between the PCL measures and antisocial conduct, including criminal recidivism. Using 95 published studies providing a merged dataset with over 15,000 participants, they observed a significant relationship between both Factor 1 (mean weighted $d = .40$) and Factor 2 (mean weighted $d = .57$) scores and the outcome of interest (violent offense recidivism risk). Factor 2 effect sizes

were significantly larger than those for Factor 1. Effect sizes were also moderated by participant characteristics, including gender, race, and setting (correctional versus hospital settings). This potentially limits the utility of the PCL-R in some settings, and for some people.

A meta-analysis addressing the incremental validity to various specialized measures relative to the PCL-R (Yang, Wong, & Coid, 2010) concluded that most specialized risk measures had comparable predictive validity, with few differences across measures. Factor 1 of the PCL (interpersonal and affective features) was less effective in risk assessment than Factor 2 (antisocial behavior history), with only two instruments—the HCR-20 and the OGRS—adding incremental validity to the PCL-R. This is consistent with the meta-analytic findings (Guy, Douglas, & Hendry, 2010) that the HCR-20 and the PCL-R performed comparably with respect to predictive accuracy (both had AUC values of .69).

Instruments designed more specifically to predict violence, and with a particular age cohort, may have superior performance to those designed to predict general offending (e.g., the Level of Service family of measures) or not designed as a risk assessment measure (e.g., the PCL-R). One relevant meta-analysis (Singh, Grann, & Fazel, 2011) observed that the SAVRY (an SPJ tool) produced the largest effect, with the LS and PCL-R having the smallest. Consistent with this, another meta-analysis (Fazel, Singh, Doll, & Grann, 2012) identified the VRAG (an actuarial measure) as well as a number of SPJ measures as having higher predictive validity than the LS and PCL-R, designed for more general outcomes (odds ratios of 6.1 versus 2.84).

The current view about the PCL-R as a risk assessment measure reflects the reality that it was developed as a measure of a personality disorder rather than as a specialized risk measure; once other such specialized risk measures have appeared and received research support, it is more appropriate to use them for the purpose of risk assessment. We note that the PCL-R has been among the actuarial measures of violence risk considered by researchers over nearly three decades. Typically, however, larger reviews (e.g., meta-analyses, meta-reviews) have included a number of actuarial measures of risk. For example, one meta-analysis focusing on risk assessment with sex offenders (Hanson & Morton-Bourgon, 2009) concluded that actuarial measures were more accurate than unstructured clinical judgment for three outcomes: any offending, violent offending, and sexual offending. SPJ measures had intermediate effect sizes. Actuarial measures were particularly effective with sexual recidivism outcomes, although there were only six SPJ studies included—and three of them used summary risk ratings as opposed to summation of scores.[8]

Other meta-analyses have reported comparable predictive effectiveness of actuarial and SPJ measures (see Olver, Stockdale, & Wormith, 2009; Viljoen, Mordell, & Beneteau, 2012). Smith, Cullen, and Latessa (2009) addressed the question of whether the Level of Service Inventory Revised (LSI-R; Andrews & Bonta, 2001) performed acceptably with a sample of female offenders. Smith et al. reviewed 25 published and unpublished datasets to evaluate the effect size of the relationship between LSI-R total score and recidivism for female offenders. The researchers found that the effect sizes in these studies were basically equivalent to the effect sizes found in samples of male offenders, indicating that the LSI-R worked comparably well with female offenders. Nevertheless, they suggested that a specialized risk assessment measure specific to females might perform even better (see also Manchak, Skeem, Douglas, & Siranosian, 2009).

These meta-analyses provided considerable evidence for the utility of the actuarial prediction of violent behavior, including criminal recidivism. Additional evidence comes from the MacArthur Risk Assessment Study (Monahan et al., 2001; Steadman et al., 1998), the largest and best-designed study to date of violent behavior in the community committed by those who had been discharged from inpatient mental health treatment facilities. The original MacArthur data were combined with additional validation data (Monahan, Steadman, Robbins et al., 2005) to yield an actuarial tool (the Classification of Violence Risk, or COVR; Monahan, Steadman, Appelbaum et al., 2005; see still Chapter 5) that is effective in predicting serious acts of violence in the community

committed by individuals with mental disorder. There was some shrinkage in the accuracy rates from derivation sample (AUC = .88) to the cross-validation sample (AUC = .63–.70, depending on the definition and measurement of outcomes), which is typical upon cross-validation.

Moving from meta-analysis to single cases, there have been recent developments in the debate concerning the applicability of actuarial measures to prediction in single cases. Some (see, e.g., Hart, Michie, & Cooke, 2007; Cooke & Michie, 2010; Hart & Cooke, 2013) have argued that the application of an actuarial prediction to a single case creates such serious statistical problems that those results are nearly meaningless when so applied. These arguments have not met with universal agreement within the field (see, e.g., Harris & Rice, 2007, 2015; Mossman, 2007, 2015), but the statistical issues involved were sufficiently complex that for several years the debate remained active. Some have concluded that the statistical arguments against the accuracy of actuarial instruments in the individual case are not persuasive (Imrey & Dawid, 2015; although cf. Hart & Cooke, 2013). It *is* important that an actuarial tool be derived and validated on large samples, both to increase its generalizability and to ensure that the confidence intervals are narrow and the risk categories do not overlap. An earlier discussion of this debate may be found elsewhere (Heilbrun et al., 2009).

One of the active debates in the field over the years has involved the way in which actuarial weighting and scoring can miss certain influences that may apply in the individual case. Relatively few studies of the impact of "clinical overrides" of actuarial risk levels—the "adjusted actuarial" procedure—have been published. Apparently the first such study (Gore, 2008) considered the accuracy of the actuarial versus "adjusted actuarial" conclusions in 383 cases in which there had been a clinical override of the Minnesota Sex Offender Screening Tool-Revised (MnSOST-R; Epperson et al., 2003) risk level. The clinical adjustment did not yield increased accuracy, but it also did not lower the overall accuracy substantially. MnSOST-R actuarial levels were more accurate than the adjusted levels, but the difference was not statistically significant. In addition, the direction of the adjustment was not related to measured predictive accuracy. Upon conducting their meta-analysis, Hanson and Morton-Bourgon (2009) concluded that clinical adjustment to actuarial prediction tended to decrease overall predictive accuracy. Wormith and colleagues (this volume) discuss research with the LS system that fails to show support for adjustment of actuarial predictions in terms of enhancing predictive accuracy, and they recommend that they be used only sparingly, under extraordinary circumstances. Scurich and Krauss (2013) found that clinical adjustment significantly increased mock juror willingness to commit an offender under a sexually violent predator (SVP) statute, but only when the adjustment was upward (i.e., indicating higher risk). Downward adjustment had no effect on the mock jurors.

In addition, the process of weighting predictors in the course of actuarial assessment has been considered in another study. Using 10 well-established risk factors for violent criminal recidivism from the HCR-20, the investigators (Grann & Långström, 2007) considered the impact of four different weighting schemes in applying these predictors to 404 forensic psychiatric patients who had been released to their communities in Sweden. These weighting approaches included Nuffield's method, bivariate and multivariate logistic regression, and an artificial neural network procedure. Upon cross-validation, they reported that simpler weighting techniques did not increase predictive accuracy over an unweighted approach. In addition, more complex weighting procedures were associated with significant shrinkage.

