

Colophon

Title	Sexual offender treatment for reducing recidivism among convicted sex offenders: a systematic review and meta-analysis
Institution	The Campbell Collaboration
Authors	Schmucker, Martin Lösel, Friedrich
DOI	10.4073/csr.2017.8
No. of pages	75
Last updated	28 July 2017

Citation	Schmucker, M, Lösel, F. Sexual offender treatment for reducing recidivism among convicted sex offenders: a systematic review and meta-analysis. Campbell Systematic Reviews 2017:8 DOI: 10.4073/csr.2017.8
ISSN	1891-1803
Copyright	© M. Schmucker & F. Lösel This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Roles and responsibilities	Martin Schmucker and Friedrich Lösel contributed to the writing and revising of this review. They will both be responsible for updating.
Editors for this review	Editors: David B. Wilson and Charlotte Gill Managing editor: Angela Higginson
Sources of support	No external funding was claimed for this review.
Declarations of interest	The authors are not engaged in the development or delivery of treatment programs for sexual offenders and are not affiliated with an institution providing such treatment. As independent researchers and experts, MS and FL are conducting evaluations of offender treatment programs. FL also serves on the Correctional Services Accreditation and Advice Panel of England and Wales and various advisory boards on treatment and prevention of offending in Germany. Both authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.

Corresponding author	Martin Schmucker Friedrich-Alexander-Universität Erlangen-Nürnberg Chair of Psychological Assessment Nägelsbachstr. 49c D-91052 Erlangen, Germany E-mail: martin.schmucker@fau.de
-----------------------------	--

Full list of author information is available at the end of the article.

Campbell Systematic Reviews

Editor-in-Chief Julia Littell, Bryn Mawr College, USA

Editors

Crime and Justice David B. Wilson, George Mason University, USA
Charlotte Gill, George Mason University, USA

Education Sandra Jo Wilson, Vanderbilt University, USA

International Birte Snilstveit, 3ie, UK
Development Hugh Waddington, 3ie, UK

Social Welfare Brandy Maynard, St Louis University, USA

Knowledge Translation Robyn Mildon, CEI, Australia
and Implementation Cindy Cai, AIR, USA

Methods Therese Pigott, Loyola University, USA
Ryan Williams, AIR, USA

Managing Editor Chui Hsia Yong, The Campbell Collaboration

Co-Chairs

Crime and Justice David B. Wilson, George Mason University, USA
Peter Neyroud, Cambridge University, UK

Education Sarah Miller, Queen's University, UK
Gary W. Ritter, University of Arkansas, USA

Social Welfare Mairead Furlong, National University of Ireland
Brandy Maynard, St Louis University, USA

Knowledge Translation Robyn Mildon, CEI, Australia
and Implementation Cindy Cai, AIR, USA

International Peter Tugwell, University of Ottawa, Canada
Development Hugh Waddington, 3ie, UK

Methods Ariel Aloe, University of Iowa, USA

The Campbell Collaboration was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. Campbell offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of Campbell's editors, librarians, methodologists and external peer reviewers contribute.

The Campbell Collaboration
P.O. Box 4404 Nydalen
0403 Oslo, Norway
www.campbellcollaboration.org

Table of contents

PLAIN LANGUAGE SUMMARY	5
EXECUTIVE SUMMARY/ABSTRACT	7
Background	7
Objectives	7
Search methods	7
Selection criteria	7
Data collection and analysis	7
Results	8
Authors' conclusions	8
BACKGROUND	9
OBJECTIVES	11
METHODS	12
Criteria for considering studies for this review	12
Search methods for identification of studies	13
Data collection and analysis	15
RESULTS	18
Description of studies	18
Synthesis of results	20
DISCUSSION	24
Overall results	24
Quality of the evidence and risks of bias	24
Results of further moderator analyses	26
AUTHORS' CONCLUSIONS	29
REFERENCES	30
References to included studies	30
Additional references	32
References to excluded studies	37
INFORMATION ABOUT THIS REVIEW	55
Review authors	55
Roles and responsibilities	56
Sources of support	56
Declarations of interest	56
Plans for updating the review	56
Author declaration	57

TABLES	58
Table 1: Characteristics of included comparisons	58
Table 2: Total mean effects and heterogeneity	63
Table 3: Moderator variables	64
FIGURES	67
Figure 1: PRISMA flow chart of the literature search and selection process	67
Figure 2: Forest plot of sexual recidivism outcomes	68
Figure 3: Forest plot of general recidivism outcomes	69
APPENDIX	70
Appendix 1: Relevant studies that appeared more recently	70
Appendix 2: Coding scheme	73

Plain language summary

Treatment of sexual offenders reduces reoffending, but more research needed to identify effective interventions

Treatment can reduce reoffending (recidivism) rates of sexual offenders. But the results of individual studies are too heterogeneous to draw a conclusion on the general effectiveness of sex offender treatment.

What is this review about?

Sexual offender treatment programs to reduce reoffending have been implemented in many countries as part of a strategy in managing this offender group. However, there are still controversies regarding their effectiveness.

This review integrates findings from six experimental and 21 quasi-experimental studies that compare groups of treated sexual offenders with equivalent control groups. These studies tested whether treated sexual offenders differed from the control groups in sexual and other reoffending.

What are the main findings of this review?

What studies are included?

Included studies compare official recidivism rates of treated sexual offenders with a comparable group of sexual offenders that have not been subjected to the respective treatment. Quasi-experimental studies were included only if they applied sound matching procedures, where the incidental assignment would not introduce bias, or where they were statistically controlled for potential biases. The treatment had to explicitly aim at reducing recidivism rates.

The review summarizes 27 studies containing 29 eligible comparisons of a treated group and a control group, containing data for 4,939 treated and 5,448 untreated sexual offenders. The studies come from seven different countries, but more than half of the studies have been carried out in North America. All eligible comparisons evaluated psychosocial treatment (mainly cognitive behavioral programs). No studies on pharmacological/hormonal treatment were found which meet the inclusion criteria.

Does treatment of sexual offenders reduce recidivism?

On average, there is a significant reduction in recidivism rates in the treated groups. The odds to sexually reoffend were 1.41 lower for treated compared to control groups. This equals a sexual recidivism rate of 10.1 percent for treated offenders compared to 13.7 percent without treatment. The mean rates for general recidivism were higher, but showed a similar reduction of roughly a quarter due to treatment.

The results from the individual studies were very heterogeneous, that is individual study features had a strong impact on the outcomes. Methodological quality did not significantly influence effect sizes.

Cognitive-behavioral as well as studies with small samples, medium to high risk offenders, more individualized treatment, and good descriptive validity revealed better effects. There was no significant difference between various settings. We found significant effects for treatment in the community and in forensic hospitals, but there is not yet sufficient evidence to draw conclusions regarding the effectiveness of sex offender treatment in prisons.

What do the findings of this review mean?

Overall, the findings are promising, but there is too much heterogeneity between the results of individual studies to draw a generally positive conclusion about the effectiveness of sex offender treatment. Applied cognitive-behavioral foundation of treatment has relatively good potential, but other features, like the risk of the treated offenders or including individualized treatment, significantly affect treatment success.

More well documented randomized trials and high-quality quasi-experiments are needed, particularly outside of North America. In addition, there is a clear need of more differentiated process and outcome evaluations.

How up-to-date is this review?

The study pool of the present analysis was based on the broad search of 2,039 documents from a review published in 2005, updated to cover studies issued prior to 2010. More recent studies were evaluated in an appendix and mostly showed similar findings as in our review. This Campbell Systematic Review was published in July 2017.

Executive summary/Abstract

Background

Sexual offender treatment programs to reduce reoffending have been implemented in many countries as part of a strategy in managing this offender group. However, there are still controversies regarding their effectiveness.

Objectives

A meta-analysis of relatively well-controlled outcome evaluations assessing the effects of treatment for male sexual offenders to reduce recidivism is conducted. The aim is to provide robust estimates of overall and differential treatment effects.

Search methods

We searched a broad range of literature databases, scanned previous reviews and primary studies on the topic, hand-searched 16 relevant journals, carried out an internet search of pertinent institutions, and personally contacted experts in the field of sex offender treatment. In total, we identified more than 3,000 documents that were scanned for eligibility.

Selection criteria

Studies had to address male sexual offenders and contain an outcome evaluation with a treated group (TG) and an equivalent control group (CG). Apart from randomized controlled trials (RCTs), also quasi-experimental designs were eligible if they applied sound matching procedures, statistically controlled for potential biases or the incidental assignment would not introduce bias. The studies had to evaluate therapeutic measures aiming at reducing recidivism. Both, psychosocial and organic treatment approaches were eligible. Case reports were not eligible and sample size had to be at least $n = 10$. To be eligible, studies had to report official recidivism data as an outcome and provide sufficient information for effect size calculation. There were no restrictions with regard to country of origin or language and both published and unpublished documents were eligible.

Data collection and analysis

For each study/comparison we coded general features, characteristics of the sample, treatment variables and methodological features. As most studies reported their results in terms of recidivism rates, we chose the odds ratio (OR) as effect size measure. If results on treatment dropouts were provided, we merged them with the treatment group results (“intent to treat” analysis). All statistical analysis of effect sizes applied a random effects model.

Results

29 comparisons drawn from 27 studies met our inclusion criteria. This study pool comprised 4,939 treated and 5,448 untreated offenders. A quarter of the studies were retrieved from unpublished sources. Most studies appeared since 2000 and more than half came from North America. The evaluations mostly addressed cognitive-behavioral sex offender treatment. No study on hormonal treatment met the inclusion criteria. Only about one fifth of the comparisons were RCTs and matching designs were rare as well. The follow-up periods ranged from 1 to 19.5 years ($M = 5.9$ years). Most frequently recidivism was defined as a new conviction and with only one exception studies presented data on sexual reoffending.

Overall, there was a positive, statistically significant effect of treatment on sexual reoffending ($OR = 1.41$, 95% CI: 1.11 to 1.78, $p < .01$). The mean effect equates to 26.3% less recidivism after treatment (sexual recidivism rate of 10.1% in treated sex offenders vs. 13.7% in the control groups). There was a comparable effect on general recidivism (26.4% less recidivism in treated groups; $OR = 1.45$, 95% CI: 1.15 to 1.83, $p < .01$). The overall effects were robust against outliers, but contained much heterogeneity.

Cognitive-behavioral programs showed a significant effect. Two RCTs on Multi-Systemic Therapy (MST) which also contains many cognitive-behavioral elements revealed a particularly large effect. Other intervention types showed weaker or no effects. There was a rather clear trend for better treatment effects of more individualized programs. There was no significant difference between various settings. We found significant effects for treatment in the community and in forensic hospitals, but there is not yet sufficient evidence to draw conclusions regarding the effectiveness of sex offender treatment in prisons.

The overall methodological quality of the studies was not significantly related to effect size. It should be noted, though, that we could not demonstrate a significant effect on sexual reoffending for the few RCTs in our study pool. Sample size was not linearly related to effect size but small studies with fewer than 50 participants had larger effects. This may suggest publication selection bias. However, studies from unpublished sources did not reveal weaker effects compared to published studies. The strongest methodological moderator was descriptive validity. Most studies lacked a detailed documentation of offender variables so that only few analyses could target this factor. With regard to offender characteristics we found no significant treatment effect for low risk participants. In contrast, medium and higher risk groups benefitted from treatment. Although the treatment of adolescents fared somewhat better than for adults, this difference was not significant. It made no difference whether offenders entered treatment voluntarily or on a mandatory basis.

Authors' conclusions

Overall, the findings are promising, but there is too much heterogeneity between the results of individual studies to draw a generally positive conclusion about the effectiveness of sex offender treatment. However, the results reveal information that is practically relevant: For example, our review confirms that cognitive-behavioral programs and multi-systemic approaches are more effective than other types of psychosocial interventions. The findings also suggest various conditions of success such as more individualization instead of fully standardized group programs, an advantage of treatment in the community or therapeutic settings instead of prisons, a focus on medium to high risk offenders, early treatment of young sexual offenders, and measures to ensure quality of implementation.

Overall, and particularly with regard to moderators, the research base on sex offender treatment is still not yet satisfactory. To enable more definite answers, more high-quality research is needed, particularly outside North America. There is a clear need of more differentiated process and sound outcome evaluations on various types of interventions (including pharmacological treatment), specific characteristics of programs, implementation, settings and participants and research methods.

Background

Sexual offending is a topic of particularly high concern in the general public, mass media and in crime policy making. Accordingly, many governments of industrialized countries have implemented not only more punitive measures but have also invested in treatment of sexual offenders to reduce recidivism. However, there is much controversy about the effectiveness of sex offender treatment, in particular with regard to methodological issues of evaluation (e.g. Marshall & Marshall, 2010; Rice & Harris, 2003; Seto et al., 2008). A general conclusion and consensus on ‘what works’ in this field is complicated by various issues (Lösel & Schmucker, 2017):

1. Sexual offending is a very heterogeneous category that contains, for example, various forms of child molesting, rape, exhibitionism, distribution and consumption of child pornography on the internet and other forms.
2. There are very different types of sexual offenders such as those with (or without) a deviant sexual preference (paraphilia), an antisocial personality, an opportunistic orientation, neuropsychological deficits, and so forth (Robertiello & Terry, 2007).
3. Although there is much research on static and dynamic risk factors for reoffending and structured risk assessment instruments (e.g. Hanson & Morton-Bourgon, 2009), the knowledge about the origins and causal mechanisms is less clear (e.g. Mann et al., 2010; Ward et al., 2005).
4. Treatment approaches are heterogeneous, ranging from psychosocial interventions such as cognitive-behavioral programs and relapse prevention or psychodynamic therapy to organic interventions such as hormonal treatment by medication or surgical castration, and some of these categories embrace rather different therapeutic measures in themselves (e.g. Marshall et al., 1998; McGrath et al., 2010).
5. Sound treatment evaluation is difficult because in various jurisdictions serious sexual offenders cannot simply be left untreated in control groups, the base rate of sexual recidivism is relatively low, and with regard to sexual reoffending longer follow-up periods are required compared to other fields of correctional intervention.

For such reasons controlled evaluations of programs for sexual offenders are less frequent than in general or violent offender treatment, particularly outside North America (Lipsey & Cullen, 2007; Lösel, 2012). However, over the last 20 years the number of studies has increased and more than a dozen systematic reviews or meta-analyses have been carried out (for overviews see Corabian et al., 2011; Lösel & Schmucker, 2014, 2017). Although there is overlap between these syntheses, they vary substantially with regard to the included primary studies, coding schemes, methods of effect size calculation and integration as well as the investigation of outcome moderators. Some meta-analyses concentrated on psychotherapeutic/psychosocial interventions only (e.g. Hanson et al., 2002), whereas others also included hormonal medication and surgical castration (Lösel & Schmucker, 2005). Within the category of psychotherapeutic/psychosocial interventions the specific treatment programs not only vary considerably but also share similarities. For example, the contents of cognitive-behavioral treatment (CBT), such as reducing deviant sexual attitudes, improving self-control, enhancing social skills, promoting perspective taking, or coping with stressors, overlap with those of relapse prevention programs that focus on avoidance or coping

with risk situations. Multi-systemic therapy for young sexual offenders and social therapy/therapeutic communities make particular use of the social context of the clients, but also incorporate elements from CBT, attachment and psychodynamic approaches. Hormonal treatment is used for subgroups of paraphilic offenders or others who are mainly motivated by sexual drive and not by dominance or other motivations. Medication is normally not used alone but accompanied by psychotherapy and other interventions. The available research syntheses also vary with regard to the countries of origin or language (e.g. most concentrated on reports in English), outcome criteria (e.g. reoffending versus other variables) and – in particular – methodological quality of the studies included.

Given this diversity of interventions, it is not surprising that the magnitude of treatment effects vary substantially (Lösel & Schmucker, 2014), although the two most comprehensive meta-analyses revealed similar results with regard to those types of treatment in which they overlapped (psychosocial interventions; Hanson et al., 2002; Lösel & Schmucker, 2005). However, due to the low number of high quality evaluations, i.e. randomized controlled trials (RCTs) or good quasi-experiments with equivalent control groups, the results of these reviews should not be seen as a definite answer to ‘what works in sexual offender treatment’ but rather as steps in a development to establish a sound evidence base. A good example for such a process is the review of Hanson et al. (2009) that showed that the Risk-Need-Responsivity (RNR) model of offender treatment (Andrews & Bonta, 2010) can be transferred from general to sexual offender treatment.

Objectives

Building on and updating our previous meta-analysis (Lösel & Schmucker, 2005) we meta-analytically integrate methodologically sound experimental and quasi-experimental studies that estimate the effects of treatment for sexual offenders on recidivism. This should provide the currently most valid international database on the effects of sexual offender treatment. Apart from the general question if treatment works for sexual offenders, the meta-analysis aims at analyzing characteristics that moderate treatment success.

Methods¹

Criteria for considering studies for this review

In order to be eligible for the meta-analysis, primary studies had to have the following characteristics:

1. *Study of male sexual offenders.* Participants had to have been convicted of a sexual offence or to have committed acts of illegal sexual behavior that would have led to a conviction if officially prosecuted. Studies on female sex offenders were not eligible. From the little that is known about female sex offending, we have to assume that it is not just a blueprint of its male counterpart (e.g. Freeman & Sandler, 2008).
2. *Evaluation of treatment.* No restrictions were made on the kind of intervention applied as long as it aimed to reduce recidivism (i.e., psychosocial as well as organic treatment modes such as hormonal medication by medroxyprogesterone or cyproterone acetate and surgical castration were eligible). However, interventions had to incorporate therapeutic measures; purely deterrent or punishing approaches were not included. Treatment did not have to be specifically tailored for sexual offenders. General offender treatment programs were eligible if the study addressed at least a subgroup of sexual offenders and reported separate results for these in both the treated and control groups.
3. *Study design.* The study had to report the same recidivism outcome for the treatment group (TG) and a control group (CG) not receiving the program under investigation. Apart from randomized studies, we included comparisons from quasi-experimental designs if there were no serious doubts regarding the equivalence of treatment and control groups. This included studies that used appropriate matching procedures, demonstrated equivalence by comparison of and/or statistical control for relevant variables. Equivalence was also assumed if the criteria of the incidental assignment did not relate to risks of reoffending such as availability of treatment in a certain region/at a certain time. These aspects were reflected in our adaptation of the Maryland Scientific Methods Scale (see Farrington et al., 2002). Level 3 or above had to be reached in order to be eligible. Our adaptation is slightly stricter and a little more differentiated at the upper end of the scale that is of special interest for the synthesis of methodologically sound studies. We used the following categories:
 - *Level 1: No control or comparison group.*
 - *Level 2: Nonequivalent comparison group.* Differences on relevant variables effecting recidivism are reported or are to be expected (e.g., treatment dropouts, subjects who refuse treatment).

¹ The methodology of this review is based on the published protocol (Lösel & Schmucker, 2009)

- *Level 3: Incidental assignment but equivalent control group.* No serious doubts that assignment resulted in equivalent groups, or sound statistical control of potential differences.²
- *Level 4: Matching procedures.* Systematic strategy to attain equivalence of the control group (e.g. theoretically sound matching or propensity score techniques).
- *Level 5: Random assignment of treated and untreated subjects.* This level also required absence of selective attrition (in case of selective attrition studies were downgraded or excluded depending on its severity).

CGs could consist of untreated offenders or offenders receiving “treatment as usual” or another kind of treatment that differed from the evaluated program in content, intensity and specificity. Waiting-list control groups were included if the design allowed testing of a program effect (see outcome measures).

4. *Measure of recidivism as outcome.* An indication of officially registered new offences had to be included as a dependent variable. Although recidivism is not a very sensitive indicator of treatment effects (e.g. Barbaree, 1997), it is politically and practically most relevant. We followed a broad definition of recidivism (sexual as well as non-sexual offences). Studies could use criteria such as arrest, charge, conviction or incarceration as long as these definitions drew on officially registered recidivism. In contrast, primary studies focusing exclusively on changes in measures of personality, attitudes, hormone levels, problem behaviors, clinical ratings of improvement, and the like were not included. Self-reported offending was also not included because of the severe risk of biased reporting (i.e. denial of offences).
5. *Sample size.* Studies had to contain a minimum total sample size of 10 persons with at least 5 offenders in each group. This also excluded case reports. We kept this criterion low because clinical studies on sex offenders are often small-scale.
6. *Sufficient data for effect size computation.* Studies had to report outcomes in a way permitting the calculation of effect size estimates.
7. *Country of origin.* No restrictions were made as to where studies were conducted. For economic reasons, we restricted our analysis to studies reported in English, German, French, Dutch, or Swedish language.
8. *Published and unpublished studies.* Published as well as unpublished studies were eligible. There were no restrictions regarding the time of publication.