The findings of these two studies (Gore, 2008; Grann & Långström, 2007) may suggest that the use of empirically validated risk factors in prediction yields an increase in accuracy over unstructured judgment, but there may be limits to the increases in accuracy that can be achieved through the specific weighting and combination of predictors that are typically present in actuarial prediction (a point made roughly 40 years ago by Dawes, 1979). Accordingly, some commentators have opposed the practice of weighting the risk factors in risk assessment instruments,

when simple unit weighting (equal weights for all risk factors) is more accurate (see Harris & Rice, 2015; Grann & Långström, 2007).

On this point, Kroner, Mills, and Reddon (2005) compared the predictive accuracy of three widely recognized actuarial tools (the PCL-R, LSI-R, and VRAG), another approach using "General Statistical Information on Recidivism," and four additional instruments that had been constructed through randomly selecting items from the total pool of original items. Intriguingly, none of the three actuarial tools or the fourth using general statistical information on recidivism was more accurate in predicting post-release failure than the four randomly generated instruments. The investigators argued that the development of better risk theory, yielding the capacity to test hypotheses and explain behavior, is necessary to advance the process of risk assessment in the criminal justice and mental health fields. A related explanation—that there is a ceiling on the predictive accuracy that can be obtained using any approach that incorporates empirically supported risk factors, regardless of how these factors are weighted or combined—is discussed in more detail later in this chapter.

Structured Professional Judgment Approaches to Violence Risk Assessment

Structured professional judgment is an approach to risk assessment that began to be used in the 1990s and has continued to evolve (see Douglas, 2019). It is more recent than actuarial risk assessment, and also less oriented toward quantifying conclusions. Rather than yielding a final numerical score, SPJ tools typically call for a risk judgment of low, moderate, or high. Assuming that an SPJ tool uses risk factors whose rating can be quantified, it is possible to study that tool by using these risk factors in an actuarial sense—by obtaining a "total score" and relating that score to observed outcome. However, it is important to test SPJ tools in the way it is recommended they be used. Rather than relying on specific weights for predictors that would be derived using an actuarial approach, the SPJ approach uses risk factors that have support in the broad empirical literature; it also makes the general assumption that the presence of a greater number of risk factors is typically associated with higher risk (Douglas & Kropp, 2002). Research on the relationship between final risk judgments and outcomes is most meaningful in testing this approach. That is what we will describe in this section.

There have been 45 published studies and at least one dissertation (McGowan, 2007) addressing the relationship between final SPJ risk judgments and violence (Douglas, 2019). Of these 45 studies, 39 have yielded findings to the effect that SPJ judgments are significantly predictive of violent recidivism.[9] In addition, 18 of these studies assessed whether the SPJ "final judgment" adds incremental predictive accuracy to the use of the tool elements combined in an actuarial fashion (Douglas, 2019). In 16 of these 18 studies, such incremental validity was observed.

Taken together, these studies provide very strong evidence that structured professional judgment is an efficacious approach to risk assessment. But how does it compare to actuarial assessment? In certain respects this is not a meaningful question, as actuarial and structured professional judgment share several important features. Both use risk factors that are specified *a priori* and can be rated reliably. In both *data selection* and *data coding*, therefore, these approaches are quite similar. It is in *data combination* that they diverge. Actuarial approaches combine data according to an established formula derived and validated on specific datasets; SPJ approaches call for the combination of data into a professional judgment that is "structured" through provision of risk factors (obtained from the broader literature but typically not from any specific dataset) and guided formulation (or integration of risk factors) that can be rated reliably.

Accordingly, we might attribute observed differences between actuarial and SPJ approaches to their divergence in data combination, as they are similar in data selection and data coding. With that in mind, we now review studies that have compared these approaches.

Actuarial vs. Structured Professional Judgment Approaches to Violence Risk Assessment

Over the last two decades, a number of meta-analyses and reviews have compared the performance of actuarial versus SPJ measures. We regard such aggregate analyses and reviews as the most stable form of evidence on this question. We will summarize the relevant findings of such studies (discussed earlier, particularly in the Validation of Actuarial Approaches section) first, and then discuss the most relevant of these in greater detail. Generally, the evidence seems to identify little meaningful difference between actuarial and SPJ approaches to predictive accuracy—even when some of the studies reviewed use summed item scores rather than summary judgments for SPJ measures. Further, to the extent that a measure is developed for a more specific purpose (e.g., predicting violence rather than general offending), or for use with a more specific population (e.g., adolescents, adults), that seems to enhance its effectiveness.

Douglas and his colleagues (2005) compared the respective predictive validities of an SPJ measure (the HCR-20) and several actuarial measures (the Violence Risk Appraisal Guide and the Hare Psychopathy Checklist-Revised and Psychopathy Checklist: Screening Version), among others. Several indices were related to violent recidivism with large effect sizes. These included HCR-20 structured risk judgments, VRAG scores, and behavioral scales of psychopathy measures (PCL-R Factor 2, Cooke and Michie's (2001) Factor 3, and PCL-SV Part 2). These findings are consistent with strong predictive performance by actuarial measures, but not with a noteworthy advantage to such measures. The HCR-20 structured risk judgments were also strongly related to violent outcomes.

A similar comparison between the HCR-20 and two actuarial measures (the Level of Service Inventory-Revised edition and the Psychopathy Checklist-Revised) was conducted on offenders in Germany (Dahle, 2006). The investigator noted only minor differences in predictive accuracy among these measures. As with the study by Douglas and his colleagues (2005), all measures performed reasonably well, but without noteworthy advantage to any particular measure. It is worth noting that in the latter study, direct entry of variables by the investigator into the regression equation resulted in significance of both the VRAG score and HCR summary risk rating, while stepwise entry (using automatic entry of the most strongly predictive variables in succession, after the variance accounted for by previously entered variables is no longer considered) yielded significance for only the HCR summary risk ratings.

In another comparison of an actuarial measure (the Static-99) with an SPJ measure (the SVR-20), de Vogel, de Ruiter, van Beek, and Mead (2004) observed significantly better predictive performance of the latter. They scored these tools using file information from 122 sex offenders who had been admitted to a Dutch forensic psychiatric hospital between 1974 and 1996 and had been released to and residing in the community for almost 12 years. Base rates of reconvictions over this period were 39% (sexual offenses), 46% (nonsexual violent offenses), and 74% (any offenses). Predictive validity for the SVR-20 was good for both the total score (AUC = .80) and the final risk judgment (AUC = .83), and reasonably good but significantly lower for the Static-99 total score (AUC = .71) and risk categories (AUC = .66).

A study with adolescents (Catchpole & Gretton, 2003) compared two actuarial approaches—the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2002; still Chapter 9) and the Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003)—with the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2002; still Chapter 18). A total of 133 participants were obtained from consecutive referrals from youth court judges. The YLS/CMI ratings were completed by probation officers; the PCL:YV and the SAVRY were coded using file data. Follow-up periods ranged between 7 and 61 months, with a mean of 35.8 months. Violent recidivism was

among the outcomes measured. For such violent recidivism, the SAVRY (AUC = .81) was most accurate predictively, followed by the PCL:SV (AUC = .73) and the YLS/CMI (AUC = .64).

The relationship between the PCL-R and the HCR-20 (Version 2) was considered in samples (k = 34) containing both (Guy, Douglas, & Hendry, 2010). The two instruments had equivalent AUC values (.69). The investigators were able to determine (through using raw data sets) that the HCR-20 added unique incremental validity to the PCL-R, but the PCL-R did not add such validity to the HCR-20—and removing the psychopathy item from the HCR-20 did not decrease its predictive accuracy.

Hanson and Morton-Bourgon (2009) conducted a meta-analysis of 118 sex offender studies and reported that actuarial measures produced the largest effect sizes, with unstructured clinical judgment having the smallest. SPJ measures (with the exception of the SVR-20, which had the largest effect size) were intermediate in the effect sizes reported, although the small number of SPJ studies meant that a meaningful comparison between actuarial and SPJ effectiveness was not possible. In another meta-analysis focusing on risk assessment of youth, the measure with the strongest effect size was the SAVRY (an SPJ tool) (.30 to .38), while lower effect sizes were observed for the YLS/CMI (an actuarial measure; .26 to .32). No meaningful differences between actuarial and SPJ measures were reported in another meta-analysis of adolescent sexual offender assessment (Viljoen, Mordell, & Beneteau, 2012). This is very consistent with findings from a different meta-analysis (Singh, Grann, & Fazel, 2011) reporting the strongest effect sizes for the SAVRY and the weakest for the LS measures.

Nine risk assessment tools in another meta-analysis (including the HCR-20 and VRAG) predicted violence at approximately a moderate level. These measures included both actuarial and SPJ approaches and considered whether any of the included risk assessment measures could improve on the accuracy associated with the PCL-R. There were few differences between measures, with only the HCR-20 (based on 16 samples) and the Offender Group Reconviction Scale (Copas & Marshall, 1998; based on two samples) adding the predictive accuracy of the PCL-R. It is noteworthy that the meta-analysis did not use the summary risk ratings (provided by the evaluators) from the SPJ measures (Yang et al., 2010).

There are a growing number of studies on which to base a conclusion about the relative predictive accuracy of actuarial measures versus SPJ approaches using a "final judgment" as the risk classification. But the available evidence at present, much of it gathered since the first version of this chapter was written a decade ago, strongly suggests that actuarial and SPJ approaches are at least comparable in predictive accuracy for violent outcomes. This is an important observation, as this book details the structure and functions of both SPJ and actuarial risk assessment tools.

Anamnestic Approaches to Violence Risk Assessment

The anamnestic approach to risk assessment does not lend itself easily to research on predictive accuracy. Indeed, the identification of risk factors, particularly those that are subject to change and hence targets for intervention, by using the individual's own history of violence is much better suited to risk management. But focusing part of the risk assessment on the individual's history of violence will remain an important step when informing a broader legal decision. Three considerations are worth noting in this respect. First, the assumption of some relationship between the number of risk factors present and the overall risk level has some support from SPJ studies that allow the evaluator to make a final judgment in light of the number and intensity of observed risk factors. Final predictions made in this context seem about as accurate as actuarial predictions. Second the process of "individualizing" the evaluation of forensic mental health assessment (FMHA) is quite important as a broad principle in forensic assessment, including risk assessment done in forensic contexts (Heilbrun, 2009). Third, there are occasionally exceptions that involve

a substantial difference in the individual's personal attributes or situation that distinguish their present from most of their history. One of the best ways to learn about such exceptions is through obtaining detailed information about a person's history of violent behavior and the circumstances surrounding each event.[10]

Although it would be difficult, it is certainly feasible to investigate the extent to which anamnestic assessments accurately identify level of risk and relevant treatment targets and lead to intervention planning that is more effective in reducing risk. Such research would be similar to a "clinical trial" in which one approach (using anamnestic risk assessment) was compared with others (e.g., a standard risk assessment tool, "usual practice" by a treatment team, and perhaps some combinations). The added demand to deliver interventions makes this a more challenging study than the typical "administer tool à observe outcome" design of a prediction study. But this is true for any investigation of risk reduction, and it would be important to use the anamnestic approach for what it should do best (identify treatment targets to reduce risk) as well as what it may do (help inform an accurate conclusion about risk level).

Future Directions

The development of specialized assessment tools continues to provide one of the best avenues to promoting empirically informed practice. The various chapters in this book provide descriptions of risk assessment tools that have been developed and validated for some of the various purposes and populations described in the present chapter. Promising directions for risk assessment include increased application of risk assessment to risk reduction planning and delivery, combining data gathered by use of risk assessment measures with data from other measures that may provide moderating influences (e.g., clinical functioning, situational variables), and the incorporation of the influences of trauma and adverse experience into risk assessment and risk reduction. Practitioners might observe a modest advantage accruing from the development of a specialized measure with a specific population, such as adolescence, or for a specific purpose, such as assessing the risk of violent offending rather than general offending. Researchers might consider the accumulating evidence that specialized measures, whatever their numerous advantages, may be hitting a ceiling in their observed accuracy defined by AUC values around .80. This might suggest that modifications in our approach to developing and using specialized risk assessment measures are needed. Both practitioners and researchers might consider the chapters that follow in light of these conclusions.

Notes

1. There are occasional exceptions in which a court does maintain jurisdiction over an individual following civil commitment, such as in cases involving individuals committed as sexually violent predators.
2. The term "protective factor" has been used imprecisely in much of the literature in this area. A recent useful clarification has been provided by Monahan and Skeem (2016), who described "risk factors" as influences directly associated with adverse outcomes, "promotive factors" as influences directly associated with favorable outcomes, and "protective factors" as those reducing the influence of risk factors (in effect, moderating the influence of risk factors).
3. The "adjusted actuarial" debate addresses whether the results of actuarial risk assessment should be modified in light of other considerations. As noted elsewhere (Heilbrun, 2009), this can involve either "adjusting" the score or the final risk level yielded by the actuarial tool at the scoring stage, or adjusting their meaning at the interpretation stage. Changing a score or risk level at the scoring stage is extremely problematic. Actuarial tools were designed and validated to be used with an established algorithm. Changing this (for whatever reason) would raise questions about the meaning of the adjusted score. It would substitute clinical judgment for the actuarial algorithm, reducing accuracy over multiple cases. If the evaluator is not clear about doing this, the result may appear actuarial when it is something different. For these reasons, adjusting actuarial results at the scoring stage should not be done. At the interpretation

stage, however, the actuarial tool may be one of multiple sources of information contributing to the final conclusion. In forensic mental health assessment, no single source of information is considered definitive and final opinions result from integration across sources (Heilbrun, 2001). But actuarial measures provide valuable data concerning violence risk, so deviation from their conclusions should involve (1) noting the actuarial results, (2) describing the deviation, (3) noting the information contributing to the different opinion, and (4) describing the logic underlying the evaluator's reasoning.