Search methods for identification of studies

The study pool of the present analysis was based on the broad search of 2,039 documents that was reported in Lösel and Schmucker (2005) and updated to cover studies issued prior to 2010. Thus it concerned at least six more years of primary research than the previous meta-analysis.³ The coding was also updated for new information where necessary.

² As an example, the treatment program evaluated by Lab et al. (1993) was originally intended for low to medium risk youths only, but in fact the risk scores of CG youths assigned to non-specific programming did not substantially differ and the groups were also comparable on a number of other characteristics. We therefore assumed that potential outcome differences could be plausibly interpreted as representing treatment effects and rated it on Level 3.

³ Carrying out and publishing a comprehensive meta-analysis takes a lot of time. Therefore, trying to keep a review updated can create a vicious cycle that is in conflict with timely publication. We are aware of a few more recent studies that are not included in our review. We also know about two studies with large samples, however, after some waiting time the latter findings have not yet been released. Therefore, we felt that the current analysis should now be published. To check the robustness of our findings, we assessed the available more recent studies and found that they were generally in accordance with our main results. The respective studies are briefly reported in Appendix 1.

The search used as many sources as possible to achieve a comprehensive international study pool that included both published and unpublished evaluations (see Schmucker & Lösel, 2011). The sources included:

1.1.1 Searches in electronic literature and research databases

We searched multiple databases which tapped different academic subjects:

- C2-SPECTR
- Center for Sex Offender Management (CSOM) documents database
- Cochrane Library
- Dissertation Abstracts International
- ERIC
- KrimLit Beta II
- MedLine
- National Criminal Justice Reference Service (NCJRS)
- PAVNET Online
- PsycInfo
- Psyn dex
- Social Services Abstracts
- Sociological Abstracts
- UK National Health Service National Research Register

While such databases usually only cover published reports, some of the databases also refer to unpublished material. Usually the search combined four different keyword clusters: 1) (abnormal/delinquent) sexual behavior (e.g. sexual, paraphilia, molestation etc.); 2) criminal conduct and population (e.g. criminal, offenders, prison etc.); 3) therapeutic intervention (therapy, treatment, corrections etc.) and 4) outcome research (e.g. effectiveness, outcomes, recidivism etc.). The following terms were used to search four topical clusters:

Search cluster	Terms used
Cluster 1: (abnormal/delinquent) sexual behavior	sex* or paraphil* or rape or rapist or molest* or exhibitionis* or voyeur* or pedophil* or incest* or fetish* or necrophil* or frotteur*
Cluster 2: criminal conduct and population	offen* or crim* or delinquen* or perpetrator* or prison*
Cluster 3: therapeutic intervention	treat* or therapy or psychotherapy or intervention or training or correction* or rehabilitation or prevention or management
Cluster 4: outcome research	evaluation or evaluate or evaluated or outcome or outcomes or effect or effects or effectiveness or impact or recidivism or re-offen* or reoffen* or follow-up or followup or relapse

Search terms were individually adapted to the specific layout and search options the databases allowed for in order to construct manageable, but albeit comprehensive results.

1.1.2 Searching other resources

Apart from electronic databases our search used a number of additional sources to further enhance the comprehensiveness of our search:

- *Previous reviews* on sexual offender treatment were scanned for included studies.
- *Primary studies* were scanned for cross references (*snowball method*).

- *Handsearches of pertinent journals.* Available journals that are known to publish articles relevant to the topic at hand were searched manually. This search included 16 journals (e.g. Aggression and Violent Behavior; Criminal Behaviour and Mental Health; Journal of Sexual Aggression; Psychology, Crime & Law; Sexual Abuse).
- *Internet search.* We also conducted internet searches primarily to find unpublished material. Obviously, the internet cannot be searched in full as it constitutes a rather loosely organized pool of information (Schmucker & Lösel, 2011). We visited the internet representations of pertinent institutions (e.g. Departments of Corrections, Ministries of Justice), searched them for information on relevant studies and followed indications of such research until we could locate the referenced material.
- *Personal inquiries.* We personally contacted experts in the field of sexual offender research and asked for own or other studies that would contribute to our study pool.

1.1.3 Managing studies located incidentally

Sometimes relevant studies are found incidentally (e.g. in the course of another but related literature search; content alerts of journals and the like). There is the danger that incidentally located studies might bias the study pool depending on the special research interests or typically scanned sources. One might decide to drop such studies from the pool. On the other hand, the aim of a comprehensive review is to include all studies that are available. Our decision was to include such incidentally located studies but to document that they were identified in this way. However, we found that we had either located such studies by our systematic literature search as well or they did not meet the eligibility criteria on closer inspection.

Data collection and analysis

1.1.4 Data extraction and management

A broad range of variables were coded for descriptive purposes although not all relevant variables were reported in all reports. The coding of study characteristics followed a detailed coding manual that was extended from our previous meta-analysis (Lösel & Schmucker, 2005).

For each study/comparison general features (e.g. type and year of publication, country), characteristics of the sample (e.g. age, offence types, voluntariness of treatment participation, risk of reoffending), treatment variables (e.g. basic treatment concept, setting, format of the treatment) and methodological features (e.g. Maryland Scale rating, follow-up interval). Table 1 shows the main basic variables describing the pool of included comparisons.

All studies were independently coded by the first author and a trained member of our research team with experience in the field of offender treatment evaluation. Inter-rater agreement varied across the variables but was overall similar to our previous meta-analysis (Lösel & Schmucker, 2005). Especially important categories such as treatment type or quality of evaluation design reached nearly 100% and no variable was below 60%. Relatively low agreement was mostly due to discrepancies regarding the missing status of a variable (e.g. author affiliation was coded as missing more often by the second coder due to a lack of knowledge of affiliation networks for sex offender treatment specifically). In case of disagreement of the coders we had a group discussion in the research team to reach consensus.

1.1.5 Measures of treatment effect

Usually, the outcomes are reported in terms of recidivism rates for treated and untreated participants. We thus chose the odds ratio (*OR*) as effect size measure (see Fleiss et al., 1994). The following formulas were used for recidivism rates (*p*) and absolute number of successes and failures in the treated group (TG) and comparison group (CG) respectively:

$$OR = \frac{p_{CG} \times (1 - p_{TG})}{p_{TG} \times (1 - p_{CG})} \quad \text{and} \quad OR = \frac{TG_{Success} \times CG_{Failure}}{TG_{Failure} \times CG_{Success}}$$

If any of these frequencies equaled zero, 0.5 was added to each frequency. Some studies reported more sophisticated statistical analyses that controlled for differences between TGs and CGs. In such cases, we used these results instead of the simple recidivism rates. In logistic regression, the coefficients equal the natural log of the *OR* (*LOR*), and as an exponent to *e* this equals the *OR* (see Fleiss, 1994). The result for the treatment variable could thus be transferred directly. In Cox regression, results are reported in the form of a risk ratio, which is similar but not identical to the *OR*. We used the risk ratio (*RR*) to estimate a recidivism rate for the CG corrected for the group differences considered in the Cox regression model ($p_{CG} = RR \times p_{TG}$ or $p_{CG} = RR / p_{TG}$, depending on the coding of the treatment variable in the primary study). We then calculated the *OR* substituting the estimated CG recidivism rate following the above formula. Few studies reported other test statistics that could not be transformed readily into *ORs*. In these cases, we used standard procedures to calculate Cohen's *d* (see Lipsey & Wilson, 2001) and then converted these into odds ratios using

$$LOR = \frac{\pi}{\sqrt{3}} \times d \quad (\text{Hasselblad \& Hedges, 1995, Formula 4, re-arranged}) \quad \text{and} \quad OR = e^{LOR}.$$

Studies often reported multiple outcome variables. Different domains of recidivist behavior (i.e., sexual, violent, or general recidivism) were always analyzed separately. If a study reported different indicators of failure (i.e., charge, arrest, or conviction) for a common construct of interest, we would code effect sizes separately and then average them to a single effect size. In fact, this did not occur for any of the studies included in the final sample. To check whether different definitions of recidivism systematically relate to effect sizes, we subjected this to a moderator analysis and found no significant impact (see results section).

Some studies reported separate results for different offender types or risk groups, but did not meet criteria for independent comparisons as defined below. Here, we calculated effect sizes separately for the subgroups and used the weighted average to obtain a study effect size (see Fleiss, 1994).

Whenever possible, participants who dropped out of treatment were included in the treatment group ("intent to treat" analysis).

1.1.6 Unit of analysis issues

Sometimes references report more than one study. We then referred to the individual studies as the units of analysis. If a study contained multiple dependent (sub-)samples, we used the comparison with the highest internal validity. For example, if a study compared recidivism rates for the total sample of treated/untreated participants and additionally matched a subsample of these groups on relevant characteristics, we would use the latter comparison. Some primary studies present results for different independent subsamples (e.g. separated according to offence types). In those cases we used the subsamples as units of analysis when this would improve equivalence between treated and control groups and the report allowed for a differentiated coding of the individual subsamples regarding the coding variables (see below). Following this approach we extracted 29 comparisons from 27 studies that met our inclusion criteria. In total, the 29 comparisons comprise 4,939 treated and 5,448 untreated offenders.

1.1.7 Data synthesis and moderator analyses

Statistical analyses were conducted on the natural log of the *OR* (Fleiss, 1994; Lipsey & Wilson, 2001). To integrate effect sizes, we applied the weighting procedures based on the standard error of individual effect sizes (Hedges & Olkin, 1985). Analyses were conducted with IBM SPSS Statistics using the macros for meta-analysis written by David Wilson (see Lipsey & Wilson, 2001). Because of the expected heterogeneity of effect size distributions, we applied a random effects model. The random variance component (τ^2) was estimated via the method-of-moments procedure. All moderator analyses were carried out under the assumption of a mixed effects model (see also Lipsey & Wilson, 2001; Wilson, 2001). Continuous moderator variables were analyzed using weighted regression, again under the mixed effects model and weighted β coefficients (β_w) are reported. When only one moderator variable is in the regression model β_w equals the simple (weighted) correlation. Data were inspected for outliers and when necessary analyses were controlled for the presence of outliers and extreme values.

Results

Description of studies

1.1.8 Results of the search

Altogether, for the current integration more than 3,000 documents were identified. Figure 1 shows the PRISMA flow chart (Moher et al., 2009) of the literature search and selection process. After duplicates removed, 2851 documents were left for screening. Whenever titles or abstracts of located material did not clearly suggest that the study was ineligible we retrieved the full report to determine eligibility. 38 documents would have remained in the pool for further full-text screening but could not be located. Of the remaining 2813 documents 2373 were excluded because the title or the abstract clearly indicated that they were not eligible. Of the remaining 440 documents the full-texts were assessed for eligibility. At this stage another 413 documents were excluded. A number of these studies did not evaluate a concrete (sexual) offender treatment with respect to outcomes ($n = 143$). Of the remaining 270 studies, most had to be excluded because they did not incorporate adequate control groups.⁴ In the end only 27 studies met the inclusion criteria and comprised the study pool.

1.1.9 Characteristics of included studies

Table 1 contains an overview of the basic characteristics of included comparisons. They were predominantly reported in the last two decades. Nearly a half appeared since 2000 and only four studies were dated before 1990. Due to the lag between treatment and outcomes that is required in follow up studies the time of treatment implementation was often considerably earlier. Although our search identified eligible studies from seven countries, more than a half came from Canada and the USA. The majority of the comparisons were extracted from published journal articles. However, as mentioned, we took effort to include unpublished studies and these constituted almost one fourth of the pool.

Treatment characteristics. Most studies addressed the evaluation of cognitive-behavioral treatments (CBTs). Only eight programs were classified in other categories. In contrast to our previous meta-analysis (Lösel & Schmucker, 2005), no study on hormonal treatment met the more rigorous inclusion criteria of the present meta-analysis. However, some of the programs in the current pool applied additional medication in individual cases. Treatment took place in institutional as well as community settings and all but three programs were specifically designed for sexual offenders.

We coded whether treatment occurred in group and/or individual sessions on a five point scale. In most programs, treatment was solely ($k = 9$) or mainly ($k = 8$) carried out in a group format. Eight programs (27.6 %) contained predominantly individual sessions.

⁴ The reference section contains a list of these studies with the respective reasons for exclusion.

The duration of treatment ranged from a minimum of eight weeks to a maximum of 281 weeks ($M = 73.34$, $SD = 69.21$, $Median = 37.5$). Obviously, the treatment length differed between settings with outpatient treatment having the shortest durations ($M = 52.54$, $SD = 41.58$, $Median = 30.8$) and treatment in prison settings the longest ($M = 98.50$, $SD = 91.24$, $Median = 78.0$). The length of treatment could not be determined in 9 cases, i.e. almost one third of the studies did not provide information on a very basic variable.

Some features of the interventions were not well documented. Especially, coding of treatment integrity was rarely possible and if so this mostly meant that studies reported positive indicators of treatment integrity. Only one study (Hanson et al., 2004) reported problems in implementing the treatment, but in 18 studies (62.1%) there was simply no information on this aspect. It was also rarely reported whether aftercare services had been offered.

Offender characteristics. Regarding the age of the treated offenders, a majority of programs addressed adults only. However, this information could not always be extracted with sufficient certitude. The mean age of the treated offenders across all comparisons was 31.13 years ($SD = 7.97$). Usually the samples were rather homogeneous in age, but again this aspect was not always clearly reported.

With regard to sexual offending, nearly half of the programs and evaluations included mixed groups of rapists and child molesters ($k = 14$). Sometimes other sexual offenders also participated in the program ($k = 6$). Only one comparison referred to rapists and another one to exhibitionists exclusively. Seven comparisons only included child molesters and/or incest offenders. For eight comparisons no further information of offence type (apart from being sexual offenders) was available.

Meta-analyses on general offender treatment have shown that the risk of recidivism is negatively related to effect size (e.g. Lösel, 2012). Therefore, we tried to estimate the mean risk of treated offenders for each comparison. Mostly, there was no report on proper risk assessments in the studies. However, many studies reported some information on variables that are relevant for risk. We used the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997) to evaluate this information. The RRASOR was originally designed for individual risk judgments. We used the items of the RRASOR to estimate the mean risk for the treated group by translating group statistics of the relevant variables (information on prior convictions, age distribution, and victim characteristics in the study sample) into item scores and added them up to the total score. This was possible for 17 comparisons ($M = 1.98$; $SD = 0.63$ across comparisons). We then recoded these scores into three risk categories with low risk ranging to a score of 1.5 and high risk at a score of 2.5 or above. According to recidivism data reported by Hanson (1997) and Doren (2004) this renders a low risk group with estimated 5 year recidivism rates of roughly below 10%, a medium risk group with estimated 5 year recidivism rates between approximately 10% and 20%, and a higher risk group with estimated 5 year recidivism rates of about 20% and above. Three comparisons reported other risk assessments that could be grouped in these categories as well. Another four comparisons provided information that allowed an approximate risk classification. Table 1 shows the risk classification for those 24 comparisons. Five comparisons did not allow for any risk estimate. One might argue that our high risk category does not represent the offenders at very high risk and could be termed “elevated risk” or high-medium risk as this is done in some studies. However, our risk scores do not refer to individual offenders as in practical risk assessments, but are only used for a rough differentiation between groups as a whole. Against this background, we assume that the comparisons in our high risk category contain a substantial proportion of offenders at highest risk.

Methodological characteristics. Sample sizes ranged widely between a very small sample of 16 (Borduin et al., 1990) and a very large sample of 2,557 (Friendship et al., 2003). On average studies included 358 ($SD = 586.73$) offenders but in fact more than half of the comparisons (51.7 %) included fewer than 150 participants ($Median = 136$).

Only about one fifth of the comparisons were randomized controlled trials (RCTs) and studies with matching procedures to ensure equivalence of treated and untreated offenders were rare as well.

More than half of the comparisons drew on incidentally assigned samples (Level 3 on the Maryland Scale). Although sample characteristics and/or statistical control procedures justified the assumption of equivalence between treatment and control groups (see Methods section), the large proportion of Level 3 studies bears a risk of findings biased by weaker designs. Most studies had a rather long follow-up period. The mean time at risk ranged from 12 to 234 months with 24 comparisons (82.8 %) reporting follow ups of more than three years. On average the follow up time was 70.26 months or 5.9 years. Except for one study (Robinson, 1995) all reported sexual recidivism as an outcome. Most commonly, recidivism was defined as a new conviction but other definitions such as rearrest, new charges or reincarceration were used as well. Three studies integrated different indicators to establish whether a new offense had occurred or not.

We also coded what Lösel and Köferl (1989) introduced as “Descriptive Validity” (DV) of an evaluation (see also Farrington, 2006; Gill, 2011). This is not a characteristic of the study method itself but refers to the accuracy of information provided in a research report. Overall, there was often a lack of information and clarity about the treatment evaluated and details regarding the population and methods used. On a scale from 0 (very low) to 3 (excellent) the overall transparency was on average 1.21 ($SD = 0.68$). The descriptive validity was especially low for reporting on the actual implementation of the treatment at hand ($M = 0.48$; $SD = 0.69$) which points back to the high number of missing information regarding treatment integrity. For other areas the documentation was better, but not ideal (DV for “treatment concept”: $M = 1.41$; $SD = 0.91$; DV for “evaluation methods”: $M = 1.48$; $SD = 0.74$). Only outcome reporting had better values regarding DV ($M = 2.38$; $SD = 0.98$); however, this was due to our eligibility criteria as studies that did not allow for a reasonably accurate estimate of effect size were not included.

Synthesis of results

1.1.10 Total effects

Of the 29 comparisons included in the analyses, 28 reported on sexual recidivism outcomes (see Table 2). Figure 1 gives an overview of the *ORs* and confidence intervals for these comparisons as well as the overall mean. The forest plot shows considerable differences between effect sizes and this heterogeneity was significant; $Q(df = 27) = 52.05, p < .01$. According to Higgins et al.’s (2003) I^2 -measure nearly half of the observed heterogeneity cannot be attributed to sampling errors but represents systematic differences between the studies. Integration of the results according to a random effects model revealed a highly significant mean *OR* of 1.41 (95% CI: 1.11 to 1.78, $p = .005$). The treated offenders recidivated sexually at a mean rate of 10.1% (n -weighted average). The mean *OR* indicated that without treatment the recidivism rate would have been at 13.7%, i.e. treatment reduced recidivism by 3.6 percentage points or 26.3%.

Too few studies reported on violent ($k = 7$) or non-sexual recidivism ($k = 7$) to allow for adequate integration on these outcomes. However, 14 comparisons presented data on general recidivism (see Table 2 and Figure 2). As in sexual offending there was considerable and significant heterogeneity across outcomes in general recidivism; $Q(df = 12) = 23.66, p = .03$. The mean effect size was *OR* = 1.45 (95% CI: 1.15 to 1.83, $p = .002$). In terms of recidivism rates the n -weighted average in general reoffending for the treated groups was 32.6%. According to the estimated mean effect the respective rate is 41.2% without treatment. This is a reduction of 8.6 percentage points or 26.4% in general recidivism.

1.1.11 Sensitivity analyses: Exclusion of outliers

The forest plots of Figures 1 and 2 suggest that the significant heterogeneity might be due to outliers. In order to test the robustness of the effects we supplemented the calculation of the total effects with an analysis excluding extreme values. To identify outliers we drew on the procedure developed by Huffcutt and Arthur (1995) for meta-analysis. This takes into account the extremeness of a value (i.e. its deviation from the grand mean) as well as the respective sample size. For small samples larger deviations may be expected by chance, while for larger samples even

small deviations can be unlikely (i.e. “extreme”) and influence results considerably. For every study the Sample-Adjusted Meta-Analytic Deviancy (*SAMD*) statistic was calculated, both with respect to effects in sexual and general recidivism. For sexual recidivism three effects stood out of the other effect sizes (Borduin et al., 2009; Greenberg et al., 2000a; McGrath et al., 1998). Excluding those comparisons from the integration resulted in a marginally lower mean *OR* of 1.38 (95% CI: 1.12 to 1.70). This effect was still significant at $p = .003$. While the effect size distribution became more homogenous with the outliers excluded ($I^2 = 35.4\%$), it was still significantly heterogeneous; $Q (df = 24) = 37.18, p = .05$. For any recidivism one study showed an extreme value (Borduin et al., 2009). Excluding this reduced the total effect to $OR = 1.40$ (95% CI: 1.14 to 1.71). Again, the effect remained significant at $p = .001$, and heterogeneity was reduced, $I^2 = 32.7\%$, $Q (df = 12) = 17.83, p = .12$.