4. Anamnestic assessment is described here as a distinct kind of risk assessment. However, it is a useful complement to other aspects of risk appraisal in any evaluation. In fact, it is similar to the formulation aspect of SPJ, which attempts to answer *why* a person has been violent, and acknowledged as such by SPJ developers (Douglas, Hart, Webster, & Belfrage, 2013). In effect, it provides a way of individualizing the risk assessment that increases both accuracy and associated credibility—both important considerations in forensic mental health assessment.
5. They reported effect sizes of .26 for the LSI-R and .21 for the PCL-R, a difference the authors suggest is significant using fail-safe calculations. Fail-safe analyses indicated that there would need to be 37 additional PCL-R studies with the maximum observed effect size, or 22 LSI-R effect sizes of 0, in order for their predictive performance to be equal.
6. It might also be noted that there is some judgment required in rating certain items on each of these tools, but the items are combined through numerical scores, and the total score is considered to be the indicator of risk—so each qualifies as an "actuarial" tool in this respect. It is also somewhat ironic that the PCL-R has come to be considered by some as a benchmark for the performance of risk assessment tools; it was not developed as a risk assessment measure, but as a way of assessing a specific personality disorder.
7. The LCSF is a procedure that uses only information from the file to rate an individual's irresponsibility, self-indulgence, interpersonal intrusiveness, and social rule-breaking.
8. This is a common methodological problem in aggregate reviews of the effectiveness of SPJ measures. Such measures are designed to use professional judgment ("summary judgment") after considering item content rather than sum item scores. To the extent that summary judgment enhances SPJ effectiveness (discussed elsewhere in this chapter), using summed item scores rather than summary judgments artificially truncates the effectiveness of SPJ measures.
9. These studies include Arbach-Lucioni, Andres-Pueyo, Pomarol-Clotet, & Gomar-Sones, 2011; Belfrage, Strand, Ekman, & Hasselborg, 2012; Braithwaite, Charette, Crocker, & Reyes, 2010; Catchpole & Gretton, 2003; de Vogel & de Ruiter, 2005, 2006; de Vogel, de Ruiter, Hildebrand, Bos, & van de Ven, 2004; de Vogel, de Ruiter, van Beek, & Mead, 2004; de Vries Robbé, de Vogel, & de Spa, 2011; de Vries Robbé, de Vogel, Koster, & Bogaerts, 2014; de Vries Robbé, de Vogel, Douglas, & Nijman, 2015; de Vries Robbé, de Vogel, Wever, Douglas, & Nijman, 2016; Desmarais, Nicholls, Wilson, & Brink, 2012; Dolan & Rennie, 2008; Douglas, Ogloff, & Hart, 2003; Douglas, Yeomans, & Boer, 2005; Enebrink, Långström, & Gumpert, 2006; Folino, 2015; Gammelgård, Koivisto, Eronen, & Kaltiala-Heino, 2008; Ho et al., 2015; Hogan & Olver, 2016; Jovanović, Toševski, Ivković, Damjanović, & Gašić, 2009; Kropp & Hart, 2000; Langton, Hogue, Daffern, Mannion, & Howells, 2009; Lodewijks, de Ruiter, & Doreleijers, 2008; Lodewijks, Doreleijers, & de Ruiter, 2008; Lodewijks, Doreleijers, de Ruiter, & Borum, 2008; Meyers & Schmidt, 2008; Neal, Miller, & Shealy, 2015; Neves, Goncalves, & Palma-Oliveira, 2011; O'Shea, Picchioni, Mason, Sugarman, & Dickens, 2014; O'Shea et al., 2016; Nonstad et al., 2010; Pedersen, Rasmussen, & Elsass, 2010; Penney, Lee, & Moretti, 2010; Rajlic & Gretton, 2010; Sada et al., 2016; Schaap, Lammers, & de Vogel, 2009; Schmidt, Campbell, & Houlding, 2011; Sjöstedt & Långström, 2002; Strub, Douglas, & Nicholls, 2014; van den Brink, Hooijschuur, van Os, Savenije, & Wiersma, 2010; Verbrugge, Goodman-Delahunty, & Frize, 2011; Viljoen et al., 2008, 2016; Vincent, Chapman, & Cook, 2011.
10. The presence of an exceptional circumstance or substantial change in personal characteristics might suffice to cause the evaluator to avoid using a risk assessment measure that does not allow the incorporation of such a difference into the formal evaluation.

References

Abbiati, M., Palix, J., Gasser, J., & Moulin, V. (2019). Predicting physically violent misconduct in prison: A comparison of four risk assessment instruments. *Behavioral Sciences and the Law*, *37*, 61–77.

Ægosdóttir, S., White, M. J., Spengler, P. M., Maugherman, L. A., Cook, R. S., Nichols, C. N., . . . Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*, 341–382.

American Educational Research Association, American Psychological Association, & The National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Andrews, D. A., & Bonta, J. (1995). *The level of service inventory-revised*. Toronto, ON: Multi-Health Systems.

Andrews, D. A, & Bonta, J. (2001). *Level of service inventory-revised (LSI-R): User's manual*. Toronto, ON: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexis.

Andrews, D. A, Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19–52.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *Level of service/case management inventory (LS/CMI): An offender assessment system user's manual*. North Tonawanda, NY: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, S. (2006). Recent past and near future of risk/need assessment. *Crime and Delinquency*, *52*, 7–27.

Arbach-Lucioni, K., Andres-Pueyo, A., Pomarol-Clotet, E., & Gomar-Sones, J. (2011). Predicting violence in psychiatric inpatients: A prospective study with the HCR-20 violence risk assessment scheme. *Journal of Forensic Psychiatry & Psychology*, *22*, 203–222.

Belfrage, H., Strand, S., Ekman, L., & Hasselborg, A. (2012). Assessing risk of patriarchal violence with honour as a motive: Six years experience using the PATRIARCH checklist. *International Journal of Police Science & Management*, *14*, 20–29.

Bonta, J., & Andrews, D. (2017). *The psychology of criminal conduct* (6th ed.). New York: Routledge.

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, *124*, 123–142.

Borum, R., Bartel, P., & Forth, A. (2002). *Manual for the structured assessment of violence risk in youth: Consultation version*. Tampa: University of South Florida, Florida Mental Health Institute.

Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the short-term assessment of risk and treatability (START). *International Journal of Forensic Mental Health*, *9*, 271–281.

Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior*, *36*, 567–590.

Catchpole, R. E. H., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior*, *30*, 688–708.

Childs, K., Frick, P., Ryals, J., Lingonblad, A., & Villio, M. (2014). A comparison of empirically based and structured professional judgment estimation of risk using the structured assessment of violence risk in youth. *Youth Violence and Juvenile Justice*, *12*, 40–57.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, *13*, 171–188.

Cooke, D. J., & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and Human Behavior*, *34*, 269–274.

Cooke, D. J., Wozniak, E., & Johnstone, L. (2008). Casting light on prison violence in Scotland: Evaluating the impact of situational risk factors. *Criminal Justice and Behavior*, *35*, 1065–1078.

Copas, J., & Marshall, P. (1998). The offender group reconviction scale: The statistical reconviction score for use by probation officers. *Journal of the Royal Statistical Society*, *47C*, 159–171.

Daffern, M., & Howells, K. (2007). The prediction of imminent aggression and self-harm in personality disordered patients of a high security hospital using the HCR-20 clinical scale and the dynamic appraisal of situational aggression. *International Journal of Forensic Mental Health*, *6*, 137–143.

Dahle, K. P. (2006). Strengths and limitations of actuarial prediction of criminal reoffence in a German prison sample: A comparative study of LSI-R, HCR-20 and PCL-R. *International Journal of Law and Psychiatry*, *29*, 341–442.

Dawes, R. (1979). The robust beauty of improper linear models. *American Psychologist*, *34*, 571–582.

Desmarais, S. L., Nicholls, T., Wilson, C. M., & Brink, J. (2012). Reliability and validity of the short-term assessment of risk and treatability (START) in assessing risk for inpatient aggression. *Psychological Assessment*, *24*, 685–700.

de Vogel, V., & de Ruiter, C. (2005). The HCR-20 in personality disordered female offenders: A comparison with a matched sample of males. *Clinical Psychology & Psychotherapy*, *12*, 226–240.

de Vogel, V., & de Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law*, *12*, 321–336.

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2009). *SAPROF: Guidelines for the assessment of protective factors for violence risk* (English version). Utrecht, The Netherlands: Forum Educatief.

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2012). *SAPROF: Guidelines for the assessment of protective factors for violence risk* (2nd ed.). Utrecht, The Netherlands: Vander Hoeven Stichting.

de Vogel, V., de Ruiter, C., Hildebrand, M., Bos, B., & van de Ven, P. (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health*, *3*, 149–165.

de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and static 99 in a Dutch sample of treated sex offenders. *Law and Human Behavior*, *28*, 235–251.

de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients: A retrospective validation study of the SAPROF. *International Journal of Forensic Mental Health*, *10*, 178–186.

de Vries Robbé, M., de Vogel, V., & Douglas, K. (2013). Risk factors and protective factors: A two-sided dynamic approach to violence risk assessment. *Journal of Forensic Psychiatry & Psychology*, *24*, 440–457.

de Vries Robbé, M., de Vogel, V., Douglas, K. S., & Nijman, H. L. I. (2015). Changes in dynamic risk and protective factors for violence during inpatient forensic psychiatric treatment: Predicting reductions in postdischarge community recidivism. *Law and Human Behavior*, *39*, 53–61. doi:10.1037/lhb0000089.

de Vries Robbé, M., de Vogel, V., Koster, K., & Bogaerts, S. (2014). Assessing protective factors for sexually violent offending with the SAPROF. *Sexual Abuse: A Journal of Research and Treatment*, *27*, 51–70.

de Vries Robbé, M., de Vogel, V., Wever, E. C., Douglas, K. S., & Nijman, H. L. I. (2016). Risk and protective factors for inpatient aggression. *Criminal Justice and Behavior*, *43*, 1364–1385.

Dolan, M. C., & Rennie, C. E. (2008). The "structured assessment of violence risk in youth" as a predictor of recidivism in a United Kingdom cohort of adolescent offenders with conduct disorder. *Psychological Assessment*, *20*, 35–46.

Douglas, K. S. (2019). Evaluating and managing risk for violence using structured professional judgment. In A. Day, C. Hollin, & D. Polaschek (Eds.), *Wiley international handbook of correctional psychology* (pp. 429–445). Hoboken, NJ: Wiley & Sons.

Douglas, K. S., Hart, S. D., Groscup, J. L., & Litwack, T. R. (2014). Assessing violence risk. In I. Weiner & R. K. Otto (Eds.), *The handbook of forensic psychology* (4th ed., pp. 385–441). Hoboken, NJ: Wiley & Sons.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 (version 3): Assessing risk for violence, user guide*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, *29*, 617–658.

Douglas, K., & Ogloff, J. R. P. (2003). The impact of confidence on the accuracy of structured professional and actuarial violence risk judgments in a sample of forensic psychiatric patients. *Law and Human Behavior*, *27*, 573–587.

Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services*, *54*, 1372–1379.

Douglas, K. S., & Skeem, J. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, *11*, 347–383.

Douglas, K. S., Yeomans, M., & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*, *32*, 479–510.

Enebrink, P., Långström, N., & Gumpert, C. H. (2006). Predicting aggressive and disruptive behavior in referred 6- to 12-year-old boys: Predictive validation of the EARL-20B risk/needs checklist. *Assessment*, *13*, 356–367.

Epperson, D., Kaul, J., Huot, S., Goldman, R., & Alexander, W. (2003). *Minnesota sex offender screening tool-revised (MnSOST-R) technical paper: Development, validation, and recommended risk level cut scores*. Currently unpublished manuscript. Retrieved April 9, 2008, from www.psychology.iastate.edu/~dle/TechUpdatePaper12-03.pdf

Fazel, S., Singh J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *British Medical Journal*, *345*, e4692. doi:10.1136/bmj.e4692

Folino, J. (2015). Predictive efficacy of violence risk assessment instruments in Latin-America: Predictive efficacy of violence risk assessment instruments in Latin-America. *The European Journal of Psychology Applied to Legal Context*, *7*, 51–58.

Forth, A., Kosson, D., & Hare, R. (2003). *The psychopathy checklist: Youth version.* Toronto, ON: Multi-Health Systems.

Gammelgård, M., Koivisto, A., Eronen, M., & Kaltiala-Heino, R. (2008). The predictive validity of the structured assessment of violence risk in youth (SAVRY) among institutionalised adolescents. *Journal of Forensic Psychiatry & Psychology*, *19*, 352–370.

Garcia-Mansilla, A., Rosenfeld, B., & Cruise, K. (2011). Violence risk assessment and women: Predictive accuracy of the HCR-20 in a civil psychiatric sample. *Behavioral Sciences and the Law*, *29*, 623–633.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? A lesson in knowledge cumulation. *Criminal Justice and Behavior*, *29*, 397–426.

Gore, K. (2008). Adjusted actuarial assessment of sex offenders: The impact of clinical overrides on predictive accuracy. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *68*(7B), 4824.

Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behavior*, *27*, 97–114.

Grann, M., & Långström, N. (2007). Actuarial assessment of violence risk: To weigh or not to weigh? *Criminal Justice and Behavior*, *34*, 22–36. http://dx.doi.org/10.1177/0093854806290250

Grann, M., & Wedin, I. (2002). Risk factors for recidivism among spousal assault and spousal homicide offenders. *Psychology, Crime & Law*, *8*, 5–23.

Gray, N., Taylor, J., & Snowden, R. (2011). Predicting violence using structured professional judgment inpatients with different mental and behavioral disorders. *Psychiatry Research*, *187*, 248–253.

Grove, W., & Meehl, P. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, & Law*, *2*, 293–323.

Grove, W., Zald, D., Lebow, B., Snitz, B., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, *12*, 19–30.

Guy, L., Douglas, K. S., & Hendry, M. (2010). The role of psychopathic personality disorder in violence risk assessments using the HCR-20. *Journal of Personality Disorders*, *24*, 551–580.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessment for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, *21*, 1–21.

Hanson, R. K., Morton-Bourgon, K. E., & Harris, A. (2003). Sexual offender recidivism risk: What we know and what we need to know. *Annals of the New York Academy of Sciences*, *989*, 154–166.

Hare, R. (1991). *The Hare Psychopathy Checklist-revised.* Toronto, ON: Multi-Health Systems.

Hare, R. (2003). *The Hare Psychopathy Checklist-revised* (2nd ed.). Toronto, ON: Multi-Health Systems.

Harris, G. T., & Rice, M. E. (2007). Characterizing the value of actuarial violence risk assessments. *Criminal Justice and Behavior*, *34*, 1638–1658.

Harris, G. T., & Rice, M. E. (2013). Bayes and base rates: What is an informative prior for actuarial violence risk assessment? *Behavioral Sciences and the Law*, *31*, 103–124.

Harris, G. T., & Rice, M. E. (2015). Progress in violence risk assessment and communication: Hypothesis versus evidence. *Behavioral Sciences and the Law*, *33*, 128–145.

Harris, G. T., Rice, M., & Cormier, C. (2002). Prospective replication of the "violence risk appraisal guide" in predicting violent recidivism among forensic patients. *Law and Human Behavior*, *26*, 377–394.

Harris, P. (2006). What community supervision officers need to know about actuarial risk assessment and clinical judgment. *Federal Probation*, *70*, 8–14.

Hart, S. (2009). Evidence-based assessment of risk for sexual violence. *Chapman Journal of Criminal Justice*, *1*, 143–165.

Hart, S. D., & Cooke, D. (2013). Another look at the (im-)precision of individual risk estimates made using actuarial risk assessment instruments. *Behavioral Sciences and the Law*, *31*, 81–102.