Overall, our sensitivity analysis showed that the mean effect sizes were relatively robust. The study of Borduin et al. (2009) on MST was an outlier in both outcome criteria and therefore should be considered with particular caution, but due to its relatively small sample size it had not much impact on the overall effect in our sensitivity analysis. As the effect size distribution for sexual recidivism remained heterogeneous, a more differentiated analysis of moderator effects was carried out.

1.1.12 Moderator analyses

The moderator analyses were based on a mixed effects model. Due to the rather small number of comparisons those analyses suffer from low statistical power. Nevertheless it seemed worthwhile to explore on variables that may systematically influence the results because this is relevant for a more detailed future development of sexual offender treatment. Table 3 gives an overview of methodological, offender and treatment variables and their impact on differences between study results.

1.1.12.1 Methodological variables

As we included studies that used different definitions for recidivism, we tested whether the recidivism measure used would be related to systematic outcome differences. At $Q (df = 4) = 2.94, p = .57$, there was no significant impact on study effect sizes and the heterogeneity of the effect size distribution was not reduced when applying this characteristic as a moderating variable.

Overall, design quality had no systematic effect on results. Neither the comparison between randomized and quasi-experimental designs nor the more differentiated distinction according to the Maryland Scale yielded any significant differences between mean effects ($p = .80$ and $p = .94$, respectively) and the correlation between study effect size and methodological quality was minuscule ($\beta_w = -.06, z = -0.34, p = .73$). However, the effect of treatment was statistically significant only for the designs at Level 3 of the Maryland Scale. For the few RCTs the effect was a bit smaller and not statistically significant. This reflects the low number of RCTs and hence the lower statistical power. High heterogeneity among randomized trials also contributed to lower statistical power, $Q (df = 4) = 14.39, p < .01$ (see also Figure 1). While the two randomized studies on MST of juvenile offenders (Borduin et al., 1990, 2005) showed untypically strong treatment effects, the remaining three RCTs revealed weak to even negative results (Marques et al., 2005; Ortmann, 2002; Romero & Williams, 1983).

Although general recidivism outcomes were not the target of our moderator analyses, it should be noted that these showed a different picture with regard to methodological quality (see Figure 2). Here, there was a significant difference between randomized and non-randomized designs, $Q (df = 1) = 5.91, p = .02$. RCTs had a strong treatment effect ($k = 4, OR = 3.46, 95\% CI: 1.63$ to $7.34, p = .001$), whereas quasi-experimental designs revealed no significant outcomes ($k = 10, OR = 1.30, 95\% CI: 0.98$ to $1.74, p = .07$). This reverse picture is obviously due to different subsets of primary studies. Those two randomized studies showing the worst outcomes for sexual recidivism (Marques et al. 2005; Romero & Williams, 1983) did not present data on general recidivism. Marques et al. reported findings on violent recidivism which showed even worse results ($OR = 0.64$). Therefore,

we assume that if all randomized studies had reported on general recidivism the effect would have been much smaller than mentioned above.

Recidivism base rate – defined as the mean recidivism rate in TG and CG – was an important moderator. The higher the rate of reoffending in a study sample, the larger the resultant effect sizes ($\beta_w = .39, z = 2.27, p = .02$). This is in fact closely related to the a priori risk of treated offenders with higher risk (see section on offender variables).

There were no systematic differences due to the length of follow up. However, two counteracting processes may be reflected in this variable. On the one hand, longer follow up periods are logically related to higher recidivism rates (in our sample the correlation was $r = .35$). Recidivism outcomes thus have a higher range in which effects can be demonstrated. On the other hand, the longer the follow up, the more other influences may have an impact in the life of a treated offender, thus supposedly reducing the effect of treatment. Following these thoughts we calculated a partial correlation between effect size and length of follow up with control for the recidivism base rate. It showed a clearer albeit still not significant negative trend ($\beta_w = -.27, z = -1.48, p = .14$; corrected for outliers: $\beta_w = -.39, z = -1.94, p = .052$).

Analyses on sample size also revealed complex results. There was only a small and non-significant linear relation to treatment effects with larger samples doing slightly worse ($\beta_w = -.05, z = -0.29, p = .77$). Eliminating the particularly large studies with $N > 1000$ (Duwe & Goldman; 2009; Friendship et al. 2003) raised the correlation which remained non-significant though ($\beta_w = -.19, z = -1.04, p = .30$). However, as Table 3 shows, there is one category that clearly stands out: Studies with small samples ($n \leq 50$) had very strong effects compared to all studies with larger samples ($p = .001$). Among the comparisons with larger samples there was no systematic relationship between sample and effect size ($\beta_w = .14, z = 0.67, p = .50$).

The strongest moderating effect in the methodological domain resulted for descriptive validity (DV, quality of reporting on the study). The 4-point scale rating of DV correlated with effect size at $\beta_w = .46, z = 2.78, p = .01$, indicating that unsatisfactory reports went along with worse outcomes. A closer inspection showed that this was mainly due to imprecise reporting on the treatment concept ($p = .01$) and the evaluation outcomes ($p = .02$). While the latter is probably related to conservative effect size estimation procedures, the former aspect points towards treatment integrity.

There was no difference in mean effects with regard to publication type, $Q(df = 2) = 2.59, p = .27$, or publication status, $Q(df = 1) = 0.01, p = .94$.

1.1.12.2 Treatment variables

The analyses on the treatment characteristics showed a significant effect for the general treatment concept applied. This is mainly a function of two evaluations on MST which were carried out by the program developers and had extremely large effects. Repeating the analyses on differences between the general treatment approach without those two studies revealed a non-significant result, $Q(df = 2) = 0.51, p = .78$. Of the remaining treatment approaches cognitive-behavioral programs showed a modest but significant effect on sexual recidivism. Other psychotherapeutic approaches did not yield a statistically significant treatment effect. This may be due to the low number of studies conducted on such therapies. The time of treatment implementation does not make a difference. There is no indication that treatment effects became larger in more recent time.

As Table 3 shows, there are only few treatment features that clearly differentiate effective treatment. This is in part due to the few comparisons available for moderator analyses and the low power of the respective tests. However, there are some other findings that deserve mentioning. For example, while there was no clear indication of effect size differences across different settings ($p = .16$), we only found significant effects for outpatient treatments and those provided in hospitals. Treatment in prison settings yielded a lower and non-significant mean effect.

The comparisons of specialized (versus non-specialized) sex offender treatment and authors' affiliation with the treatment both showed no significant moderator effect. But when testing the individual categories, only treatment tailored for sexual offenders and only evaluations conducted by authors affiliated with the treatment revealed significant mean effects in sexual recidivism.

There was a rather clear trend for better treatment effects of programs that have a more individualized approach ($\beta_w = .41, z = 2.47, p = .01$). In part this was due to the two trials on MST which represent a highly individualized approach. However, there remained a considerable tendency after exclusion of those studies ($\beta_w = .31, z = 1.67, p = .09$).

Treatment duration did not play a role regarding effect size; there was even a non-significant negative relation ($\beta_w = -.15, z = -0.72, p = .47$). Controlling for different settings, outliers, or offender risk did not substantially alter this picture.

1.1.12.3 Offender variables

Most studies lack a detailed description of offender variables or their analyses are not differentiated enough to allow for a detailed investigation of their impact on effect size. For example, we could not even perform a sensible analysis regarding the type of offence committed. Therefore, only three offender variables have been looked at in detail.

Regarding offender age, there was a significant treatment effect for both adults and adolescents. Although treatments that refer to adolescents fared somewhat better than those for adults, this difference was not significant ($p = .17$). If the analysis drew on the mean age of the treated participants, there was a tendency for younger groups benefiting more from treatment ($\beta_w = -.30, z = -1.80, p = .07$). However, this was mainly due to the two evaluations of MST that targeted adolescents. Excluding these, the age effect disappears ($\beta_w = -.11, z = -0.60, p = .55$). Another result refers to treatment recruitment. It made no difference whether offenders entered treatment voluntarily or on a mandatory basis ($OR = 1.33$ vs. $OR = 1.32$).

One of the strongest moderating effects is related to the risk of reoffending. The higher the risk, the higher the resulting treatment effect was. Treatments for low risk participants showed no effect at all. For the three risk categories there was a strong linear relationship ($\beta_w = .46, z = 2.59, p < .001$) and the results proved rather stable against outlier corrections. However, it must be noted that our risk classification is only a rough estimate and only three studies fitted into the highest category.

Discussion

Overall results

The above meta-analysis revealed a significant mean odds ratio of 1.41 for sexual recidivism. Only 10.1% of treated offenders reoffended whereas without treatment the recidivism rate would be 13.7%. That is a difference of 3.6 percentage points or 26.3%. For the more general outcome of any recidivism the mean effect was in the same range, even somewhat higher. Excluding outlier results only slightly reduced the mean effects and they remained significant, both for sexual and any recidivism. Thus, the total effects seem to be robust. Drawing on a sample of 29 rather well-controlled comparisons the results suggest that treatment can effectively reduce recidivism in sexual offenders.

The present mean effect in sexual recidivism is smaller than the one we found in our previous meta-analysis which included 80 comparisons ($OR = 1.70$; Lösel & Schmucker, 2005). However, the previous review contained many studies with nonequivalent comparison groups. It also incorporated studies on surgical castration and pharmacological treatment. Studies on surgical castration showed very large effect sizes but had various methodological shortcomings (apart from ethical and legal problems of the intervention itself). Excluding those studies the mean OR in our previous review was 1.38, and when the analyses were restricted to psychosocial interventions only it further decreased to $OR = 1.32$. As only psychosocial interventions fulfilled the stricter eligibility criteria in the current meta-analysis, the present mean effect is even a little stronger than in the previous meta-analysis.

Quality of the evidence and risks of bias

1.1.13 Study design

Although the overall results suggest a desirable mean effect of treatment this cannot be simply generalized because of the considerable heterogeneity in the findings of the primary studies. In addition, only six studies (five with sexual offending as outcome) were RCTs. Eight further studies at least used individual matching procedures to render equivalence between treatment and comparison groups. Although the effect size of those studies was in the same range as for the methodologically weaker studies, both the RCTs and the studies with an individual matching failed to yield statistical significance. In both cases this may be due to low statistical power (few studies and often only small sample sizes). The RCTs also showed very heterogeneous results, which further reduces statistical power. Obviously, there is no unambiguous trend in the best studies available. Accordingly, more RCTs are needed in order to get more valid data on the true effects of sexual offender treatment. On the other hand one should consider the arguments of Marshall and Marshall (2007) against a too narrow focus on RCTs in this field; for counter-arguments see Seto et al. (2008). An RCT that is not adequately designed to address the practice of psychotherapy may have limited value (e.g. Seligman & Levant, 1998; Hollin, 2008) and various threats to internal validity may also occur in RCTs (e.g. Lösel, 2007). In addition, RCTs for sexual offenders become increasingly difficult because various countries require mandatory treatment when the offence or sentence exceeds a specific level of seriousness. This often makes it impossible to form a randomized control group. Therefore, we suggest carrying out more RCTs on sexual offender treatment, but when an RTC is not feasible for legal or practical reasons one should also apply

sound quasi-experimental designs. Such a strategy has already been recommended in Campbell's (1969) groundbreaking article on program evaluation (see also Shadish et al., 2002).

The basic evaluation design was not a significant moderator in our meta-analysis. This is in contrast to findings in other fields of criminology (Weisburd et al., 2001), but not an exception in offender treatment research (Lipsey & Cullen, 2007; Lösel, 2012). In the present meta-analysis other features had a clearer influence on effect sizes.

1.1.14 Other risks of bias

Beyond the overall quality of the evaluation design our meta-analysis may contain various other risks of bias. For example, one third of the studies had only small sample sizes with up to 50 offenders. Those had higher effects than evaluations based on larger samples. This is usually regarded as a sign of publication bias. However, the difference in the present meta-analysis was not simply a function of an evaluation being published or not. Publication status did not exert an influence on effect size and the small sample effect was visible in published as well as unpublished studies. It is possible, though, that there is an "internal" publication bias, i.e. it may be more difficult to "hide" the results of a larger study. In contrast, the results of small scale studies may never be reported at all, not even as an unpublished report, especially if those results are negative and the researcher has a strong interest in not making the results visible. In fact, only one of the unpublished studies drew on a small sample (14%) compared to 30% among published studies. Thus, the results of this meta-analysis may be upwardly biased due to publication selection bias.

An alternative explanation of the small sample effect may be that treatment implementation is better monitored and easier controlled in a small scale setting. There are some other findings in our review that fit well with this implementation hypothesis: Evaluations that focused on only one program, implemented in one location revealed somewhat better results than studies that evaluated different programs across different institutions. Usually the latter indicates that program implementation was not well controlled (Greenberg et al., 2002; Riddijs & Timmermann, 2000) or that it was in fact weak (Hanson et al., 2004). Only two of the multi-location evaluations indicated a well-controlled implementation (Friendship et al., 2003; Guarino-Ghezzi & Kimball, 1998). Those two showed relatively good outcomes among the multi-location evaluations. Also, model projects that can be assumed to have a tight grip on program implementation fared slightly better than routine applications of treatment. This is in accordance with the literature on general offender treatment (Lösel, 2012) and also related to the issue of a potential influence of authorship.

The finding that only evaluations by authors affiliated to the program had a significant effect is in accordance with other criminological research (e.g. Eisner, 2009; Petrosino & Soydan, 2005). On the one hand this could be a matter of treatment integrity: It is likely that those who evaluate their own work pay more attention to proper program implementation. In fact, three quarters of the comparisons showing positive indicators of treatment integrity come from authors affiliated with the program in some way. On the other hand authors affiliated with the treatment may also be more reluctant to report negative results and may selectively analyze and publish favorable results. As we do not have detailed data on such processes, we only can alert the reader to such potential risks. For example, both studies on MST had a sound RCT design, but they were carried out by the program developers themselves and showed extremely large effect sizes (with Odds Ratios of about 20). However, the two studies also had other features that are connected with higher effects in the moderator analyses: They targeted young and rather high risk adolescent offenders, contained small samples and controlled for treatment integrity. Relatively positive results on MST have been reported in general offender treatment as well (Curtis et al., 2004). However, these are also predominantly studies by the program developers. In addition, the effects of MST were especially high in efficacy studies (demonstration projects) compared to effectiveness studies in real practice. When Littell et al. (2005) conducted a review on MST they drew a more skeptical picture because they only identified one fully independent evaluation and this showed no positive effect. Not surprisingly, Littell et al.'s critical conclusions have been challenged by Henggeler et al. (2006). However, independent evaluations of MST in Scandinavia have also shown contradicting effects (Ogden et al., 2007; Sundell et al. 2008).

In general there was no clear interaction between author affiliation and publication status in our data. But again, these results only refer to reports that were made available to us and there might be a “hidden” publication effect that goes beyond “officially published or not.” Overall, there was not enough valid information on treatment implementation and therefore this topic could not be properly tested.

1.1.15 Descriptive validity

Insufficient information in the documentation of details of the evaluation was very common in the current study set. This problem hinders more detailed moderator analyses and is in itself related to treatment effects. Studies that had more shortcomings in their reports showed lower effects than the better documented studies. The correlation between documentation quality (descriptive validity) and effect size can be tracked down to two aspects. First, it is a consequence of outcome reporting. Whenever possible effects were estimated for a comparison, but sometimes data had to be partially reconstructed from what was reported in a study. To ensure that the reconstruction would not overestimate the effects this was done in a conservative manner, so smaller effects in those comparisons could be expected. The second – and probably stronger – influence regarding the quality of documentation comes from the lack of detail on the treatment concept under consideration. The clearer a treatment concept was documented the higher the treatment effect. Again, this underlines the importance of treatment integrity. One can assume that in those cases that did not sufficiently report on the treatment, the concept may have been less elaborated or not properly implemented. Although this interpretation is somewhat speculative, the issue of descriptive validity should be seriously taken into account in future research.

The influence of methodological variables reduces the power to detect important content variables or may be confounded with such variables (Lipsey, 2003). Due to the limited number of available comparisons a meaningful statistical control for confounded variables was not possible in this meta-analysis. In spite of these limits, there are some moderating effects that deserve further attention.

Results of further moderator analyses

1.1.16 Treatment characteristics

Various treatment concepts that are used in practice were only represented by single studies or not at all. For example, no evaluation of pharmacological treatment fulfilled the eligibility criteria for our study pool. With regard to cyproterone acetate (CPA) or selective serotonin reuptake inhibitors (SSRIs) we found no controlled studies that examined their effectiveness on sexual offender recidivism. With regard to medroxyprogesterone acetate (MPA), there are at least some controlled studies. However, these evaluations mostly draw upon non-equivalent control groups and none of them fulfilled the criteria for the current review. To our knowledge there is one RCT on MPA treatment with sexual offenders (McConaghy et al., 1988). But with regard to the recidivism outcomes the randomized design is so severely disturbed that it renders the groups clearly non-equivalent. The RCT only holds for a less strict outcome criterion (“reduction in anomalous behavior”) that was not eligible for the present analysis. While other meta-analyses found favorable effects for hormonal medication (Hall, 1995; Lösel & Schmucker, 2005) these effects were based on relatively weak studies. It is therefore essential that the promising findings from previous meta-analyses be confirmed in evaluations with stronger research designs.

Only evaluations of psychosocial treatments met the inclusion criteria of this meta-analysis. Among the various therapeutic approaches one program stands out: The two evaluations of MST for juvenile sexual offenders showed extraordinarily strong effects and differed significantly from other approaches. The risk of a bias in these studies has been discussed above. Against this background the two MST studies on sexual offenders need replication in independent evaluations.

The majority of evaluations in the present study pool addressed cognitive-behavioral treatments (CBTs). Although CBT is not at all a homogeneous concept (Marshall & Marshall, 2010), there is a relatively broad study base to draw conclusions. The 20 comparisons evaluating sexual recidivism showed a significant, albeit moderate mean effect. This is in line with most of the previous meta-analyses on sexual offender treatment (e.g. Hall, 1995; Hanson et al., 2002; Lösel & Schmucker, 2005) and on general offender treatment (Landenberger & Lipsey, 2005; see also Lösel, 2012 and Wilson et al., 2005). Other approaches did not reach significant effects. In fact, there were only few evaluations of other treatment approaches that fulfilled the inclusion criteria. However, even among the CBT approaches the effects varied considerably and the only RCT on CBT that reports sexual recidivism outcomes (Marques et al., 2005) did not show a positive treatment effect. Although CBT approaches have been advocated over the last decades, the effects are not as clear cut as one might wish for “best practice”. It seems that the principal treatment approach in itself is not the clearest moderator and other variables may be more relevant for outcome differences.

Many of the treatment-related variables in the current meta-analysis did not provide clear cut differences between evaluations. However, there was a tendency that outpatient treatment fared better than treatment in prisons. The difference in favor of community programs is in agreement with the general research on ‘what works’ in correctional treatment (e.g. Andrews & Bonta, 2010; Koehler et al., 2013; Lipsey and Cullen, 2007; Lösel, 2012; Lösel & Koehler, 2014). This may be due to iatrogenic ‘contamination effects’ in the prison subculture, a lack of deterrence, a deferred transfer of learned contents to the world outside, difficulties during resettlement and other influences (Durlauf & Nagin, 2011; Gatti et al., 2009; Lösel et al., 2012; Markson et al., 2015). Our results on prison-based treatment are highly relevant for practice but they are difficult to interpret. Although there was no significant mean effect, prison-based programs did not fare significantly worse than treatment in other settings. Therefore, some issues of treatment context need to be emphasized: First, the primary studies did not directly compare treatment in prison vs. in the community, but TGs and CGs within the prison context. Second, institutionalized treatment in hospitals showed a significant effect on sexual reoffending. Third, one of the few primary studies in our pool that demonstrated a significant result was a prison-based CBT program (Duwe & Goldman, 2009: *OR* = 1.46). Therefore, it is necessary to investigate what program, organizational, relational and offender differences can make sexual offender treatment in prisons more promising.