Hart, S. D., Douglas, K. S., & Guy, L. S. (2017). The structured professional judgment approach to violence risk assessment: Origins, nature, and advances. In M. Rettenberger & L. Craig (Series Eds.) & D. Boer (Vol. Ed.), *The Wiley-Blackwell handbook on the assessment, treatment and theories of sexual offending: Volume 1. Assessment* (pp. 643–666). Oxford: Wiley-Blackwell.

Hart, S. D., Michie, C., & Cooke, D. (2007). Precision of actuarial risk assessment instruments: Evaluating the "margins of error" of group versus individual predictions of violence. *British Journal of Psychiatry*, *190*(Suppl. 49), s60–s65.

Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal context. *Law and Human Behavior*, *21*, 347–359.

Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic, Plenum Press.

Heilbrun, K. (2009). *Evaluation for risk of violence in adults*. New York: Oxford University Press.

Heilbrun, K., Douglas, K., & Yasuhara, K. (2009). Violence risk assessment: Core controversies. In J. Skeem, K. Douglas, & S. Lilienfeld (Eds.), *Psychological science in the courtroom: Controversies and consensus* (pp. 333–357). New York: Guilford Press.

Heilbrun, K., Rogers, R., & Otto, R. (2002). Forensic assessment: Current status and future directions. In J. Ogloff (Ed.), *Psychology and law: Reviewing the discipline* (pp. 120–147). New York: Kluwer Academic, Plenum Press.

Ho, R. M., Cheung, H. H., Lai, T. T., Tam, V. F., Yan, C. K., Chan, W. L., & Yuen, K. K. (2015). Use of the historical, clinical, risk management-20 to assess the risk of violence by discharged psychiatric patients. *Hong Kong Medical Journal*, *21*, 45–47.

Hogan, N. R., & Olver, M. E. (2016). Assessing risk for aggression in forensic psychiatric inpatients: An examination of five measures. *Law and Human Behavior*, *40*, 233–243.

Hogan, N. R., & Olver, M. E. (2019). Static and dynamic assessment of violence risk among discharged forensic patients. *Criminal Justice and Behavior*, *46*, 923–928.

Hoge, R., & Andrews, D. (2002). *The youth level of service/case management inventory manual and scoring key*. Toronto, ON: Multi-Health Systems.

Imrey, P. B., & Dawid, A. P. (2015). A commentary on statistical assessment of violence recidivism risk. *Statistics and Public Policy*, *2*, 1–18.

Jovanović, A. A., Toševski, D. L., Ivković, M., Damjanović, A., & Gašić, M. J. (2009). Predicting violence in veterans with posttraumatic stress disorder. *Vojnosanitetski Pregled: Military Medical & Pharmaceutical Journal of Serbia & Montenegro*, *66*, 13–21.

Kroner, D., Mills, J., & Reddon, J. (2005). A coffee can, factor analysis, and prediction of antisocial behavior: The structure of criminal risk. *International Journal of Law and Psychiatry*, *28*, 360–374.

Kropp, P. R., & Hart, S. D. (2000). The spousal assault risk assessment (SARA) guide: Reliability and validity in adult male offenders. *Law and Human Behavior*, *24*, 101–118.

Kropp, P. R., & Hart, S. D. (2015). *Spousal assault risk assessment guide user's manual* (3rd ed.). Vancouver, BC: Protect International Risk and Safety Services, Inc.

Kropp, P., Hart, S., & Lyon, D. (2002). Risk assessment of stalkers: Some problems and possible solutions. *Criminal Justice and Behavior*, *29*, 590–616.

Langton, C. M., Hogue, T. E., Daffern, M., Mannion, A., & Howells, K. (2009). Prediction of institutional aggression among personality disordered forensic patients using actuarial and structured clinical risk assessment tools: Prospective evaluation of the HCR-20, VRS, static-99, and risk matrix 2000. *Psychology, Crime & Law*, *15*, 635–659.

Leistico, A., Salekin, R., DeCosta, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, *32*, 28–45.

Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, T. A. H. (2008). Gender differences in violent outcome and risk assessment in adolescent offenders after residential treatment. *International Journal of Forensic Mental Health*, *7*, 133–146.

Lodewijks, H. P. B., Doreleijers, T. A. H., & de Ruiter, C. (2008). SAVRY risk assessment in relation to sentencing and subsequent recidivism in a Dutch sample of violent juvenile offenders. *Criminal Justice and Behavior*, *35*, 696–709.

Lodewijks, H. P. B., Doreleijers, T. A. H., de Ruiter, C., & Borum, R. (2008). Predictive validity of the structured assessment of violence in youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry*, *31*, 263–271.

Manchak, S. M., Skeem, J. L., Douglas, K. S., & Siranosian, M. (2009). Does gender moderate the predictive utility of the level of service inventory-revised (ISA-R) for serious violent offenders? *Criminal Justice and Behavior*, *36*, 425–442.

Marczyk, G., Heilbrun, K., Lander, T., & DeMatteo, D. (2003). Predicting juvenile recidivism with the PCL: YV, MAYSI, and YLS/CMI. *International Journal of Forensic Mental Health*, *2*, 7–18.

McGowan, M. (2007). The predictive validity of violence risk assessment within educational settings. *Dissertation Abstracts International: Section A: Humanities and Social Sciences*, *68*(3A), 876.

Melton, G., Petrila, J., Poythress, N., Slobogin, C., Otto, R., Mossman, D., & Condie, L. (2018). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (4th ed.). New York: Guilford Press.

Meyers, J. R., & Schmidt, F. (2008). Predictive validity of the structured assessment for violence risk in youth (SAVRY) with juvenile offenders. *Criminal Justice and Behavior*, *35*, 344–355.

Michel, S., Riaz, M., Webster, C., Hart, S., Levander, S., Muller-Isberner, R., . . . Hodgins, S. (2013). Using the HCR-20 to predict aggressive behavior among men with schizophrenia living in the community: Accuracy of prediction, general and forensic settings, and dynamic risk factors. *International Journal of Forensic Mental Health*, *12*, 1–13. doi:10.1080/14999013.2012.760182

Monahan, J., & Skeem, J. (2016). Risk assessment in criminal sentencing. *Annual Review of Clinical Psychology*, *12*, 489–513.

Monahan, J., Steadman, H., Appelbaum, P., Grisso, T., Mulvey, E., Roth, L., . . . Silver, E. (2005). *Classification of Violence Risk: Professional manual*. Lutz, FL: PAR.

Monahan, J., Steadman, H., Robbins, P. C., Appelbaum, P., Banks, S., Grisso, T., . . . Silver, E. (2005). Prospective validation of the multiple iterative classification tree model of violence risk assessment. *Psychiatric Services*, *56*, 810–815.

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P. C., Mulvey, E., . . . Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, *62*, 783–792.

Mossman, D. (2007). Avoiding errors about "margins of error". *British Journal of Psychiatry*, *191*, 561.

Mossman, D. (2015). From group data to useful probabilities: The relevance of actuarial risk assessment in individual instances. *Journal of the American Academy of Psychiatry and the Law*, *43*, 93–102.

Neal, T. M., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior*, *42*, 952–968.

Neves, A. C., Goncalves, R. A., & Palma-Oliveira, J. M. (2011). Assessing risk for violent and general recidivism: A study of the HCR-20 and the PCL-R with a non-clinical sample of Portuguese offenders. *International Journal of Forensic Mental Health*, *10*, 137–149.