One relevant issue may be the treatment format. In practice sexual offender treatment takes place in groups for the most part. In a thorough discussion Ware et al. (2009) provide plausible arguments for this approach. Not least, practical and financial reasons have to be considered. However, our findings suggest that the inclusion of individual sessions reveals better results. There may be confounding variables at work. For example, excluding the MST evaluations reduced the effect of individualization and the relation is probably not fully linear, i.e. a complete individualization may not be the golden principle either. However, it seems that supplementing group treatment with individualized sessions may better fit the responsivity principle of appropriate offender treatment (Andrews & Bonta, 2010). Unfortunately, there is no systematic research on the question whether an individualized or a group format is better for sexual offenders. However, there are various reasons for better effects of programs with individualized elements. First, some offenders may “hide” in group sessions. Second, using group sessions means that the same needs are targeted for all participants. This goes against the concept of individual needs and specific responsivity (Andrews et al., 2011), especially in mixed groups with very heterogeneous offender types. Third, supplemental individual sessions allow to tailor treatment more specifically (Drake & Ward, 2003) and to strengthen therapeutic alliances (Marshall et al., 2003; Ward & Maruna, 2007). Since general research on psychotherapy has clearly shown that relational issues and therapist characteristics are as important as the treatment model (e.g. Orlinsky et al., 1994), offender treatment needs to recognize that one size may not fit all (Lösel, 2012). Accordingly, treatment manuals should provide sufficient scope for flexibility and innovation (Marshall, 2009).

It would be desirable to more clearly disentangle the effect of the treatment format also for other variables; e.g. there is no controlled research on a fixed versus rolling format. Unfortunately, our study pool is too small to allow for analytical models enabling us to control for confounding variables in a more appropriate manner. In our previous meta-analysis that had less strict

inclusion criteria and thus a bigger study pool we could control for a number of other variables. As a consequence the impact of group versus individual treatment was less clear when we applied hierarchical regression analyses (Lösel & Schmucker, 2005). Therefore, at this stage we recommend to further investigate whether individualization is connected to better treatment outcomes or not. This kind of research should be related to analyses of the influence of other treatment process variables (see also Harkins & Beech, 2007; Pratt, 2010).

1.1.17 Offender characteristics

Regarding offender characteristics there is a trend for younger sex offenders to gain more from treatment. Again, this has to be interpreted with caution due to possible confounding variables. For example, younger offenders are also at higher risk for reoffending. Nevertheless, our findings indicate that early interventions in the career of sexual offenders are particularly worthwhile. The treatment of adolescent or young adult offenders can also benefit more from protective factors in the family or natural social context (Lösel & Bender, in press; Lösel & Farrington, 2012; Lösel et al. 2012).

The risk of reoffending was the strongest predictor of a positive treatment effect in the current analysis. The result of better effects in offenders at higher risk is in line with findings from general offender treatment (Lipsey & Cullen, 2007; Lösel, 2012). Hanson et al. (2009) applied the Risk-Need-Responsivity model to sexual offender treatment and found that programs were most effective when they fulfilled all three principles. The risk principle taken alone did not reach a significant result, but Hanson et al. rated the risk only dichotomously. Probably our differentiated risk rating led to more homogeneous categories and therefore better statistical power. However, as mentioned above, the category of 'high risk' in our review should be regarded cautiously because it does not mean that all of these offenders were at very high risk. For example, psychopathic offenders who would qualify as highest risk groups are particularly difficult to treat and often excluded from treatment programs (Lösel, 1998). At the other end of the risk level our findings suggest no significant effect. For offenders at low risk of reoffending the recidivism rate is so small that treatment cannot add much to further reduce reoffending.

Another variable deserves attention because it failed to produce a moderating effect: voluntary vs. non-voluntary treatment participation did not differ in their outcomes. Although the mean effect of studies with non-voluntary treatment was not significant this seems to be mainly a consequence of low statistical power (only six comparisons fell in that category). In fact, the mean effect is just the same as for voluntary treatment and in both categories the outcomes are highly heterogeneous. This means that a) offenders brought to treatment via external pressures such as judicial orders may benefit from treatment, and b) that voluntariness in itself is not a sufficient condition for successful treatment. Our finding points to the important role of change motivation as a process (e.g. Prochaska & Levesque, 2002) and techniques such as motivational interviewing (Miller & Rolnick, 2002). Unfortunately, treatment descriptions were not detailed enough to code and analyze this issue in more detail.

Authors' conclusions

Taken together the above analyses of reasonably well-controlled evaluations suggest that treatment of sexual offenders can be effective, but the results are not homogeneous. In particular, treatment in prisons and pure group formats seem to be less promising and require more differentiated research. Our findings are also supported by several more recent studies that were not included in this review (see Appendix 4). However, there is still a lack of high quality studies to unambiguously demonstrate treatment effectiveness. Future research must continue to evaluate sexual offender treatment in studies that use methodologically sound designs and are preferably independently authored and well documented. Good documentation is important because this is the key to a more thorough understanding of causal mechanisms in treatment practice. Due to the heterogeneity between primary studies, the investigation of outcome moderators needs much more attention. For example, although there is much research on the characteristics and subtypes of sexual offenders, this is rarely taken into account in treatment evaluation. In addition, we need more research on the processes of therapy with sexual offenders (Marshall & Burton, 2010) and focused tests of certain treatment features such as individualization, motivation and institutional context (Lösel, 2012). There are also too few evaluations that investigate recidivism not only as a dichotomous category but consider multiple criteria such as survival time, frequency and harm of the respective offences (e.g. Olver et al., 2012). Instead of sweeping controversies about the effectiveness of sex offender treatment more differentiated perspectives are needed (Koehler & Lösel, 2015). As it is common in other areas of psychotherapy and psychosocial intervention, research and practice should ask more frequently what works with whom, in what contexts, under what conditions, with regard to what outcomes, and also why. Although our review does not provide a definite answer to such differentiated questions, it suggests that sexual offender treatment has made progress towards an evidence-oriented crime policy.

References

References to included studies

[Independent comparisons drawn from the studies are reported in brackets]

- Bakker, L., Hudson, S. M., Wales, D. S., & Riley, D. (1998). *"And there was Light": Evaluating the Kia Marama Treatment Programme for New Zealand sex offenders against children*. Christchurch, NZ: Psychological Service, Department of Corrections.
- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, *34*, 105-113.
- Borduin, C. M., Schaeffer, C. M., & Heiblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: Effects on youth social ecology and criminal activity. *Journal of Consulting and Clinical Psychology*, *77*(1), 26-37.
- Duwe, G., & Goldman, R. A. (2009). The impact of prison-based treatment on sex offender recidivism: Evidence from Minnesota. *Sexual Abuse: A Journal of Research and Treatment*, *21*, 279-307.
- Friendship, C., Mann, R. E., & Beech, A. R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence*, *18*(7), 744-759.
- Greenberg, D. M., Da Silva, J.-A., & Loh, N. (2002). *Evaluation of the Western Australian Sex Offender Treatment Unit (1987-1999): A quantitative analysis*. Forensic Research Unit, Department of Psychiatry and Behavioural Sciences & Crime Research Centre: University of Western Australia. [Comparisons: (a) = Child molesters, (b) rapists]
- Guarino-Ghezzi, S., & Kimball, L. M. (1998). Juvenile sex offenders in treatment. *Corrections Management Quarterly*, *2*(1), 45-54.
- Hanson, R. K., Broom, I., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Sciences*, *36*(2), 87-96.
- Hanson, R. K., Steffy, R. A., & Gauthier, R. (1992). Long-term follow-up of child molesters: Risk predictors and treatment outcome.
- La Macaza Clinic (2002). *Criterion 8 - Program Follow-up and ongoing assessment*. Unpublished manuscript.
- Lab, S. P., Shields, G., & Schondel, C. (1993). Research note: An evaluation of juvenile sexual offender treatment. *Crime and Delinquency*, *39*(4), 543-553.

- Looman, J., Abracen, J., & Nicholaichuk, T. P. (2000). Recidivism among treated sexual offenders and matched controls: Data from the Regional Treatment Centre (Ontario). *Journal of Interpersonal Violence, 15*(3), 279-290.
- Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C., & von Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment, 17*, 79-107.
- Marshall, W. L., & Barbaree, H. E. (1988). The long-term evaluation of a behavioral treatment program for child molesters. *Behaviour Research and Therapy, 26*, 499-511. [Comparisons: (a) = Molesters of nonfamilial female children, (b) = Molesters of nonfamilial male children]
- Marshall, W. L., Eccles, A., & Barbaree, H. E. (1991). The treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. *Behaviour Research and Therapy, 26*, 129-135.
- McGrath, R. J., Hoke, S. E., & Vojtisek, J. E. (1998). Cognitive-behavioral treatment of sex offenders. A treatment comparison and long-term follow-up study. *Criminal Justice and Behavior, 25*, 203-225.
- Nicholaichuk, T. P. (1996). Sex offender treatment priority: An illustration of the risk/need principle. *Forum on Corrections Research, 8*(2), 38-41.
- Ortmann, R. (2002). *Sozialtherapie im Strafvollzug: Eine experimentelle Längsschnittstudie zu den Wirkungen von Strafvollzugsmaßnahmen auf Legal- und Sozialbewährung*. [Social therapy in prisons: An experimental longitudinal study on the effects of treatment in prisons on legal and social outcomes. Freiburg: Max-Planck-Institute for Foreign and International Criminal Law.
- Procter, E. (1996). A five-year outcome evaluation of a community-based treatment program for convicted sexual offenders run by the probation service. *Journal of Sexual Aggression, 2*, 3-16.
- Rice, M. E., Harris, G. T., & Quinsey, V. L. (1993). Evaluating treatment programs for child molesters. In J. Hudson & J. V. Roberts (Eds.), *Evaluating justice: Canadian policies and programs* (pp. 189-203). Toronto: Thompson.
- Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders* (Research Report No. R-41). Ottawa, Ontario, Canada: Correctional Service of Canada.
- Romero, J. J., & Williams, L. M. (1983). Group psychotherapy and intensive probation supervision with sex offenders: A comparative study. *Federal Probation, 47*(4), 36-42.
- Ruddijs, F., & Timmerman, H. (2000). The Stichting ambulante preventie projecten method: A comparative study of recidivism in first offenders in a Dutch outpatient setting. *International Journal of Offender Therapy and Comparative Criminology, 44*(6), 725-739.
- Schmid, P. (1988). *Was geschieht mit den Sexualstraftätern in der Psychiatrie? Darstellung und Bewertung der psychiatrischen Behandlung von Sexualstraftätern im Psychiatrischen Landeskrankenhaus Bad Schussenried in den Jahren 1978 - 1987*. [What happens with sexual offenders in psychiatry: Description and evaluation of the psychiatric treatment of sexual offenders in the psychiatric clinic of Bad Schussenried between 1978 and 1987]. Unpublished Dissertation, Universität Tübingen, Germany.

- Taylor, R. (2000). *A seven-year reconviction study of HMP Grendon Therapeutic Community* (Research Findings No. 115). London: Home Office Research and Statistics Directorate.
- Worling, J. R., & Curwen, T. (2000). Adolescent sexual offender recidivism: success of specialized treatment and implications for risk prediction. *Child Abuse & Neglect*, 24(7), 965-982.
- Ziethen, F. (2002). *Rückfallpräventive Effizienz der sozialtherapeutischen Behandlung von Sexualstraftätern. Evaluation der Sozialtherapie in der JVA Berlin-Tegel*. [The effects of social therapy on recidivism of sexual offenders: Evaluation of the social-therapeutic prison Berlin –Tegel]. Unpublished Diplomarbeit, Freie Universität Berlin, Germany.

Additional references

- Andrews, D.A., & Bonta, J. (2010). *The psychology of criminal conduct*, 5thed. Cincinnati, OH: Anderson.
- Andrews, D.A., Bonta, J., & Wormith, S. (2011). The Risk-Need-Responsivity (RNR) model: Does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behavior*, 38, 735-755.
- Barbaree, H.E. (1997). Evaluating treatment efficacy with sexual offenders: The insensitivity of recidivism studies to treatment effects. *Sexual Abuse: A Journal of Research and Treatment*, 9, 111-128.
- Campbell, D.T. (1969). Reforms as experiments. *American Psychologist*, 24, 409-429.
- Cooper, H.M., & Hedges, L.V. (Eds.). (1994). *The Handbook of research synthesis*. New York: Russell Sage Foundation.
- Corabian, P., Dennett, L., & Harstall, C. (2011). Treatment for convicted adult male sex offenders: an overview of systematic reviews. *Sexual Offender Treatment*, 6 (1), online journal.
- Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004). Multisystemic treatment: A meta-analysis of outcome studies. *Journal of Family Psychology*, 18(3), 411-419.
- Doren, D.M. (2004). Stability of the interpretative risk percentages for the RRASOR and Static-99. *Sexual Abuse: A Journal of Research and Treatment*, 16, 25-36.
- Drake, C.R., & Ward, T. (2003). Practical and theoretical rules for the formulation based treatment of sexual offenders. *International Journal of Forensic Psychology*, 1, 71-84.
- Durlauf, S.N., & Nagin, D. (2011). Imprisonment and crime: Can both be reduced? *Criminology and Public Policy*, 10, 13-54.
- Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology*, 5, 163-183.
- Farrington, D.P. (2006). Methodological quality and the evaluation of anticrime programs. *Journal of Experimental Criminology*, 2, 329-337.
- Farrington, D.P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In L.W. Sherman, D.P. Farrington, B.C. Welsh and D.L. MacKenzie (Eds.), *Evidence-based crime prevention* (pp. 13-21). London: Routledge.

- Fleiss, J. L. (1994). Measures of effect size for categorical data. In L. V. Hedges (Ed.), *The handbook of research synthesis* (pp. 245-260). New York: Russell Sage Foundation.
- Freeman, N.J., & Sandler, J.C. (2008). Female and male sex offenders: A comparison of recidivism patterns and risk factors. *Journal of Interpersonal Violence, 23*, 1394-413.
- Gatti, U., Tremblay, R.E., & Vitaro, F. (2009). Iatrogenic effects of juvenile justice. *Journal of Child Psychology and Psychiatry, 50*, 991-998.
- Gill, C.E. (2011). Missing links: How descriptive validity impacts the policy relevance of randomized controlled trials in criminology. *Journal of Experimental Criminology, 7*, 201-224.
- Hall, G.C.N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*, 802-809.
- Hanson, R.K. (1997). *The development of a brief actuarial scale for sexual offense recidivism*. Ottawa, Ont: Public Works and Government Services of Canada.
- Hanson, R. K., Broom, I., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Sciences, 36*, 87-96.
- Hanson, K., Burgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36*, 865-891.
- Hanson, R.K., Gordon, A., Harris, A.J.R., Marques, J.K., Murphy, W.D., Quinsey, V.L., & Seto, M.C. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 14*, 169-194.
- Hanson, R.K., & Morton-Bourgon, K.E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*, 1-21.
- Harkins, L., & Beech, A.R. (2007). A review of the factors that can influence the effectiveness of sexual offender treatment: Risk, need, responsivity, and process issues. *Aggression and Violent Behavior, 12*, 616-627.
- Hasselblad, V., & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin, 117*, 167-178.
- Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis*, Orlando, Academic Press.
- Henggeler, S.W., Schoenwald, S.K., Swenson, C.C., & Borduin, C.M. (2006). Methodological critique and meta-analysis as a Trojan horse. *Children and Youth Services Review, 28*, 447-457.
- Higgins, J. P. T., Simon, G. T., Deeks, J. J., & Altman, D. G. (2003). Measuring Inconsistency In Meta-Analyses. *BMJ: British Medical Journal, 327*(7414), 557-560.
- Hollin, C.R. (2008). Evaluating offending behaviour programmes: Does only randomization glister? *Criminology and Criminal Justice, 8*, 89-106.
- Huffcutt, A.I., & Arthur, W.J. (1995) Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology, 80*, 327-334.

- Koehler, J., & Lösel, F. (2015). A differentiated view on the effects of sex offender treatment. *British Medical Journal* (eLetter), <http://www.bmj.com/content/350/bmj.h199/rr-o>.
- Koehler, J.A., Lösel, F., Humphreys, D.K., & Akoensi, T.D. (2013). A systematic review and meta-analysis on the effects of young offender treatment programs in Europe. *Journal of Experimental Criminology*, 9, 19-43.
- Landenberger, N.A., and Lipsey, M.W. (2005). The positive effects of cognitive-behavioral programs for offenders: a meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1, 451-76.
- Lipsey, M.W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *Annals of the American Academy of Political and Social Science*, 587, 69-81.
- Lipsey, M.W., & Cullen, F.T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, 3, 297-320.
- Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Littell, J. H., Campbell, M., Green, S., & Toews, B. (2005). Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17. 2005, Issue 4. *Cochrane Database of Systematic Reviews*, Issue 4.
- Lösel, F. (1998). Treatment and management of psychopaths. In D.J. Cooke, A.E. Forth and R.B. Hare (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 303-354). Dordrecht, NL: Kluwer Academic Publishers.
- Lösel, F. (2007). Doing evaluation in criminology: Balancing scientific and practical demands. In R.D. King and E. Wincup (Eds.), *Doing research on crime and justice*, 2nd ed. (pp. 141-170). Oxford, UK: Oxford University Press.
- Lösel, F. (2012). Offender treatment and rehabilitation: What works? In M. Maguire, R. Morgan and R. Reiner (Eds.), *The Oxford handbook of criminology*, 5th ed. (pp. 986-1016). Oxford, UK: Oxford University Press.
- Lösel, F., & Bender, D. (in press). Protective factors against crime and violence in adolescence. In P. Sturmei (Ed.), *The Wiley handbook of violence and aggression*. New York: Wiley.
- Lösel, F., & Farrington, D.P. (2012). Direct protective and buffering protective factors in the development of youth violence. *American Journal of Preventive Medicine*, 43 (2S1), 8-23.
- Lösel, F., & Koehler, J. (2014). Can prisons reduce reoffending? A meta-evaluation of custodial and community treatment programs. *Presentation at the 14th Conference of the European Society of Criminology, 10-13 September 2014, Prague, CZ*.
- Lösel, F., & Köferl, P. (1989). Evaluation research on correctional treatment in West Germany: A meta-analysis. In H. Wegener, F. Lösel and J. Haisch (Eds.), *Criminal behavior and the justice system* (pp. 334-355). New York: Springer.
- Lösel, F., Pugh, G., Markson, L., Souza, K., & Lanskey, C. (2012). *Risk and protective factors in the resettlement of imprisoned fathers with their families*. Final research report. Norwich, UK: Ormiston Children and Families Trust.
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, 1, 117-146.

- Lösel, F., & Schmucker, M. (2009). *Protocol for a systematic review of the effects of sexual offender treatment*. The Campbell Collaboration library of systematic reviews, Retrieved from <http://campbellcollaboration.org/lib/project/53/>
- Lösel, F., & Schmucker, M. (2014). Treatment of sex offenders. In G. Bruinsma and D. Weisburd (Eds.), *Encyclopedia of Criminology and Criminal Justice* (pp. 5323-5332). New York: Springer.
- Lösel, F., & Schmucker, M. (2017). Treatment of sexual offenders: Concepts and empirical evaluations. In T. Sanders (Ed.), *The Oxford handbook on sex offences and sex offenders* (pp. 392-414). New York: Oxford University Press.
- Mann, R.E., Hanson, R.K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, 22, 191-217.
- Markson, L., Lösel, F., Souza, K., & Lanskey, C. (2015). Male prisoners' family relationships and resilience in resettlement. *Criminology and Criminal Justice*, in press, online; DOI: 10.1177/1748895814566287).
- Marshall, W.L. (2009). Manualization: A blessing or a curse? *Journal of Sexual Aggression*, 15, 109-120.
- Marshall, W.L., & Burton, D. (2010). The importance of therapeutic processes in offender treatment. *Aggression and Violent Behavior*, 15, 141-149.
- Marshall, W.L., Fernandez, Y.M., Hudson, S.M., & Ward, T. (Eds.), (1998). *Sourcebook of treatment programs for sexual offenders*. New York: Plenum Press.
- Marshall, W. L., & Marshall, L. E. (2007). The utility of the Random Controlled Trial for evaluating sexual offender treatment: The gold standard or an inappropriate strategy? *Sexual Abuse: A Journal of Research and Treatment*, 19, 175-191.
- Marshall, W.L., & Marshall, L.E. (2010). Can treatment be effective with sexual offenders or does it do harm? A response to Hanson (2010) and Rice (2010). *Sexual Offender Treatment*, 5 (2), online.
- Marshall, W. L., Serran, G. A., Fernandez, Y. M., Mulloy, R., Mann, R. E. & Thornton, D. (2003). Therapist characteristics in the treatment of sexual offenders: Tentative data on their relationship with indices of change. *Journal of Sexual Aggression*, 9, 25-30.
- McGrath, R.J., Cumming, G.F., Burchard, B.L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuse management: The Safer Society 2009 North American Survey*. Brandon, VT: The Safer Society Press.
- Miller, W., & Rollnick, S. (2002). *Motivational interviewing: preparing people for change*, 2nd ed. New York: Guilford.
- Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6): e1000097. doi:10.1371/journal.pmed1000097
- Petrosino, A., & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1, 435-50.

- Polizzi, D.M., MacKenzie, D.L., & Hickman, L.J. (1999). What works in adult sex offender treatment? A review of prison- and non-prison-based treatment programs. *Journal of Offender Therapy and Comparative Criminology*, 43, 357-374.
- Pratt, T.C. (2010). Meta-analysis in criminal justice and criminology: What it is, when it's useful, and what to watch out for. *Journal of Criminal Justice Education*, 21, 152-168.
- Prochaska, J.O., & Levesque, D.A. (2002). Enhancing motivation of offenders at each stage of change and phase of therapy. In M. Mc Murrin (Ed.), *Motivating offenders to change* (pp. 57-73). Chichester, UK: Wiley.
- Ogden, T., Hagen, K., & Andersen, O. (2007). Sustainability of the effectiveness of a programme of Multisystemic Treatment (MST) across participant groups in the second year of operation. *Journal of Children's Services*, 2, 4-14.
- Orlinsky, D.E., Grawe, K., & Parks, B.K. (1994). Process and outcome in psychotherapy. In A.E. Bergin and S.L. Garfield (Eds), *Handbook of psychotherapy and behavior change*, 4th ed. (pp. 270-376). New York: Wiley.
- Rice, M.E., & Harris, G.T. (2003). The size and signs of treatment effects in sex offender therapy. In R.A. Prentky, E.S. Janus and M.C. Seto (Eds.), *Annals of the New York Academy of Sciences* (pp. 428-440). New York: New York Academy of Sciences.
- Robertiello, G., & Terry, K.J. (2007). Can we profile sex offenders? A review of sex offender typologies. *Aggression and Violent Behavior*, 12, 508-518.
- Schmucker, M., & Lösel, F. (2011). Meta-analysis as a method of systematic reviews. In D. Gadd, S. Karstedt and S.F. Messner (Eds.), *The Sage handbook of criminological research methods* (pp. 425-443). Thousand Oaks, CA: Sage.
- Seligman, M. E., & Levant, R. F. (1998). Managed care policies rely on inadequate science. *Professional Psychology: Research and Practice*, 29(3), 211.
- Seto, M.C., Marques, J.K., Harris, G.T., Chaffin, M., Lalumière, M.L., Miner, M., Berliner, M. H., Rice, M.E., Lieb, R. and Quinsey, V.L. (2008). Good science and progress in sex offender treatment are intertwined. *Sexual Abuse: A Journal of Research and Treatment*, 20, 247-255.
- Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Sundell, K., Hansson, K., Löfholm, C.A., Olsson, T., et al. (2008). The transportability of Multisystemic Therapy to Sweden: Short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, 22, 550-560.
- Ward, T., & Maruna, S. (2007). *Rehabilitation: Beyond the risk-paradigm*. London: Routledge.
- Ward, T., Polaschek, D.L.L., & A. R. Beech, A.R. (2005). *Theories of sexual offending*. Chichester, UK: Wiley.
- Ware, J., Mann, R.E., & Wakeling, H.C. (2009). Group versus individual treatment: What is the best modality for treating sexual offenders? *Sexual Abuse in Australia and New Zealand*, 1, 70-79.

Weisburd, D., Lum, C.M., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *The Annals of the American Academy of Political and Social Science*, 578, 50-70.

Wilson, D.B. (2001). Meta-analytic methods for criminology. *Annals of the American Academy of Political and Social Science*, 578 (1), 71-89.

Wilson, D. B., Bouffard, L. A., & MacKenzie, D. L. (2005). A Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders. *Criminal Justice and Behavior*, 32(2), 172–204.

References to excluded studies

The following list presents those identified reports that were excluded after the full text was obtained. These reports presented an outcome evaluation of treatment for sexual offenders but did not meet the eligibility criteria of our systematic review. The references are marked by numbers which refer to the following reasons for exclusion:

- [1] No control group
- [2] No adequate control group (e.g. treatment dropouts)
- [3] Non-equivalent control group (i.e. rated at level 2 of our methods rating)
- [4] No official recidivism measure as outcome
- [5] Insufficient presentation of results
- [6] Complementary article to an included study

Abel, G. G., Mittelman, M., Becker, J. V., Rathner, J., & Rouleau, J. L. (1988). Predicting child molesters' response to treatment. In R. A. Prentky & V. L. Quinsey (Eds.), *Human sexual aggression: Current perspectives* (pp. 223-234). New York: New York Academy of Science. [1]

Adams, S. (2001). Specialized Sex Offender Probation in Cook County Links Supervision, Treatment *On Good Authority* (Vol. 4, pp. 1-4). [3]

Alaska Department of Corrections. (1996). Sex Offender Treatment Program: Initial Recidivism Study. Anchorage: Alaska Department of Corrections Offender Programs, Alaska Justice Statistical Analysis Unit, Justice Center and University of Alaska. [5]

Allam, J. (1998). Community-based treatment for sex offenders: An evaluation. Birmingham, UK: University of Birmingham. [1]

Allam, J. (1998). Effective practice in work with sex offenders: A reconviction study comparing treated and untreated offenders. Birmingham: West Midlands Probation Service Sex Offender Unit. [3]

Aytes, K. E., Olsen, S. S., Zakrajsek, T., Murray, P., & Ireson, R. (2001). Cognitive/behavioral treatment for sexual offenders: An examination of recidivism. *Sexual Abuse: A Journal of Research and Treatment*, 13(4), 223-231. [3]

Backer, J. (1990). *Nachuntersuchung kastrierter und im Maßregelvollzug untergebrachter Sexualstraftäter aus Westfalen-Lippe*. (Dissertation), Christian-Albrecht-Universität Kiel, Kiel. [1]

Bancroft, J., Tennent, G., Loucas, K., & Cass, J. (1974). The control of deviant sexual behaviour by drugs. *British Journal of Psychiatry*, 125, 310-315. [4]

- Barbaree, H. E., & Seto, M. C. (1998). The ongoing follow-up of sex offenders treated at the Warkworth Sexual Behaviour Clinic. Toronto: Centre for Addiction and Mental Health. [3]
- Barbaree, H. E., Seto, M. C., & Maric, A. (1996). Effective sex offender treatment: The Warkworth Sexual Behaviour Clinic. *Forum on Corrections Research*, 8(3). [3]
- Barbaree, H. E., Seto, M. C., & Maric, A. (1996). Sex offender characteristics, response to treatment, and correctional release decisions at the Warkworth Sexual Behavior Clinic. Toronto. [3]
- Barnes, J. M. (2000). *Recidivism in sex offenders: A follow-up comparison of treated and untreated sex offenders released to the community in Kentucky*. (Dissertation), University of Louisville, Louisville, KY. [3]
- Barnes, J. M., & Peterson, K. D. (1997). The Kentucky Sex Offender Treatment Program: Commonwealth of Kentucky, Justice Cabinet, Department of Corrections, Division of Mental Health. Kentucky State Reformatory, Kentucky. [3]
- Barry, D. J., & Ciccone, J. R. (1975). Use of depo-provera in the treatment of aggressive sexual offenders: preliminary report of three cases. *Bulletin of the American Academy of Psychiatry and the Law*, 3(3), 179-184. [1]
- Beech, A. R., Erikson, M., Friendship, C., & Ditchfield, J. (2001). A six-year follow-up of men going through probation-based sex offender treatment programmes *Findings* (Vol. 144). London: Home Office Research, Development and Statistics Directorate. [1]
- Beech, A. R., Fisher, D., & Beckett, R. C. (1999). Step 3: An evaluation of the Prison Sex Offender Treatment Programme. London: Home Office. [1, 4]
- Beech, A. R., Fisher, D., Beckett, R. C., & Scott-Fordham, A. (1998). An evaluation of the Prison Sex Offender Treatment Programme. London: Home Office. [4]
- Bench, L. L., Kramer, S. P., & Erickson, S. (1997). Discriminant Analysis of Predictive Factors in Sex Offender Recidivism. In B. K. Schwartz & H. R. Cellini (Eds.), *Sex Offender* (Vol. 2: New Insights, Treatment Innovations and Legal Developments, pp. 15-11-15-15). [2]
- Berlin, F. S., Hunt, W. P., Malin, H. M., Dyer, A., Lehne, G. K., & Dean, S. (1991). A five-year plus follow-up survey of criminal recidivism within a treated cohort of 406 pedophiles, 111 exhibitionists and 109 sexual aggressives: Issues and outcomes. *American Journal of Forensic Psychiatry*, 12(3), 5-28. [1]
- Berlin, F. S., & Meinecke, C. F. (1981). Treatment of sex offenders with antiandrogenic medication: conceptualization, review of treatment modalities, and preliminary findings. *American Journal of Psychiatry*, 138, 601-607. [2]
- Berliner, L., Schram, D., Miller, L. L., & Milloy, C. D. (1995). A sentencing alternative for sex offenders: A study of decision making and recidivism. *Journal of Interpersonal Violence*, 10(4), 487-502. [3]
- Berndtson, T. (1986). Androcur-Behandlung rückfallgefährdeter Vergewaltiger. In J. Heinrichs (Ed.), *Vergewaltigung. Die Opfer und die Täter* (pp. 116-117). Braunschweig: Holtzmeier. [1]
- Berner, W. (2001). Institutionelle Therapie bei sexueller Delinquenz. In V. Sigusch (Ed.), *Sexuelle Störungen und ihre Behandlung* (2. ed., pp. 501-516). Stuttgart: Thieme. [2]
- Berner, W., & Karlick-Bolten, E. (1986). *Verlaufsformen der Sexualkriminalität. 5-Jahres-Katamnesen bei Sexualdelinquenten unter Berücksichtigung von Frühsozialisation*,

vorausgegangener Delinquenz, psychiatrisch-psychologischer Diagnostik und Therapie. Stuttgart: Enke. [3]

- Bessler, C. (2002). Jugendliche Sexualstraftäter. Gruppentherapeutisches Behandlungsprogramm im Psychiatrisch-Psychologischen Dienst des Justizvollzugs des Kantons Zürich. [Juvenile sex offenders: Group therapy program in a forensic psychiatric-psychological service in Zurich]. *Kriminalistik*, 56(4), 263-268. [1]
- Bianchi, M. D. (1990). Fluoxetine Treatment of Exhibitionism. *American Journal of Psychiatry*, 147(8), 1089-1090. [1]
- Bingham, J. E., Turner, B. W., & Piotrowski, C. (1995). Treatment of sexual offenders in a outpatient community-based program. *Psychological Reports*, 76(3, Pt 2), 1195-1200. [1]
- Bluglass, R. (1980). Indecent exposure in the West Midlands. In D. West (Ed.), *Sex offenders in the criminal justice system* (pp. 171-180). Cambridge: Cambridge Institute of Criminology. [3]
- Böhme, A. (1935). *Psychotherapie und Kastration*. München: Lehmann. [1]
- Borduin, C. M., & Schaeffer, C. M. (2001). Multisystemic treatment of juvenile sexual offenders: A progress report. *Journal of Psychology and Human Sexuality*, 13(3/4). [6]
- Boulanger, H. (1985). *Die Sozial- und Legalbewährung von Sexualdelinquenten nach beantragter Kastration in Schleswig-Holstein*. (Dissertation), Universität Kiel, Kiel. [3]
- Bradford, J. M. W., & Pawlak, A. (1993). Double-blind placebo crossover study of cyproterone acetate in the treatment of the paraphilias. *Archives of Sexual Behavior*, 22, 383-402. [4]
- Bradford, J. M. W., & Pawlak, A. (1993). Effects of cyproterone acetate on sexual arousal patterns of pedophiles. *Archives of Sexual Behavior*, 22(6), 629-641. [1, 4]
- Brannon, J. M., & Troyer, R. (1995). Adolescent sex offenders: Investigating adult commitment-rates four years later. *International Journal of Offender Therapy and Comparative Criminology*, 39(4), 317-326. [1]
- Bremer, J. (1959). *Asexualisation. A follow-up study of 244 cases*. New York: MacMillan. [1]
- Briken, P., Berner, W., Noldus, J., Nika, E., & Michl, U. (2000). Therapie mit dem LHRH-Agonisten Leuprorelinacetat bei Paraphilien und sexuell aggressiven Impulshandlungen. *Nervenarzt*, 71(5), 380-385. [1]
- Bruinsma, F. (1993). Hilfen für die Täter. In G. Ramin (Ed.), *Inzest und sexueller Mißbrauch. Beratung und Therapie. Ein Handbuch* (pp. 355-395). Paderborn: Junfermann. [1, 4]
- Bulla, R. (1977). Verhaltenstherapie bei Sexualstraftätern. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 77-83). Stuttgart: Eigenverlag. [1, 4]
- Buresova, A., Bartova, D., & Svestka, J. (1990). Comparison of pharmacotherapeutic procedures in the treatment of sexual deviant behaviour. *Activitas Nervosa Superior*, 32(4), 299-301. [2]
- Byrne, S. M. (1999). *Treatment efficacy of a juvenile sexual offender treatment program*. (Dissertation), Memorial University of Newfoundland, St. Johns, Newfoundland, Canada. [3]

- Campbell, J. S., & Lerew, C. (2002). Juvenile sex offenders in diversion. *Sexual Abuse: A Journal of Research and Treatment*, 14(1), 1-17. [1]
- Carbo, R. A. (2002, October 30, 2002). [Effectiveness of the Sex Offender Accountability and Responsibility (SOAR) Program]. [1]
- Carter, D. L., & Prentky, R. A. (1993). Forensic treatment in the United States: A survey of selected forensic hospitals: Massachusetts Treatment Center. *International Journal of Law and Psychiatry*, 16(1-2), 117-132. [1]
- Clelland, S. R., Studer, L. H., & Reddon, J. R. (1998). Follow-up of rapists treated in a forensic psychiatric hospital. *Violence and Victims*, 13(1), 79-86. [2]
- Cook, D. A., Fox, C. A., Weaver, C. M., & Rooth, F. G. (1991). The Berkeley Group: Ten years' experience of a group for non-violent sex offenders. *British Journal of Psychiatry*, 158, 238-243. [2]
- Cooper, A. J., Sandhu, S., Loszтын, S., & Cernovsky, Z. (1992). A double-blind placebo controlled trial of medroxyprogesterone acetate with seven pedophiles. *Canadian Journal of Psychiatry*, 37, 687-693. [4]
- Cornu, F. (1973). *Katamnesen bei kastrierten Sittlichkeitsdelinquenten aus forensisch-psychiatrischer Sicht*. Basel: Karger. [3]
- Couture, J., Parker, T., Couture, R., & Laboucane, P. (2001). A Cost-Benefit Analysis of Hollow Water's Community Holistic Circle Healing Process (pp. ISBN-0-662-30035-30031, 30107). Canada; Ontario: Ministry of the Solicitor General, Ottawa (Ontario).; Aboriginal Healing Foundation, Ottawa (Ontario). [2]
- Craft, M. (1980). Case history: Sexual offenders treated with cyproterone acetate - a follow up evaluation. *British Journal of Sexual Medicine*, 7, 41 & 54. [1]
- Craissati, J., Falla, S., McClurg, G., & Beech, A. (2002). Risk, reconviction rates and pro-offending attitudes for child molesters in a complete geographical area of London. *Journal of Sexual Aggression*, 8(1), 22-38. [3]
- Craissati, J., & McClurg, G. (1997). The Challenge Project: A treatment program evaluation for perpetrators of child sexual abuse. *Child Abuse & Neglect*, 21(7), 637-648. [3]
- Crolley, J., Roys, D., Thyrer, B. A., & Bordnick, P. S. (1998). Evaluating outpatient behavior therapy of sex offenders: A pretest- posttest study. *Behavior Modification*, 22(4), 485-501. [1]
- Czerny, J.-P., Briken, P., & Berner, W. (2002). Antihormonal treatment of paraphilic patients in German forensic psychiatric clinics. *European Psychiatry*, 17(2), 104-106. [1, 4]
- Daniel, C. J. (1987). Shame aversion therapy and social skills training with an indecent exposor. In B. J. McGurk, D. M. Thornton & M. Williams (Eds.), *Applying psychology to imprisonment: Theory and practice* (pp. 247-254). London, England: Her Majesty's Stationery Office Books. [1]
- Davies, J. (1999, 13-15 October 1999). *Aboriginal Sex Offender Treatment Program Greenough Regional Prison*. Paper presented at the Best Practice Interventions in Corrections for Indigenous People Conference convened by the Australian Institute of Criminology in conjunction with the Department for Correctional Services SA, Adelaide, Australia. [2]

- Davies, T. S. (1974). Cyproterone acetate for male hypersexuality. *Journal of International Medical Research*, 2(2), 159-163. [1]
- Di Fazio, R., Abracen, J., & Looman, J. (2001). Group versus individual treatment of sex offenders: A comparison. *Forum on Corrections Research*, 13(1), 56-59. [3]
- Dieckmann, G. (1977). Die Behandlung von Sexualdelinquenten durch stereotaktische Hypothalamotomie. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 129-136). Stuttgart: Eigenverlag. [1]
- Dimnek, B., & Duncker, H. (1996). Zur Rückfallgefährdung durch Patienten des Maßregelvollzugs. *Recht und Psychiatrie*, 14, 50-56. [1]
- Dolde, G. (1996). Zur Bewährung der Sozialtherapie im Justizvollzug von Baden-Württemberg: Tendenzen aus einer neueren Rückfalluntersuchung. *Zeitschrift für Strafvollzug und Straffälligenhilfe*, 45, 290-297. [2]
- Donato, R., & Shanahan, M. (1999). The economics of implementing intensive in-prison sex-offender treatment programs. Canberra, Australia: Australian Institute of Criminology. [2]
- Dünel, F. (1979). Sozialtherapeutische Behandlung und Rückfälligkeit in Berlin-Tegel. *Monatschrift für Kriminologie und Strafrechtsreform*, 62(6), 322-337. [3]
- Dünel, F., & Geng, B. (1994). Rückfall und Bewährung von Karrieretätern nach Entlassung aus dem sozialtherapeutischen Behandlungsvollzug und dem Regelvollzug. In M. Steller, K.-P. Dahle & M. Basqué (Eds.), *Straftäterbehandlung: Argumente für eine Revitalisierung in Forschung und Praxis* (pp. 35-74). Pfaffenweiler: Centaurus. [3]
- Eastman, B. J. (1997). *An investigation of the efficacy of adolescent sex offender treatment*. (Dissertation), Virginia Commonwealth University, Richmond. [2]
- Eher, R., Dwyer, M., Prinoth, S., Fruhwald, S., & Gutierrez, K. (1997). Sexualstraftäter im Maßregelvollzug und deren Angehörige: Ergebnisse gemeinsamer Therapiesitzungen. *Psychiatrische Praxis*, 24(4), 190-195. [1, 4]
- Eher, R., Dwyer, M., Prinoth, S., Wagner, E., Fruhwald, S., & Gutierrez, K. (1997). Sexualstraftäter im Massregelvollzug und deren Angehörige: Ergebnisse gemeinsamer Therapiesitzungen. *Imprisoned sex offenders and their relatives: Results of common treatment sessions*. *Psychiatrische Praxis*, 24(4), 190-195. [1, 4]
- Eher, R., Lackinger, F., Frühwald, S., & Frottier, P. (2006). Beziehungsorientierte Psychotherapie bei entlassenen Straftätern und einer Gruppe von Sexualstraftätern--Ergebnisse einer 7-Jahreskatamnese. [Relationship oriented psychotherapy with released offenders and a group of sexual offenders - Results of a 7 year follow up]. *Recht & Psychiatrie*, 24(2), 83-90. [2]
- Eibl, E. (1977). Bericht über 51/2-jährige Erfahrung in der Behandlung und Überwachung von ca. 300 Sexualtätern in und außerhalb des Maßregelvollzuges unter besonderer Berücksichtigung der Penis-Plethysmographie (PPG). In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 59-64). Stuttgart: Eigenverlag. [2]