Nonstad, K., Nesset, M. B., Kroppan, E., Pedersen, T. W., Nöttestad, J. A., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the short-term assessment of risk and treatability (START) in a Norwegian high secure hospital. *International Journal of Forensic Mental Health*, *9*, 294–299.

Ogloff, J., & Daffern, M. (2006). The dynamic appraisal of situational aggression: An instrument to assess risk for imminent aggression in psychiatric inpatients. *Behavioral Sciences and the Law*, *24*, 799–813.

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*(4), 329–353.

Olver, M., & Wong, S. (2011). A comparison of static and dynamic assessment of sexual offender risk and need in a treatment context. *Criminal Justice and Behavior*, *38*, 113–126.

O'Shea, L., Picchioni, M., Mason, F., Sugarman, P., & Dickens, G. (2014). Differential predictive validity of the Historical, Clinical and Risk Management Scales (HCR–20) for inpatient aggression. *Psychiatry Research*, *220*, 669–678.

O'Shea, L. E., Thaker, D., Picchioni, M. M., Mason, F. L., Knight, C., & Dickens, G. L. (2016). Predictive validity of the HCR-20 for violent and non-violent sexual behaviour in a secure mental health service. *Criminal Behaviour and Mental Health*, *26*, 366–379. doi:10.1002/cbm.1967

Otto, R. K. (2000). Assessing and managing violence risk in outpatient settings. *Journal of Clinical Psychology*, *56*, 1239–1262.

Pedersen, L., Rasmussen, K., & Elsass, P. (2010). Risk assessment: The value of structured professional judgments. *International Journal of Forensic Mental Health*, *9*, 74–81.

Penney, S. R., Lee, Z., & Moretti, M. M. (2010). Gender differences in risk factors for violence: An examination of the predictive validity of the structured assessment of violence risk in youth. *Aggressive Behavior*, *36*, 390–404.

Penney, S. R., Marshall, L. A., & Simpson, A. I. F. (2016). The assessment of dynamic risk among forensic psychiatric patients transitioning to the community. *Law and Human Behavior*, *40*, 374–386.

Quinsey, V., Harris, G., Rice, M., & Cormier, C. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Quinsey, V., Harris, G., Rice, M., & Cormier, C. (2005). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.

Rajlic, G., & Gretton, H. (2010). An examination of two sexual recidivism risk measures in adolescent offenders. *Criminal Justice & Behavior*, *37*, 1066–1085. doi:10.1177/0093854810376354

Sada, A., Robles-García, R., Martínez-López, N., Hernández-Ramírez, R., Tovilla-Zarate, C. A., López-Munguía, F., . . . Fresán, A. (2016). Assessing the reliability, predictive and construct validity of historical, clinical and risk management-20 (HCR-20) in Mexican psychiatric inpatients. *Nordic Journal of Psychiatry*, *70*, 1–6.

Schaap, G., Lammers, S., & de Vogel, V. (2009). Risk assessment in female forensic psychiatric patients: A quasi-prospective study into the validity of the HCR-20 and PCL-R. *The Journal of Forensic Psychiatry & Psychology*, *20*, 354–365.

Schmidt, F., Campbell, M., & Houlding, C. (2011). Comparative analyses of the YLS/CMI, SAVRY, and PCL-YV in adolescent offenders: A 10-year follow-up into adulthood. *Youth Violence and Juvenile Justice*, *9*, 23–42.

Scurich, N., & Krauss, D. (2013). The effect of adjusted actuarial risk assessment on mock-jurors' decisions in a sexual predator commitment proceeding. *Jurimetrics*, *53*, 395–413.

Silver, E. (2001). *Mental illness and violence: The importance of neighborhood context*. El Paso, TX: LFB Scholarly Publishing.

Singh, J. P., Grann, M., & Fazel, S. A. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*, 499–513.

Sjöstedt, G., & Långström, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law*, *8*, 25–40.

Skeem, J., Schubert, C., Odgers, C., Mulvey, E., Gardner, W., & Lidz, W. (2006). Psychiatric symptoms and community violence among high-risk patients: A test of the relationship at the weekly level. *Journal of Consulting and Clinical Psychology*, *74*, 967–979.

Smith, P. H., Cullen, E. T., & Latessa, E. J. (2009). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders. *Criminology and Public Policy*, *8*(1), 1601–1626.

Steadman, H. (1982). A situational approach to violence. *International Journal of Law and Psychiatry*, *5*, 171–186.

Steadman, H., Mulvey, E., Monahan, J., Robbins, P. C., Appelbaum, P., Grisso, T., . . . Silver, E. (1998). Violence by people discharged from acute psychiatric facilities and by others in the same neighborhoods. *Archives of General Psychiatry*, *55*, 393–401.

Strub, D. S., Douglas, K. S., & Nicholls, T. L. (2014). The validity of version 3 of the HCR-20 violence risk assessment scheme amongst offenders and civil psychiatric patients. *International Journal of Forensic Mental Health*, *13*, 148–159.

van den Brink, R. H. S., Hooijschuur, A., van Os, T. W. D. P., Savenije, W., & Wiersma, D. (2010). Routine violence risk assessment in community forensic mental healthcare. *Behavioral Sciences & the Law*, *28*, 396–410.

Verbrugge, H. M., Goodman-Delahunty, J., & Frize, M. C. J. (2011). Risk assessment in intellectually disabled offenders: Validation of the suggested ID supplement to the HCR-20. *International Journal of Forensic Mental Health*, *10*, 83–91.

Viljoen, J., Mordell, S., & Beneteau, J. (2012). Prediction of adolescent sexual reoffending: A meta-analysis of the J-SOAP-II, ERASOR, J-SORRAT-II, and static-99. *Law and Human Behavior*, *36*, 423–438.

Viljoen, S., Nicholls, T. L., Roesch, R., Gagnon, N., Douglas, K., & Brink, J. (2016). Exploring gender differences in the utility of strength-based risk assessment measures. *International Journal of Forensic Mental Health*, *15*, 1–15.

Viljoen, J. L., Scalora, M., Cuadra, L., Bader, S., Chávez, V., Ullman, D., & Lawrence, L. (2008). Assessing risk for violence in adolescents who have sexually offended: A comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY. *Criminal Justice and Behavior*, *35*, 5–23.

Vincent, G. M., Chapman, J., & Cook, N. (2011). Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior*, *38*, 42–62.

Vitacco, M., Gonsalves, V., Tomony, J., Smith, B., & Lishner, D. (2012). Can standardized measures of risk predict inpatient violence? Combining static and dynamic variables to improve accuracy. *Criminal Justice and Behavior*, *39*, 589–606.

Walters, G. (2003). Predicting criminal justice outcomes with the psychopathy checklist and lifestyle criminality screening form: A meta-analytic comparison. *Behavioral Sciences & the Law*, *21*, 89–102.

Walters, G., White, T., & Denney, D. (1991). The lifestyle criminality screening form: Preliminary data. *Criminal Justice and Behavior*, *18*, 406–418.

Webster, C., Douglas, K., Eaves, D., & Hart, S. (1997). *HCR-20. Assessing risk for violence, version 2*. Burnaby, BC: Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Webster, C., Martin, M., Brink, J., Nicholls, T., & Desmarais, S. (2009). *Short-term assessment of risk and treatability (START): Clinical guide for evaluation risk and recovery* (Version 1.1). Hamilton, ON: St. Joseph's Healthcare.