- Elsner, K. (2001). Gruppenbehandlung von Sexualstraftätern im Maßregelvollzug. [Group treatment of sex offenders in forensic psychiatric institutions]. In J. Hoyer & H. Kunst (Eds.), *Psychische Störungen bei Sexualdelinquenten* (pp. 153-181). Lengerich: Pabst. [1, 4]
- Emmanuel, N. P., Lydiard, R. B., & Ballenger, J. C. (1991). Fluoxetine treatment of voyeurism. *American Journal of Psychiatry*, *148*(7), 950. [1]
- Emory, L. E., Cole, C. M., & Meyer, W. J. (1992). The Texas experience with DepoProvera: 1980-1990. *Journal of Offender Rehabilitation*, *18*(3-4), 125-139. [3]
- Fahndrich, E. (1974). Cyproteronacetat in der Behandlung von Sexualdeviationen bei Männern. *Deutsche Medizinische Wochenschrift*, *99*(6), 234-2378 passim. [1, 4]
- Federoff, J. P., Wisner-Carlson, R., Dean, S., & Berlin, F. S. (1992). Medroxy-progesterone acetate in the treatment of paraphilic sexual disorders: Rate of relapse in paraphilic men treated in long-term group psychotherapy with or without medroxy-progesterone acetate. *Journal of Offender Rehabilitation*, *18*, 109-123. [3]
- Fehlenberg, D. (1997). Zur Psychotherapie paraphiler Sexualstraftäter im Maßregelvollzug. *Recht und Psychiatrie*, *15*, 159-167. [1]
- Field, L. H. (1973). Benperidol in the treatment of sexual offenders. *Medicine, Science and the Law*, *13*(3), 195-196. [1]
- Flowers, J. V. (1979). The differential outcome effects of simple advice, alternatives and instructions in group psychotherapy. *International Journal of Group Psychotherapy*, *29*(3), 305-316. [4]
- Fransblow, J., & Smiley, W. C. (1996). *Rational emotive behavior therapy applications to the treatment of sex offenders and violent offenders*. Paper presented at the International Congress of Psychologists, Montreal, CAN. [2]
- Frenken, J. (1994). Treatment of Incest Perpetrators: A Five-Phase Model. *Child Abuse and Neglect*, *18*(4), 357-365. [1]
- Friendship, C., Mann, R. E., & Beech, A. R. (2003). Evaluation of the prison-based Sex Offender Treatment Programme (SOTP). London: Home Office Research and Statistics Directorate. [6]
- Frisbie, L. V. (1969). Another look at sex offenders in California. Sacramento, CA: State of California, Department of Mental Hygiene. [3]
- Gabbert, T. (1987). Rehabilitationsergebnisse mit offenem und teilstationärem Maßregelvollzug. *Forensia*, *8*, 81-89. [1, 4]
- Gagné, P. (1981). Treatment of sex offenders with medroxyprogesterone acetate. *American Journal of Psychiatry*, *138*, 644-646. [1, 4]
- Glander, H. J. (1981). Die antiandrogene Therapie von Sexualdeviationen aus andrologischer Sicht. *Das Deutsche Gesundheitswesen*, *36*, 1873-1878. [2]
- Gordon, A. (1989). Research on sex offenders: Regional Psychiatric Centre. *Forum on Corrections Research*, *1*, 20-21. [1]
- Gordon, A., Holden, R., & Leis, T. (1991). Managing and treating sex offenders: Matching risk and needs with programming. *Forum on Corrections Research*, *3*(4). [3]

- Gordon, A., & Nicholaichuk, T. (1996). Applying the risk principle to sex offender treatment. *Forum on Corrections Research*, 8. [3]
- Goudsmit, W., & Reicher, J. W. (1980). Sozialtherapie schwerstgestörter Delinquenten auf psychoanalytischer Grundlage. In V. Sigusch (Ed.), *Therapie sexueller Störungen* (2. neubearbeitete und erweiterte ed.). Thieme: Stuttgart. [2]
- Greenberg, D. M., Bradford, J. M., Curry, S., & O'Rourke, A. (1996). A comparison of treatment of paraphilias with three serotonin reuptake inhibitors: A retrospective study. *Bulletin of the American Academy of Psychiatry and the Law*, 24(4), 525-532. [2, 4]
- Gretenkord, L. (1981). Mehrdimensionale Therapie eines Sexualdelinquenten in einer forensischen Klinik. *Monatschrift für Kriminologie und Strafrechtsreform*, 64(6), 353-360. [1]
- Gretenkord, L. (1994). Gewalttaten nach Maßregelvollzug (§ 63 StGB). In M. Steller, K.-P. Dahle & M. Basqué (Eds.), *Straftäterbehandlung: Argumente für eine Revitalisierung in Forschung und Praxis* (pp. 75-89). Pfaffenweiler: Centaurus. [1]
- Groth, N. (1983). Treatment of the sexual offender in a correctional institution. In J. G. Greer & I. R. Stewart (Eds.), *The sexual aggressor* (pp. 160-176). New York: Van Nostrand Reinbold. [3]
- Gutierrez Lobos, K., Eher, R., Gruenhut, C., Holzinger, A., & Bankier, B. (1999). Soziales Netzwerk und die Erfuellung sozialer Bedürfnisse bei Sexualstraftaetern im Massregelvollzug Social networks and social support in a sample of mentally ill sex offenders. *Psychiatrische Praxis*, 26(2), 85-88. [1, 4]
- Gutierrez-Lobos, K., Eher, R., Grunhut, C., Holzinger, A., & Bankier, B. (1999). Soziales Netzwerk und die Erfüllung sozialer Bedürfnisse bei Sexualstraftätern im Maßregelvollzug. *Psychiatrische Praxis*, 26(2), 85-88. [1, 4]
- Hagan, M. P., & Gust-Brey, K. L. (1999). A ten-year longitudinal study of adolescent rapists upon return to the community. *International Journal of Offender Therapy and Comparative Criminology*, 43(4), 448-458. [1]
- Hagan, M. P., King, R. P., & Patros, R. L. (1994). The efficacy of a serious sex offenders treatment program for adolescent rapists. *International Journal of Offender Therapy and Comparative Criminology*, 38(2), 141-150. [1]
- Hall, G. C. N. (1995). The preliminary development of a theory-based community treatment for sexual offenders. *Professional Psychology: Research and Practice*, 26(5), 478-483. [3]
- Hansen, H., & Lykke-Olsen, L. (1997). Treatment of dangerous sexual offenders in Denmark. *Journal of Forensic Psychiatry*, 8(1), 195-199. [2]
- Hanson, R. K., & Nicholaichuk, T. P. (2000). A cautionary note regarding Nicholaichuk et al. (2000). *Sexual Abuse: A Journal of Research and Treatment*, 12(4), 289-293. [3]
- Hanson, R. K., Steffy, R. A., & Gauthier, R. (1993). Long-term recidivism of child molesters. *Journal of Consulting and Clinical Psychology*, 66, 646-652. [6]
- Hanstein, W. (1997). Therapieabbruch und Rückfälle bei Sexualstraftätern. In K. Klees & W. Friedebach (Eds.), *Hilfen für missbrauchte Kinder. Interventionsansätze im Überblick* (pp. 229-240). Weinheim: Beltz. [1]
- Harsch, S. (2000). Vergleichende Studie von Sexualstraftätern im Strafvollzug und in der forensischen Psychiatrie. In A. Eckert, M. Wolfersdorf, G. Ropers, E. Spancken & R. Szczyzny

(Eds.), *Forschung in Psychiatrischen Krankenhäusern Beiträge 1. Postersymposium Arbeitsgruppe Forschung der Bundesdirektorenkonferenz 1999 in Bedburg-Hau*. Regensburg: Roderer. [1]

Hauch, M., Galedary, G., Haag, A., Lohse, H., & Schorsch, E. (1984). Ambulante psychotherapeutische Behandlung von Sexualstraftätern. *Sexualpädagogik und Familienplanung*, 12(6), 8-9. [1, 4]

Hebestreit, J. (1977). Psychologische Nach- und Voruntersuchungen -- Testpsychologische Befunde bei Hypothalatomie. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 167-171). Stuttgart: Eigenverlag. [1, 4]

Hedderman, C., & Sugg, D. (1996). Does treating sex offenders reduce offending? London. [3]

Hedlund, E. (1993). Gruppterapi med man som begatt sexbrott. Inga återfall rapporteras i uppföljning. [Group therapy for male sex offenders. No recurrence is reported in a follow-up]. *Läkartidningen*, 90(5), 379-382. [1]

Heim, N. (1977). Kastration bei Sexualstraftätern - eine kritische Betrachtung! In G. Nass (Ed.), *Kriminalätiologie und Prophylaxe: Wandlungen der weiblichen Kriminalität; Folgen der chirurgischen Kastration; Hirnschädigung durch Rauschgift; Kriterien zur Strafumwandlung* (pp. 7-20). Kassel: Verlag Gesellschaft für vorbeugende Verbrechensbekämpfung. [1]

Heim, N. (1977). Vollzugskrankenhaus Hohenasperg: Kastration bei Sexualstraftätern - Ergebnisse einer Erkundungsstudie. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 97-110). Stuttgart: Eigenverlag. [1]

Heim, N. (1980). *Die Kastration und ihre Folgen bei Sexualstraftätern*. Göttingen: Schwartz. [1]

Hendriks, J., & Bijleveld, C. (2005). Recidive van jeugdige zedendelinquenten na residentiële behandeling. *Tijdschrift voor Seksuologie*, 29, 150-160. [1]

Hendriks, J., & Bijleveld, C. (2005). Recidive van jeugdige zedendelinquenten: poliklinisch behandelden versus niet-behandelnden. *Tijdschrift voor Seksuologie*, 29, 215-225. [3]

Hersh, K. R., Baldwin, K., & Gray-Little, B. (1999). *Treatment completion and recidivism among incarcerated sex offenders*. Paper presented at the Annual Conference of the Association for the Treatment of Sexual Abusers, Orlando, FL. [2]

Hiob, J. (1977). Die Kastrationsbehandlung und andere Verfahren zu Therapie der Sexualstraftäter im Berliner Strafvollzug. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 91-96). Stuttgart: Eigenverlag. [1]

Hudson, S. M., Wales, D. S., Bakker, L., & Ward, T. (2002). Dynamic risk factors: the Kia Marama evaluation. *Sexual Abuse: A Journal of Research and Treatment*, 14(2), 103-119. [1]

Huot, S. J. (1997). Research summary: Sex offender treatment and recidivism. Minnesota, MN: Minnesota Department of Corrections. [3]

- Huot, S. J. (October 2002). *Recidivism, recidivism, recidivism! An update of several Minnesota recidivism studies*. Paper presented at the 21st Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers on 'Best Practices: Clinical and Research Collaborations', Montréal, Québec, Canada. [3]
- Janner, J. (1959). Über die Bedeutung psychiatrischer Betreuung verurteilter Sexualdelinquenten. *Schweizerische Zeitschrift für Strafrecht*, 74, 310-319. [3]
- Jenkins-Hall, K. D., Osborn, C. A., Anderson, C. S., Anderson, K. A., & Shockley-Smith, C. (1989). The Center for Prevention of Child Molestation. In D. R. Laws (Ed.), *Relapse prevention with sex offenders* (pp. 269-291). New York: Guilford Press. [1, 4]
- Jones, R. L., Winkler, M. X., Kacin, E., Salloway, W. N., & Weissman, M. (1998). Community-based sexual offender treatment for inner-city African-American and Latino youth. In W. L. Marshall & Y. M. Fernandez (Eds.), *Sourcebook of treatment programs for sexual offenders. Applied clinical psychology* (pp. 457-476). New York, NY, US: Plenum Press. [1]
- Jost, F. (1975). On the treatment of abnormal sexual behaviour with the anti-androgen cyproterone acetate (1971 to 1975). *Der informierte Arzt*, 3, 303-309. [1]
- Jost, F. (1979). Zur Behandlung und Resozialisierung von Sexualdelinquenten mit Antiandrogenen und Psychotherapie. *Schweizer Archiv für Neurologie, Neurochirurgie und Psychiatrie*, 124(2), 243-253. [1]
- Kafka, M. P., & Prentky, R. (1992). Fluoxetine treatment of nonparaphilic sexual addictions and paraphilias in men. *Journal of Clinical Psychiatry*, 53(10), 351-358. [1, 4]
- Kahn, T. J., & Chambers, H. J. (1991). Assessing reoffense risk with juvenile sexual offenders. *Child Welfare*, 70(3), 333-345. [1]
- Kansas Department of Corrections. (2000). Offender programs evaluation: Volume IV. Topeka: Kansas Department of Corrections. [3]
- Kaul, J., Huot, S. J., Epperson, D., & Dornfeld, M. (1994). Sex offenders released in 1988. St. Paul, MN: Minnesota Department of Corrections. [3]
- Kemper, J. (1992). *Sexualtherapeutische Praxis Teil II: Forschungsergebnisse und Materialien*. München: Pfeiffer. [1, 4]
- Kiersch, T. A. (1990). Treatment of sex offenders with Depo-Provera. *Bulletin of the American Academy of Psychiatry and the Law*, 18, 179-187. [1]
- Kruttschnitt, C., Uggens, C., & Shelton, K. (2000). Predictors of Desistance among Sex Offenders: The Interaction of Formal and Informal Social Controls. *Justice Quarterly*, 17(1), 61-87. [1]
- Langelüddeke, A. (1963). *Die Entmannung von Sittlichkeitsverbrechern*. Berlin: de Gruyter. [1]
- Langevin, R., Paitich, D., Hucker, S. J., Newman, S., Ramsay, G., Pope, S., . . . Anderson, C. (1979). The effect of assertiveness training, Provera, and sex of therapist in the treatment of genital exhibitionism. *Journal of Behavior Therapy and Experimental Psychiatry*, 10, 275-282. [3]
- Laschet, U., & Laschet, L. (1967). Antiandrogentherapie der pathologisch gesteigerten und abartigen Sexualität des Mannes. *Klinische Wochenschrift*, 45(6), 324-325. [1, 4]
- Laschet, U., & Laschet, L. (1971). Klinische Ergebnisse über die Hemmung der Sexualität durch Antiandrogene. *Journal of Neuro-visceral Relations (Suppl.)*, 10, 388-393. [1, 4]

- Leipziger, K. (2000). *Forensische Psychiatrie am Bezirkskrankenhaus Bayreuth. Beschreibung und Untersuchung der Rahmenbedingungen, Konzepte und Behandlungsergebnisse bei nach Paragraph 63 StGB im Maßregelvollzug untergebrachten Patienten im Bezirkskrankenhaus Bayreuth unter besonderer Berücksichtigung der Gruppe Sexualstraftäter*. Regensburg: Roderer. [1]
- Letourneau, E. J., Henggeler, S. W., Borduin, C. M., Schewe, P. A., McCart, M. R., Chapman, J. E., & Saldana, L. (2009). Multisystemic therapy for juvenile sexual offenders: 1-year results from a randomized effectiveness trial. *Journal of Family Psychology, 23*(1), 89-102. [4]
- Lidberg, L. (1986). Sexualstraftaten in Schweden: Strafmaßnahmen, Rückfälligkeit und Behandlung. In J. Heinrichs (Ed.), *Vergewaltigung. Die Opfer und die Täter* (pp. 114-116). Braunschweig: Holtzmeier. [4]
- Lindsay, W. R., Marshall, I., & Neilson, C. (1998). The treatment of men with a learning disability convicted of exhibitionism. *Research in Developmental Disabilities, 19*(4), 295-316. [2]
- Lindsay, W. R., Neilson, C. Q., Morrison, F., & Smith, A. H. W. (1998). The treatment of six men with a learning disability convicted of sex offences with children. *British Journal of Clinical Psychology, 37*(1), 83-98. [2]
- Lindsay, W. R., & Smith, A. H. W. (1998). Responses to treatment for sex offenders with intellectual disability: A comparison of men with 1- and 2-year probation sentences. *Journal of Intellectual Disability Research, 42*(5), 346-353. [2]
- Lindsay, W. R., Smith, A. H. W., Law, J., Quinn, K., Anderson, A., Smith, A., . . . Allan, R. (2002). A treatment service for sex offenders and abusers with intellectual disability: Characteristics of referrals and evaluation. *Journal of Applied Research in Intellectual Disabilities, 15*(2), 166-174. [1]
- Lohse, H., & Röbbeling, G. (1980). Therapie von nichtinhaftierten Sexualstraftätern *Psychotherapie und Gesellschaft, Ausbildung, Sozial- und Drogentherapie, Gruppentherapie (Klinische Psychologie und Psychotherapie: Kongreßbericht Berlin 1980)* (Vol. 5). Tübingen: Deutsche Gesellschaft für Verhaltenstherapie e.V. (DGVT). [1, 4]
- Lowden, K., Hetz, N., Harrison, L., Patrick, D., English, K., & Pasini-Hill, D. (2003). Evaluation of Colorado's prison therapeutic community for sex offenders: A report of findings (pp. 218). Denver, CO: Office of Research and Statistics, Division of Criminal Justice. [3]
- Maletzky, B. M. (1991). *Treating the sexual offender*. Newbury Park, CA: Sage. [1, 4]
- Maletzky, B. M. (1991). The use of medroxyprogesterone acetate to assist in the treatment of sexual offenders. *Annals of Sex Research, 4*(2), 117-129. [3]
- Maletzky, B. M. (1993). Factors associated with success and failure in the behavioral and cognitive treatment of sexual offenders. *Annals of Sex Research, 6*, 241-258. [1]
- Maletzky, B. M., & Field, G. (2003). The biological treatment of dangerous sexual offenders. A review and preliminary report of the Oregon pilot depo-Provera program. *Aggression and Violent Behavior, 8*(4), 391-412. [3]
- Maletzky, B. M., & Steinhauser, C. (2002). A 25-year follow-up of cognitive/behavioral therapy with 7,275 sexual offenders. *Behavior Modification, 26*(2), 123-147. [1]
- Maletzky, B. M., Tolan, A., & McFarland, B. (2006). The Oregon depo-Provera Program: A Five-Year Follow-Up. *Sexual Abuse: Journal of Research and Treatment, 18*(3), 303-316. [3]

- Marques, J. K. (1988). The Sex Offender Treatment and Evaluation Project: California's new outcome study. In R. A. Prentky & V. L. Quinsey (Eds.), *Human sexual aggression: Current perspectives* (pp. 235-243). New York: New York Academy of Science. [6]
- Marques, J. K. (1999). How to answer the question "Does sex offender treatment work?". *Journal of Interpersonal Violence*, *14*(4), 437-451. [6]
- Marques, J. K., Day, D. M., Nelson, C., & Miner, M. H. (1989). The sex offender treatment and evaluation project: California's relapse prevention program. In D. R. Laws (Ed.), *Relapse prevention with sex offenders* (pp. 247-267). New York: Guilford Press. [6]
- Marques, J. K., Day, D. M., Nelson, C., & West, M. A. (1993). Findings and recommendations from California's experimental treatment program. In G. C. N. Hall & R. Hirschman (Eds.), *Sexual aggression: Issues in etiology, assessment, and treatment. Series in applied psychology: Social issues and questions* (pp. 197-214). Philadelphia, PA, US: Taylor & Francis. [6]
- Marques, J. K., Day, D. M., Nelson, C., & West, M. A. (1994). Effects of cognitive-behavioral treatment on sex offender recidivism: Preliminary results of a longitudinal study. *Criminal Justice and Behavior*, *21*(1), 28-54. [6]
- Marques, J. K., & Nelson, C. (1992). The relapse prevention model: Can it work with sex offenders? In R. D. Peters, R. J. McMahon & V. L. Quinsey (Eds.), *Aggression and violence throughout the life span* (pp. 222-243). Thousand Oaks, CA: Sage Publications. [6]
- Marques, J. K., Nelson, C., West, M. A., & Day, D. M. (1994). The relationship between treatment goals and recidivism among child molesters. *Behaviour Research and Therapy*, *32*(5), 577-588. [6]
- Marshall, P. (1997). A reconviction study of HMP Grendon Therapeutic Community. London: Home Office. [6]
- Marshall, W. L. (1994). Treatment effects on denial and minimization in incarcerated sex offenders. *Behaviour Research and Therapy*, *32*(5), 559-564. [4]
- Marshall, W. L., & Barbaree, H. E. (1988). An outpatient treatment program for child molesters. In R. A. Prentky & V. L. Quinsey (Eds.), *Human sexual aggression: Current perspectives* (pp. 205-214). New York: New York Academy of Science. [6]
- Marshall, W. L., Marshall, L. E., Serran, G. A., & O'Brien, M. D. (2008). Sexual offender treatment: A positive approach. *Psychiatric Clinics of North America*, *31*(4), 681-696. [2]
- Martin, I. (1999). *Efficacité d'un programme cognitif-behavioural institutionnel pour délinquants sexuels*. (Unpublished doctoral dissertation), University of Montreal. [3]
- Matthews, R., & Hines, M. (1996). Impact of Booster Programme on sex offenders. *Prison Research & Development Bulletin*, *2*, 6. [1, 4]
- McConaghy, N., Armstrong, M. S., & Blaszczynski, A. (1985). Expectancy, covert sensitization and imaginal desensitization in compulsive sexuality. *Acta Psychiatrica Scandinavica*, *72*, 176-187. [2]
- McConaghy, N., Blaszczynski, A., & Kidson, W. (1988). Treatment of sex offenders with imaginal desensitization and/or medroxyprogesterone. *Acta Psychiatrica Scandinavica*, *77*, 199-206. [4]

- McConaghy, N., Blaszczynski, A. P., Armstrong, M. S., & Kidson, W. (1989). Resistance to treatment of adolescent sex offenders. *Archives of Sexual Behavior*, 18(2), 97-107. [2]
- McGrath, R. J., Cumming, G. F., Livingston, J. A., & Hoke, S. E. (2003). Outcome of a treatment program for adult sex offenders. From prison to community. *Journal of Interpersonal Violence*, 18(1), 3-17. [3]
- McGuire, T. J. (2000). Correctional institution based sex offender treatment: A lapse behavior study. *Behavioral Sciences and the Law*, 18(1), 57-71. [3]
- McGurk, B. J., & Newell, T. C. (1987). Social skills training: Case study with a sex offender. In B. J. McGurk, D. M. Thornton & M. Williams (Eds.), *Applying psychology to imprisonment: Theory and practice* (pp. 219-225). London, England: Her Majesty's Stationery Office Books. [1]
- Meyer, A. E., Balck, F., Dahme, B., Engel, K., Haag, A., & Koch, U. (1989). Der Sonderforschungsbereich 115 für "Psychosomatik, Klinische Psychologie und Psychotherapie" in Hamburg: Ausgewählte Ergebnisse. *Zeitschrift für Psychosomatische Medizin und Psychoanalyse*, 35(4), 302-327. [1]
- Meyer, L. C., & Romero, J. (1980). A ten year follow-up of sex offender recidivism. Pennsylvania, PA: J. J. Peters Institute. [3]
- Meyer, W. J., Cole, C., & Emory, E. (1992). Depo provera treatment for sex offending behavior: An evaluation of outcome. *Bulletin of the American Academy of Psychiatry and the Law*, 20(3), 249-259. [3]
- Micheroli, R., & Bategay, R. (1985). Ambulante Behandlung von Sexualdelinquenten mit Cyproteronacetat (Androcur). *Schweizer Archiv für Neurologie, Neurochirurgie und Psychiatrie*, 136, 37-58. [1]
- Miner, M. H., & Dwyer, S. M. (1995). Analysis of dropouts from outpatient sex offender treatment. *Journal of Psychology and Human Sexuality*, 7(3), 77-93. [2]
- Miner, M. H., Marques, J. K., Day, D. M., & Nelson, C. (1990). Impact of relapse prevention in treating sex offenders: Preliminary findings. *Annals of Sex Research*, 3(2), 165-185. [6]
- Miner, M. H., Siekert, G. P., & Ackland, M. A. (1997). Evaluation: Juvenile Sex Offender Treatment Program, Minnesota Correctional Facility—Sauk Centre. Final report—Biennium 1995-1997. Minneapolis, MN: University of Minnesota, Department of Family Practice and Community Health, Program in Human Sexuality. [1]
- Minnesota Department of Corrections. (2000). Community-based sex offender program evaluation project. 1999 Report to the Legislature, revised January 2000. St. Paul: Minnesota Department of Corrections. [3]
- Mothes, C., Lenhert, J., Samini, F., & Usher, J. (1973). Chemical trial of cyproterone acetate (Androcur) in sexual deviations - collected assessments. *Schering Medical News*, 2, 26. [1]
- Müller, D. (1974). Stereotaktische Hypothalamotomie bei sexuellen Triebstörungen. Erfahrungen bei 16 Patienten. In G. Nass (Ed.), *Neue Erkenntnisse zur Behandlung abweichenden Verhaltens insbesondere sexueller Delinquenz: Neurochirurgie, analytische Sozialtherapie, Sozialpädagogik* (pp. 29-36). Kassel: Verlag Gesellschaft für vorbeugende Verbrechensbekämpfung. [1]

- Müller, D., & Lüdecke, D. K. (1977). Neurochirurgisch-neurologische und endokrinologische Auswirkungen der stereotaktischen vorderen Hypothalamotomie, Bericht über 27 Fälle. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtätern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 185-195). Stuttgart: Eigenverlag. [1, 4]
- Mulloy, R., & Smiley, W. C. (1996, August). *Recidivism and treated sexual offenders*. Paper presented at the International Congress of Psychology, August 16-21, Montréal, Canada. [3]
- Murphy, P. (1998). A therapeutic programme for imprisoned sex offenders: Progress to date and issues for the future. *Irish Journal of Psychology, 19*(1), 190-207. [1, 4]
- Murray, M. A., Bancroft, J. H., Anderson, D. C., Tennent, T. G., & Carr, P. J. (1975). Endocrine changes in male sexual deviants after treatment with anti-androgens, oestrogens or tranquillizers. *Journal of Endocrinology, 67*(2), 179-188. [1, 4]
- Nicholaichuk, T. P. (1991). A model for a rurally based day hospital treatment program for sex offenders. *International Journal of Partial Hospitalization, 7*(2), 91-100. [1, 4]
- Nicholaichuk, T. P., Gordon, A., Gu, D., & Wong, S. (2000). Outcome of an institutional sexual offender treatment program: A comparison between treated and matched untreated offenders. *Sexual Abuse: A Journal of Research and Treatment, 12*(2), 139-153. [3]
- Nicholaichuk, T. P., & Yates, P. (2002). Treatment Efficacy - Outcomes of the Clearwater Sex Offender Program. In B. K. Schwartz (Ed.), *The Sex Offender* (Vol. 4: Current Treatment Modalities and Systems Issues, pp. 7-1 - 7-18). New York: Civic Research Institute. [2]
- Nutbrown, V., & Stasiak, E. (1987). Research monograph: A retrospective analysis of O.C.I. cost effectiveness 1977-1981. *Ontario Correctional Institute, 2*, 1-16. [3]
- O'Brien, P., & Freeman, B. (1997). A national survey of NAPN treaters of adolescent sex offenders. *Journal of Offender Rehabilitation, 26*(1-2), 109-124. [1]
- Office of the Legislative Auditor. (1997). Recidivism of adult felons. A program evaluation report. St. Paul: State of Minnesota. [3]
- Ohio Department of Rehabilitation and Correction. (1996). Five year recidivism follow-up of sex offender releases. [3]
- Olver, M. E., Wong, S. C. P., & Nicholaichuk, T. P. (2009). Outcome evaluation of a high-intensity inpatient sex offender treatment program. *Journal of Interpersonal Violence, 24*(3), 522-536. [3]
- Oregon Department of Corrections. (1997). Outcome evaluation of the Jackson County Sex Offender Supervision and Treatment Program Retrieved 5/29/01, from <http://www.doc.state.or.us/research/jackrpt2.pdf> [3]
- Owen, G., & Steele, N. M. (1991). Incest offenders after treatment. In M. Q. Patton (Ed.), *Family sexual abuse: Frontline research and evaluation* (pp. 178-198). Thousand Oaks, CA: Sage. [2]
- Pasel, F. (1999). Arbeit mit Sexualstraftätern. In G. Deegener (Ed.), *Sexuelle und körperliche Gewalt. Therapie jugendlicher und erwachsener Täter* (pp. 208-221). Weinheim: Psychologie Verlags Union. [1]

- Pellerin, B., Proulx, J., Ouimet, M., Paradis, Y., McKibben, A., & Aubut, J. (1996). Étude de la récidive post-traitement chez des agresseurs sexuels judiciairisés. *Criminologie*, 29, 85-108. [2]
- Pellerin, B., Proulx, J., Ouimet, M., Paradis, Y., McKibben, A., & Aubut, J. (1996). Étude de la récidive post-traitement chez des agresseurs sexuels judiciairisés. [Study of Posttreatment Recidivism among Convicted Sex Offenders]. *Criminologie*, 29(1), 85-108. [2]
- Perilstein, R. D., Lipper, S., & Friedman, L. J. (1991). Three cases of paraphilias responsive to Fluoxetine treatment. *Journal of Clinical Psychiatry*, 52(4), 169-170. [1, 4]
- Perkins, D. (1987). A psychological treatment programme for sex offenders. In B. J. MacGurk & D. M. Thornton (Eds.), *Applying psychology to imprisonment: Theory and practice* (pp. 191-217). London: Her Majesty's Stationery Office Books. [3]
- Peters, J. J., Pedigo, J., Steg, J., & McKenna Jr., J. (1968). Group psychotherapy of the sex offender. *Federal Probation*, 32, 41-45. [3]
- Peters, J. J., & Roether, H. A. (1971). *Success and failure of sex offenders*. Philadelphia, PA: American Association for the Advancement of Science. [6]
- Pithers, W. D., & Cumming, G. F. (1989). Can relapses be prevented? Initial outcome data from the Vermont treatment program for sexual aggressors. In D. R. Laws (Ed.), *Relapse prevention with sex offenders* (pp. 313-325). New York: Guilford Press. [1]
- Pithers, W. D., Gray, A. S., Busconi, A., & Houchens, P. (1998). Children with sexual behavior problems: Identification of five distinct child types and related treatment considerations. *Child Maltreatment*, 3(4), 384-406. [2, 4]
- Pithers, W. D., Kashima, K. M., Cumming, G. F., Beal, L. S., & Buell, M. M. (1988). Relapse prevention of sexual aggression. In R. A. Prentky & V. L. Quinsey (Eds.), *Human sexual aggression: Current perspectives* (pp. 244-260). New York: New York Academy of Science. [1]
- Pitzing, H.-J. (2002). *Psychotherapeutische Ambulanz für Sexualstraftäter*. Stuttgart: Bewährungshilfe Stuttgart. [1]
- Porporino, F. J., & Nouwens, T. M. (1993). When are sex offenders at risk for reoffending? Results of two long-term follow-up studies. *Forum on Corrections Research*, 5(2). [3]
- Prentky, R. A., & Burgess, A. W. (1990). Rehabilitation of child molesters: A cost-benefit analysis. *American Journal of Orthopsychiatry*, 60(1), 108-117. [1]
- Proulx, J., Ouimet, M., Pellerin, B., Paradis, Y., McKibben, A., & Aubut, J. (1998). Posttreatment recidivism rates in sexual aggressors: a comparison between dropout and nondropout subjects. [2]
- Quayle, M., Deu, N., & Giblin, S. (1998). Sexual knowledge and sex education in a secure hospital setting. *Criminal Behaviour and Mental Health*, 8 (Suppl), 66-76. [1, 4]
- Quinsey, V. L., Chaplin, T. C., & Carrigan, W. F. (1980). Biofeedback and signaled punishment in the modification of inappropriate sexual age preferences. *Behavior Therapy*, 11, 567-576. [1, 4]
- Quinsey, V. L., Khanna, A., & Malcolm, P. B. (1998). A retrospective evaluation of the Regional Treatment Sex Offender Treatment program. *Journal of Interpersonal Violence*, 13, 621-644. [3]

- Ralston, M. E. (1990). *Navy incest perpetrator/event characteristics and treatment outcome*. (Dissertation), Florida Institute of Technology. [4]
- Rasmussen, L. A. (1995). *Factors related to recidivism among juvenile sexual offenders*. (Dissertation), The University of Utah. [3]
- Rattenbury, F. R. (1985). *The outcomes of hospitalized and incarcerated sex offenders: A study of offender types, recidivism rates, and identifying characteristics of the repeat offender*. (Unpublished doctoral dissertation), Loyola University of Chicago, Chicago, IL. [3]
- Redondo Illescas, S., & Garrido Genoves, V. (2008). Efficacy of a psychological treatment for sex offenders. [Research Support, Non-U.S. Gov't]. *Psicothema*, 20(1), 4-9. [3]
- Repp, M. (2003). *Sex Offender Probation Programs in DuPage, Lake, and Winnebago Counties*. Chicago, IL: Illinois Criminal Justice Information Authority. [3]
- Rice, M. E., Harris, G. T., & Quinsey, V. L. (1991). Evaluation of an institution-based treatment program for child molesters. *Canadian Journal of Program Evaluation*, 6(1), 111-129. [6]
- Rice, M. E., Quinsey, V. L., & Harris, G. T. (1991). Sexual recidivism among child molesters released from a maximum security psychiatric institution. *Journal of Consulting and Clinical Psychology*, 59, 381-386. [6]
- Robinson, D. (1989). Research on sex offenders: What do we know? *Forum on Corrections Research*, 1(1) [1]
- Robinson, D. (1996). Factors influencing the effectiveness of cognitive skills training. *Forum on Corrections Research*, 8(3), 6-9. [6]
- Roeder, F. D. (1974). Die Behandlung sexueller Triebtäter. Langzeitbeobachtungen an stereotaktischer Hypothalamomie. In G. Nass (Ed.), *Neue Erkenntnisse zur Behandlung abweichenden Verhaltens insbesondere sexueller Delinquenz: Neurochirurgie, analytische Sozialtherapie, Sozialpädagogik* (pp. 3-28). Kassel: Verlag Gesellschaft für vorbeugende Verbrechensbekämpfung. [1]
- Roether, H. A. (1972). *A definitive outcome study of a group psychotherapy program with probationed sex offenders*. (Dissertation), University of Pennsylvania. [6]
- Roger, D., & Masters, R. (1997). The development and evaluation of an emotion control training programme for sex offenders. *Legal and Criminological Psychology*, 2(Part 1), 51-64. [1, 4]
- Rösler, A., & Witztum, E. (1998). Treatment of men with paraphilia with a long-acting analogue of gonadotropin-releasing hormone. *New England Journal of Medicine*, 338(7), 416-422. [1, 4]
- Sagatun, I. J. (1982). Attributional effects of therapy with incestuous families. *Journal of Marital and Family Therapy*, 8(1), 99-104. [1]
- Sapp, A. D., & Vaughn, M. S. (1991). Sex Offender Rehabilitation Programs in State Prisons: A Nationwide Survey. *Journal of Offender Rehabilitation*, 17(1-2), 55-75. [1, 4]
- Schönhage, E., & Schazmann, W. (1983). Therapie von Sexualstraftätern in einem psychiatrischen Krankenhaus. *Psychiatrische Praxis*, 10, 93-96. [1]
- Schorsch, E. (1985). *Perversion als Straftat. Dynamik und Psychotherapie*. Berlin: Springer. [1]

- Schorsch, E., Galedary, G., Haag, A., Hauch, M., & Lohse, H. (1990). *Perversion als Straftat. Dynamik und Psychotherapie*. [1]
- Schram, D. D., Milloy, C. D., & Rowe, W. E. (1991). Juvenile sex offenders: A follow-up study of reoffense behavior (pp. ii, 43). Olympia, WA: Washington State Institute for Public Policy. [1]
- Schultka, H. (1977). Katamnestische Erhebungen über Sexualdelinquenten 5 - 10 Jahre nach kastrativer Operation -- ohne und mit zusätzlicher Cyproteronacetatbehandlung. In Justizministerium Baden-Württemberg (Ed.), *Maßnahmen zur Behandlung von Sexualtättern. Bericht über die überregionale Fortbildungstagung vom 12. bis 16. September 1977 im "Haus Ottilienberg" in Eppingen bei Heilbronn* (pp. 117-128). Stuttgart: Eigenverlag. [2]
- Seitz, C., & Specht, F. (2001). Legalbewährung nach Entlassung aus den sozialtherapeutischen Einrichtungen des niedersächsischen Justizvollzuges. In G. Rehn, B. Wischka, F. Lösel & M. Walter (Eds.), *Behandlung "gefährlicher Straftäter": Grundlagen, Konzepte, Ergebnisse* (pp. 348-363). Pfaffenweiler: Centaurus. [2]
- Servais, J., & Hubin, P. (1968). Synthèse des connaissances actuelles concernant un inhibiteur de la libido chez le male, la methyloestrenolone. [Synthesis of current knowledge of an inhibitor of libido in the male, the methylestrenolone]. *Encephale*, 57(4), 333-352. [1, 4]
- Shaw, T. A., Herkov, M. J., & Greer, R. A. (1995). Examination of treatment completion and predicted outcome among incarcerated sex offenders. *Bulletin of the American Academy of Psychiatry and the Law*, 23(1), 35-41. [2, 4]
- Smiley, W. C., & Mulloy, R. (1997, June). *The specific impact of REBT-based treatment on sex offenders*. Paper presented at the Poster presented at the meeting of the Canadian Psychological Association, June 11-14, Toronto, Canada. [1, 4]
- Smith, C. J., Hayler, B., & Craig, K. (2001). The Illinois Department Of Corrections' Juvenile Sex Offender Treatment Program: The Final Report Of The Program Evaluation. Chicago, IL: The Illinois Criminal Justice Information Authority. [4]
- Smith, M. A. (2000). *The impact of prison-based drug treatment on the timing and risk of reincarceration of sexual offenders*. (Dissertation), Florida State University, Tallahassee, FL. [3]
- Song, L., & Lieb, R. (1995). Washington State sex offenders: Overview of recidivism studies. Olympia, WA: Washington State Institute for Public Policy. [1;3]
- Stalans, L. J., Seng, M., Yarnold, P., Lavery, T., & Swartz, J. (2001). Process and initial impact evaluation of the Cook County Adult Probation Department's Sex Offender Program: Final and summary report for the period of June, 1997 to June, 2000: Illinois Department of Corrections. [3]
- Stalans, L. J., Seng, M., & Yarnold, P. R. (2002). Long-term impact evaluation of specialized sex offender probation programs in Lake, DuPage and Winnebago counties. Chicago, IL: Illinois Criminal Justice Information Authority. [2]
- Stein, D. J., Hollander, E., Anthony, D. T., Schneider, F. R., Fallon, B. A., Liebowitz, M. R., & Klein, D. F. (1992). Serotonergic medications for sexual obsessions, sexual addictions, and paraphilias. *Journal of Clinical Psychiatry*, 53(8), 267-271. [1, 4]
- Stephenson, M. (1991). A summary of an evaluation of the community sex offender program in the pacific region. *Forum on Corrections Research*, 3(4). [5]

- Stirpe, T. S., Wilson, R. J., & Long, C. (2001). Goal Attainment Scaling with Sexual Offenders: A Measure of Clinical Impact at Posttreatment and at Community Follow-Up. *Sexual Abuse: A Journal of Research and Treatment*, 13(2), 65-77. [1, 4]
- Studer, L. H., Alywn, A. S., & Reddon, J. R. (2005). Testosterone, sexual offense recidivism, and treatment effect among adult male sex offenders. *Sexual Abuse: Journal of Research & Treatment*, 17(2), 171. [2]
- Studer, L. H., & Reddon, J. R. (1998). Treatment may change risk prediction for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 10(3), 175-181. [2]
- Studer, L. H., Reddon, J. R., Roper, V., & Estrada, L. (1994). Sexual offender recidivism. An evaluation of the Phoenix program. Edmonton, CAN: Alberta Hospital. [2]
- Studer, L. H., Reddon, J. R., Roper, V., & Estrada, L. (1996). Phoenix: An inhospital treatment program for sex offenders. *Journal of Offender Rehabilitation*, 23(1/2), 91-97. [2]
- Stürup, G. K. (1953). Sexual offenders and their treatment in Denmark and other Scandinavian countries. *International Review of Criminal Policy*, 4, 1-19. [3]
- Stürup, G. K. (1968). Treatment of sexual offenders in Herstedvester Denmark: The rapists. *Acta Psychiatrica Scandinavica Supplement*, 204, 1-62. [3]
- Stürup, G. K. (1971). Castration: the total treatment. *International Psychiatry Clinics*, 8(4), 175-196. [1]
- Stürup, G. K. (1971). Treatment of the sex offender. Castration: the total treatment. *International Psychiatry Clinics*, 8(4), 175-196. [1]
- Swaffer, T., Hollin, C. R., Beech, A. R., Beckett, R. C., & Fisher, D. (2000). An exploration of child sexual abuser's sexual fantasies before and after treatment. *Sexual Abuse: A Journal of Research and Treatment*, 12(1), 61-68. [1, 4]
- Swanson, C. K., & Garwick, G. B. (1990). Treatment for low-functioning sex offenders: Group therapy and interagency coordination. *Mental Retardation*, 28(3), 155-161. [1]
- Thibaut, F., Kuhn, J. M., Cordier, B., & Petit, M. (1998). Les traitements hormonaux de la delinquance sexuelle. *L'Encéphale*, 24(2), 132-137. [1]
- Thomas-März, V., & Müller-Isberner, R. (1995). Psychotherapie mit Sexualstraftätern. *Zeitschrift für Strafvollzug und Straffälligenhilfe*, 44(6), 344-347. [1]
- Travers, O. (1998). Treatment v. punishment: A case for treating sex offenders in the community. *Irish Journal of Psychology*, 19(1), 226-233. [1]
- Valliant, P. M., Sloss, B. K., & Raven-Brooks, L. (1997). Effects of Brief Cognitive-Behavioral Therapy on Recidivism among Sex and Non-Sex Offenders. *Journal of Offender Rehabilitation*, 25(1-2), 163-174. [4]
- van Beek, D. (1999). *De delictscenarioprocedure bij seksueel agressieve delinquenten; een onderzoek naar de bruikbaarheid van de delictscenarioprocedure in de behandeling van seksueel sggressieve delinquenten in de Dr. Henri van der Hoevenkliniek*. Arnhem: Gouda Quint. [1]