Welsh, E., Bader, S., & Evans, S. E. (2013). Situational variables related to aggression in institutional settings. *Aggression and Violent Behavior*, *18*, 792–796.

Welsh, J., Schmidt, F., McKinnon, L., Chattha, H., & Meyers, J. (2008). A comparative study of adolescent risk assessment instruments: Predictive and incremental validity. *Assessment*, *15*, 104–115.

Wilson, C., Desmarais, S., Nicholls, T., Hart, S., & Brink, J. (2013). Predictive validity of dynamic factors: Assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior*, *37*, 377–388. doi:10.1037/lhb0000025

Yang, M., Wong, S., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*, 740–767.

Chapter 2

The Use of Measures of Psychopathy in Violence Risk Assessment

David DeMatteo, Emily Haney-Caron, Elyse Mowle, and John F. Edens

Introduction and Description of the PCL Measures

Current conceptualizations of psychopathy stem largely from Cleckley's (1941) seminal work, *The Mask of Sanity*, and the work of others around that time, including Karpman's (1946, 1948) descriptions of idiopathic versus symptomatic subtypes of psychopathy. Changing conceptualizations of psychopathy are reflected in the various iterations of the American Psychiatric Association's (APA) *Diagnostic and Statistical Manual of Mental Disorders* (DSM). The first two DSMs largely reflect Cleckley's conceptualization of the disorder, termed "sociopathic personality disturbance" in the first DSM (APA, 1952) and "personality disturbance, antisocial type" in DSM-II (APA, 1968). In DSM-III (APA, 1980), the disorder was renamed "antisocial personality disorder" and moved away from the personality traits described by Cleckley and toward more behaviorally based characteristics. Later DSMs made relatively minor changes to the primarily behaviorally oriented diagnostic features and continued to focus most heavily on criminal activity and other forms of antisocial conduct as important features. The APD criteria in DSM-5 were identical to those used in DSM-IV, and DSM-5 indicates that the disorder includes features "commonly included in traditional conceptions of psychopathy" (APA, 2013, p. 660). Additionally, one major change in the diagnosis of Conduct Disorder (CD) in the DSM-5 was the inclusion of a new specifier labeled "Limited Prosocial Emotions." This specifier can be applied to children and adolescents with CD who also show psychopathic-like characteristics (i.e., lack of remorse or guilt, callousness, shallow affect, and/or lack of concern about performance) and in many respects conceptually overlaps with the item content of Factor 1 of the Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003; described later in this chapter).

Perhaps in part because of dissatisfaction with DSM conceptualizations of psychopathy since DSM-III, since the 1980s the construct of psychopathy has been operationalized in many research and clinical contexts using instruments developed by Robert Hare and colleagues. These instruments include the Psychopathy Checklist (PCL; Hare, 1980), Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003), Psychopathy Checklist: Screening Version (PCL:SV; Hart, Cox, & Hare, 1995), and the Psychopathy Checklist: Youth Version (PCL:YV; Forth et al., 2003).

In contrast to recent DSM formulations of APD, the PCL measures place some emphasis on characteristics not overtly tied to antisocial conduct, such as interpersonal and affective features thought by many to be more representative of the disorder. There is substantial empirical evidence that the PCL-R operationalizes a construct important within the criminal justice system, but there are significant debates (e.g., Skeem & Cooke, 2010a) about the extent to which the PCL-R adequately operationalizes the core features of psychopathy. Due to space limitations, this chapter focuses primarily on PCL research relevant to violence risk assessment (for broader reviews, see Hare, 2003; Patrick, 2018; Richards et al., 2016; Skeem & Cooke, 2010a).

PCL-R and PCL:SV

The PCL-R is a 20-item construct rating scale used in research and clinical settings to assess psychopathy among adults. Standard administration involves a semi-structured interview and review of collateral data, although scoring the PCL-R without the interview is common in research settings. The PCL:SV is a 12-item tool derived from the PCL-R. PCL:SV Total scores correlate highly with PCL-R Total scores in offenders with and without intellectual disabilities (Pouls & Jeandarme, 2014), at least when used with forensic/correctional samples and scored based on the same information. Examiners rate each PCL item on a 3-point scale based on the degree to which the personality/behavior of the examinee matches the item description in the manual—0 (item does not apply to the individual), 1 (item applies to a certain extent), or 2 (item applies)—which results in scores ranging from 0 to 40 for the PCL-R (and 0 to 24 for the PCL:SV).

The PCL-R was normed on offender and forensic psychiatric samples but has been used with other populations, including community samples (e.g., DeMatteo, Heilbrun, & Marczyk, 2005, 2006; McGregor, Castle, & Dolan, 2012). Moreover, a large body of research suggests that the descriptive properties and correlates of the PCL-R are similar for male and female offenders (e.g., Gray & Snowden, 2016; see Hare, 2003, for a relevant review). Though some research found differences in item loadings and factor structure of the PCL-R for men and women (Dolan & Völlm, 2009), at least one study found adequate fit for both three- and four-factor models (Warren et al., 2003). However, psychopathic traits in women may be less predictive of violent recidivism (Weizmann-Henelius, Virkkunen, Gammelgård, Eronen, & Putkonen, 2015), though meta-analyses have produced somewhat inconsistent and conflicting results (Edens, Campbell, & Weir, 2007; Leistico, Salekin, DeCosta, & Rogers, 2008).

Research generally suggests PCL-R scores in the aggregate are not unduly influenced by the examinee's race or ethnicity (Skeem, Edens, Camp, & Colwell, 2004), with relatively similar descriptive properties and correlates of the PCL-R for Caucasian, African American, and Hispanic/Latino populations (Sullivan, Abramowitz, Lopez, & Kosson, 2006; Vachon, Lynam, Loeber, & Stouthamer-Loeber, 2012). That said, this research is limited (but growing), and at least one study found differences in the PCL-R's ability to predict violence among African American and Latino inmates relative to European American inmates (Walsh, 2013), and some have called for caution in using the PCL-R with diverse populations (e.g., Leistico et al., 2008). The predictive validity of PCL measures with various racial and ethnic groups is discussed in greater detail later in this chapter.

Early exploratory factor analyses of the PCL supported a two-factor solution, with Factor 1 primarily reflecting the interpersonal and affective characteristics of psychopathy and Factor 2 primarily reflecting antisocial and socially deviant characteristics (Harpur, Hare, & Hakstian, 1989). Later researchers proposed three-factor (e.g., Cooke & Michie, 2001; Hall, Benning, & Patrick, 2004) and four-factor models (e.g., Forth et al., 2003; Vitacco, Rogers, Neumann, Harrison, & Vincent, 2005). In early work, Cooke and Michie (2001) proposed a hierarchical three-factor model—reflecting interpersonal, affective, and lifestyle features—based on a confirmatory factor analysis of 13 PCL-R items. Studies replicated this three-factor model (e.g., Cooke, Michie, Hart, & Clark, 2004), but some researchers argue the superiority of a four-factor model that retains items explicitly tied to criminal history (e.g., Hare, 2003; Neumann, Hare, & Johansson, 2013). In recent years, the PCL-R has adopted the four-factor model that retains the antisocial factor (Hare, 2016). The first-order factors map onto a model consistent with the traditional two-factor model (Hare & Neumann, 2008); Factor 1 contains Interpersonal and Affective Factors and Factor 2 subsumes Lifestyle and Antisocial Factors.

Proponents of the three-factor model argue criminal behavior is "causally downstream" from the core features of the disorder and many factors other than psychopathy may cause such