- Waite, D., Keller, A., McGarvey, E. L., Wieckowski, E., Pinkerton, R., & Brown, G. L. (2005). Juvenile Sex Offender Re-Arrest Rates for Sexual, Violent Nonsexual and Property Crimes: A 10-Year Follow-Up. *Sexual Abuse: Journal of Research and Treatment*, 17(3), 313-331. [3]
- Walker, P., & Meyer, W. (1981). Medroxyprogesterone acetate treatment for paraphiliac sex offenders. In S. Hays, T. Roberts & K. Solway (Eds.), *Violence and the violent individual* (pp. 353-373). New York: SP Medical and Scientific Books. [1, 4]
- Washington State Institute for Public Policy. (1998). Sex offenses in Washington State: 1998 update. Olympia, WA: Washington State Institute for Public Policy. [3]
- Weiss, P. (1997). Psychotherapy of paraphilic sex offenders. *Medicine and Law*, 16(4), 753-764. [2]
- Wille, R. (1987). Zum Stand der heutigen Kastrationsforschung. *Forensia*, 8(4), 207-216. [3]
- Wille, R., & Beier, K. M. (1989). Castration in Germany. *Annals of Sex Research*, 2, 103-133. [3]
- Wilson, R. J., Picheca, J. E., & Prinzo, M. (2007). Evaluating the effectiveness of professionally-facilitated volunteerism in the community-based management of high-risk sexual offenders: Part two--A comparison of recidivism rates. *Howard Journal of Criminal Justice*, 46(4), 327-337. [3]
- Wischka, B. (1999). Behandlung von Sexualstraftätern im Niedersächsischen Justizvollzug *Schriftenreihe der JVA Lingen* (pp. 82). Lingen: Justizvollzugsanstalt Lingen. [2]
- Wolfus, B., & Stasiak, E. (1992). Report on recidivism of O.C.I. offenders released in 1985 and 1987. Brampton, ON: Ontario Correctional Institute. [2]
- Wolfus, B., & Stasiak, E. (1996). Recidivism following O.C.I. treatment: Results for offenders released in 1989 and 1991. Brampton, ON: Ontario Correctional Institute. [1]
- Wollert, R. (1988). Brief program reports: An evaluation of a communications training program within a self-help group for sexually abusive families. *Community Mental Health Journal*, 24(3), 229-235. [4]
- Worling, J. R. (1998). Adolescent sexual offender treatment at the SAFE-T Program. In W. L. Marshall, Y. M. Fernandez, S. M. Hudson & T. Ward (Eds.), *Sourcebook of treatment programs for sexual offenders* (pp. 353-365). New York: Plenum Press. [3]
- Zbytovsky, J., & Zapletalek, M. (1989). Longitudinal study of the effect, tolerance and undesired side effects of injection haloperidol decanoate applied in different psychiatric indications. *Activitas Nervosa Superior*, 31(4), 266-267. [1]
- Zgoba, K. M., & Levenson, J. (2008). Variations in the recidivism of treated and nontreated sexual offenders in New Jersey: An examination of three time frames. *Victims & Offenders*, 3(1), 10-30. [3]
- Zgoba, K. M., & Simon, L. M. J. (2005). Recidivism Rates of Sexual Offenders up to 7 Years Later: Does Treatment Matter? *Criminal Justice Review*, 30(2), 155-173. [3]
- Zohar, J., Kaplan, Z., & Benjamin, J. (1994). Compulsive exhibitionism successfully treated with Fluvoxamine: A controlled case study. *Journal of Clinical Psychiatry*, 55(3), 86-88. [1]

Information about this review

Review authors

Lead review author:

The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.

Name:	Martin Schmucker
Title:	Dr.
Affiliation:	Friedrich-Alexander-Universität Erlangen-Nürnberg
Address:	Chair for Psychological Assessment Nägelsbachstr. 49c
City, State, Province or County:	Erlangen
Postal Code:	D-91052
Country:	Germany
Phone:	+49 9131 85 64014
Email:	martin.schmucker@fau.de

Co-author:

Name:	Friedrich Lösel
Title:	Prof. Dr. Dr. h.c.
Affiliation:	Institute of Criminology, University of Cambridge, UK
Address:	Sidgwick Avenue
City, State, Province or County:	Cambridge
Postal Code:	CB3 9DA
Country:	United Kingdom
Phone:	+44 1223 335385
Email:	fal23@cam.ac.uk

Roles and responsibilities

Please give brief description of content and methodological expertise within the review team. The recommended optimal review team composition includes at least one person on the review team who has content expertise, at least one person who has methodological expertise and at least one person who has statistical expertise. It is also recommended to have one person with information retrieval expertise.

Who is responsible for the below areas? Please list their names:

- Content: Martin Schmucker, Friedrich Lösel
- Systematic review methods: Martin Schmucker, Friedrich Lösel
- Statistical analysis: Martin Schmucker
- Information retrieval: Martin Schmucker

SOURCES OF SUPPORT

No external funding was claimed for this review.

DECLARATIONS OF INTEREST

The authors are not engaged in the development or delivery of treatment programs for sexual offenders and are not affiliated with an institution providing such treatment. As independent researchers and experts, MS and FL are conducting evaluations of offender treatment programs. FL also serves on the Correctional Services Accreditation and Advice Panel of England and Wales and various advisory boards on treatment and prevention of offending in Germany.

PLANS FOR UPDATING THE REVIEW

MS and FL plan a major update in about two years' time to include more recent evaluations.

AUTHOR DECLARATION

Authors' responsibilities

By completing this form, you accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group. If an update is not submitted according to agreed plans, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to propose the update to alternative authors.

Publication in the Campbell Library

The Campbell Collaboration places no restrictions on publication of the findings of a Campbell systematic review in a more abbreviated form as a journal article either before or after the publication of the monograph version in *Campbell Systematic Reviews*. Some journals, however, have restrictions that preclude publication of findings that have been, or will be, reported elsewhere, and authors considering publication in such a journal should be aware of possible conflict with publication of the monograph version in *Campbell Systematic Reviews*. Publication in a journal after publication or in press status in *Campbell Systematic Reviews* should acknowledge the Campbell version and include a citation to it. Note that systematic reviews published in *Campbell Systematic Reviews* and co-registered with the Cochrane Collaboration may have additional requirements or restrictions for co-publication. Review authors accept responsibility for meeting any co-publication requirements.

I understand the commitment required to update a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by: Martin Schmucker

Date: 05/26/2017

Tables

TABLE 1: CHARACTERISTICS OF INCLUDED COMPARISONS

Coded variables and categories	Frequency	Percentage
<i>General study characteristics</i>		
Publication year		
1980s	4	13.8
1990s	11	37.9
Since 2000	14	48.3
Country		
USA	8	27.6
Canada	11	37.9
Great Britain	3	10.3
Germany	3	10.3
Other	4	13.8
Publication type		
Journal article	18	62.1
Book, chapter	4	13.8
Unpublished	7	24.1

Coded variables and categories	Frequency	Percentage
<i>Treatment characteristics</i>		
Mode of treatment		
Cognitive-behavioral	21	72.4
Insight oriented	2	6.9
Therapeutic community	4	13.8
(Multi)Systemic	2	6.9
Author affiliation to treatment program		
Yes	15	51.7
No	11	37.9
Unclear	3	10.3
Time of treatment implementation		
Before 1970	2	6.9
1970s	7	24.1
1980s	13	44.8
1990s	7	24.1
Sex offender specific treatment		
Yes	26	89.7
No	3	10.3
Integrity of treatment implementation		
Acceptable	10	34.5
Problematic	1	3.4
No information available	18	62.1

Coded variables and categories	Frequency	Percentage
Setting of treatment		
Prison	10	31.0
Hospital	5	17.2
Outpatient	12	41.4
Mixed	2	10.3
Individualization of treatment		
Only group treatment	9	31.0
Mainly group treatment	8	27.6
Mixed	4	13.8
Mainly individual treatment	4	13.8
Only individual treatment	4	13.8
Duration of treatment		
≤ 15 weeks	2	10.0
16 – 30 weeks	5	20.0
31 – 60 weeks	5	25.0
61 – 120 weeks	4	20.0
121 – 180 weeks	3	15.0
> 180 weeks	1	5.0
Aftercare		
Obligatory	6	20.7
Optional	5	17.2
Not offered, not reported	18	62.1
<i>Offender characteristics</i>		
Age group		
Adults	13	44.8
Adolescents	5	17.2
Mixed	1	3.4
Unclear	10	34.5

Homogeneity of age		
High	9	31.0
Medium	9	31.0
Low	6	20.7
Unclear	5	17.2
Offence type ^a		
Rape	15	51.7
Child molestation	21	72.4
Incest offences	12	41.4
Exhibitionism	7	24.1
Not specified	8	27.6
Risk category		
Low risk	8	27.6
Medium risk	12	41.4
High risk	4	13.8
Unclear	5	17.2
Treatment participation		
Voluntary	16	55.2
Non-voluntary	6	20.7
Unclear	7	24.1
<i>Methodological characteristics</i>		
Sample size		
Up to 50	7	24.1
51-150	8	27.6
151-250	5	17.2
251-500	3	10.3
More than 500	6	20.7
Design		
Level 3 (incidental assignment)	15	51.7
Level 4 (matching procedure)	8	27.6
Level 5 (randomization)	6	20.7

Follow up period (months)		
12 - 24	4	13.8
25 - 36	1	3.4
37 - 60	11	37.9
61 - 84	4	13.8
> 84	9	31.0
Type of reoffence ^a		
Sexual	28	96.6
Violent	7	24.1
Non-sexual	7	24.1
Any	14	48.3
Definition of recidivism		
Arrest	5	17.2
Charge	7	24.1
Conviction	11	37.9
Multiple definitions	3	10.3
Not indicated	3	10.3
Overall transparency of report (“descriptive validity”)		
Very low	3	10.3
Medium	18	62.1
Fair	7	24.1
Excellent	1	3.4

$k = 29$

^a individual comparisons may cover multiple categories

TABLE 2: TOTAL MEAN EFFECTS AND HETEROGENEITY

Outcome	<i>k</i>	<i>OR</i>	<i>CI</i> _{95%}	<i>Q</i>	<i>I</i> ²	Recidivism (%)	
						TG ^a	CG ^b
Sexual recidivism	28	1.41**	1.11 – 1.78	52.05**	48.1%	10.1	13.7
Any recidivism	14	1.45**	1.15 – 1.83	23.66*	45.1%	32.6	41.2

Note. *k* = number of comparisons; *OR* = mean odds ratio (random effects model with $\tau^2 = 0.14$ for sexual and $\tau^2 = .06$ for any recidivism); *CI*_{95%} = 95 % confidence interval; *Q* = test of homogeneity (χ^2 , *df* = *k* – 1); *I*² = Percentage of heterogeneity not due to chance; TG = treated group; CG = comparison group

^a *n*-weighted average ^b estimated recidivism rate

* *p* < .05, ** *p* < .01

TABLE 3: MODERATOR VARIABLES

Variables	<i>k</i>	<i>Q_{bet}</i>	<i>OR</i>	CI _{95%}		Continuous variables ^a	
				lower	upper	<i>β_w</i>	<i>z</i>
<i>Methodological characteristics</i>							
Definition of recidivism outcome		2.94					
Arrest	5		0.98	0.46	– 2.09		
Charge	7		1.65 [†]	0.93	– 2.93		
Conviction	10		1.69 [*]	1.12	– 2.54		
Multiple definitions	3		1.05	0.58	– 1.89		
Not indicated	3		1.59	0.63	– 4.01		
Design		0.13				–.06	–0.34
Level 3 (incidental assignment)	15		1.49 [*]	1.04	– 2.14		
Level 4 (matching procedure)	8		1.36	0.88	– 2.13		
Level 5 (randomization)	5		1.36	0.70	– 2.62		
Sample size ^a		2.30				–.05	–0.29
Up to 50	9		2.14 [*]	1.19	– 3.84		
51-150	8		1.27	0.75	– 2.15		
151-250	4		1.36	0.65	– 2.85		
251-500	2		1.23	0.59	– 2.60		
More than 500	5		1.32	0.85	– 2.04		
Scope of the evaluation		2.37					
Single program & location	20		1.62 ^{**}	1.20	– 2.18		
Multiple programs & locations	8		1.07	0.70	– 1.65		
Recidivism base rate	26					.39	2.27 [*]
Follow up period	28					–.03	–0.17
Overall transparency of report	28					.46	2.78 ^{**}

Variables	<i>k</i>	<i>Q_{bet}</i>	<i>OR</i>	CI _{95%}		Continuous variables ^a	
				lower	upper	<i>β_w</i>	<i>z</i>
<i>Treatment characteristics</i>							
Treatment approach		9.93*					
Cognitive-behavioral	20		1.38**	1.08	– 1.75		
Insight oriented	2		0.97	0.36	– 2.59		
Therapeutic community	4		1.24	0.69	– 2.22		
Multisystemic	2		21.76***	3.70	– 128.02		
Specific sex offender treatment		0.29					
Yes	26		1.44**	1.12	– 1.84		
No	2		1.11	0.45	– 2.74		
Setting of treatment		5.22					
Prison	9		1.25	0.85	– 1.83		
Hospital	5		1.74*	1.04	– 2.91		
Outpatient	12		1.73*	1.11	– 2.72		
Mixed	2		0.54	0.19	– 1.51		
Individualization of treatment		6.10				.41	2.47*
Only group treatment	8		1.01	0.66	– 1.55		
Mainly group treatment	8		1.38	0.89	– 2.13		
Mixed	4		1.87*	1.04	– 3.36		
Mainly individual treatment	4		1.82	0.87	– 3.82		
Only individual treatment	4		3.15*	1.14	– 8.74		
Author affiliation to treatment		2.54					
Yes	15		1.71**	1.18	– 2.47		
No	11		1.09	0.73	– 1.64		
Status of treatment program		2.05					
Model project	4		2.40*	1.12	– 5.15		
Routine practice	24		1.33*	1.03	– 1.73		
Time of treatment implementation	28					–.08	–0.48
Duration of treatment	19					–.15	–0.72

Variables	<i>k</i>	<i>Q_{bet}</i>	<i>OR</i>	CI _{95%}		Continuous variables ^a	
				lower	upper	<i>β_w</i>	<i>z</i>
<i>Offender characteristics</i>							
Age group ^a		1.83				-.30	-1.80 [†]
Adolescents only	5		2.97*	1.16	– 7.59		
Adults only	13		1.48*	1.03	– 2.12		
Treatment participation		0.00					
Voluntary	15		1.33 [†]	0.96	– 1.83		
Non-voluntary	6		1.32	0.73	– 2.37		
Risk level		9.12*				.46	2.59**
Low risk	8		1.00	0.68	– 1.47		
Medium risk	12		1.33 [†]	0.96	– 1.84		
High risk	3		3.95***	1.77	– 8.84		

k = number of comparisons; *Q_{bet}* = test of between group differences (χ^2 -distributed with *df* = number of categories – 1); *OR* = odds ratio; CI_{95%} = 95 % confidence interval; *β_w* = regression weight for continuous moderator variables.

^a Analyses refer to the continuously coded variable for age (mean), time of treatment implementation, duration of treatment, follow-up period, and sample size. For all other variables, if presented, the regression weights refer to the ordered categories presented in the table. Analyses were conducted using meta-analytic regression methods applying a mixed effects model. Since the regression models only included the tested moderator variable as single predictor in each of the analyses the significance testing of the *β*-weight via z-test equals the result for the full regression model which is therefore not individually presented.

[†]*p* < .10; **p* < .05; ***p* < .01; ****p* < .001

Figures

Figure 1: Prisma flow chart of the literature search and selection process

(Figure template adapted from Moher et al., 2009)

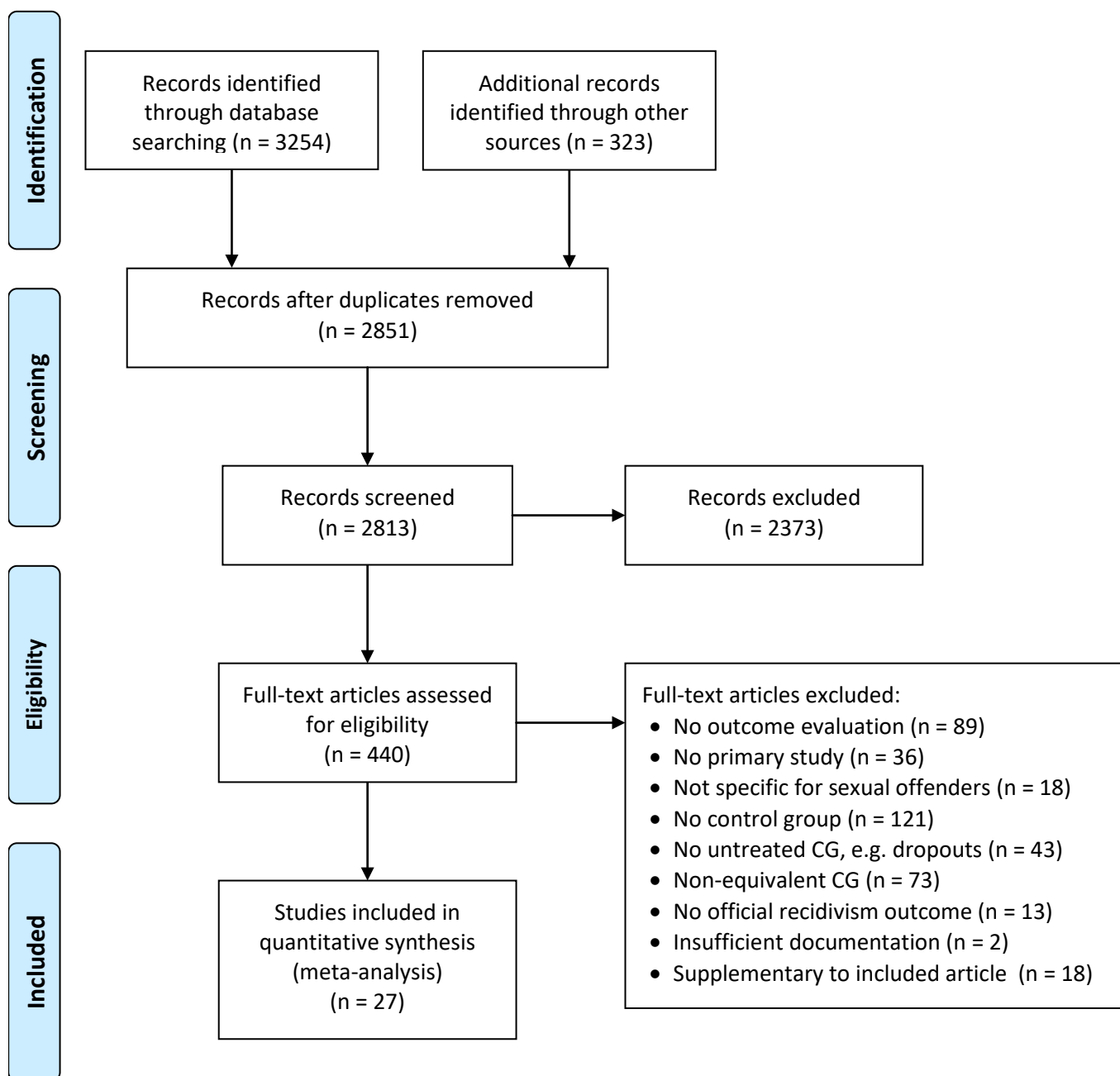


FIGURE 2: FOREST PLOT OF SEXUAL RECIDIVISM OUTCOMES

($k = 28$; Mean odds ratio and 95%-confidence interval)

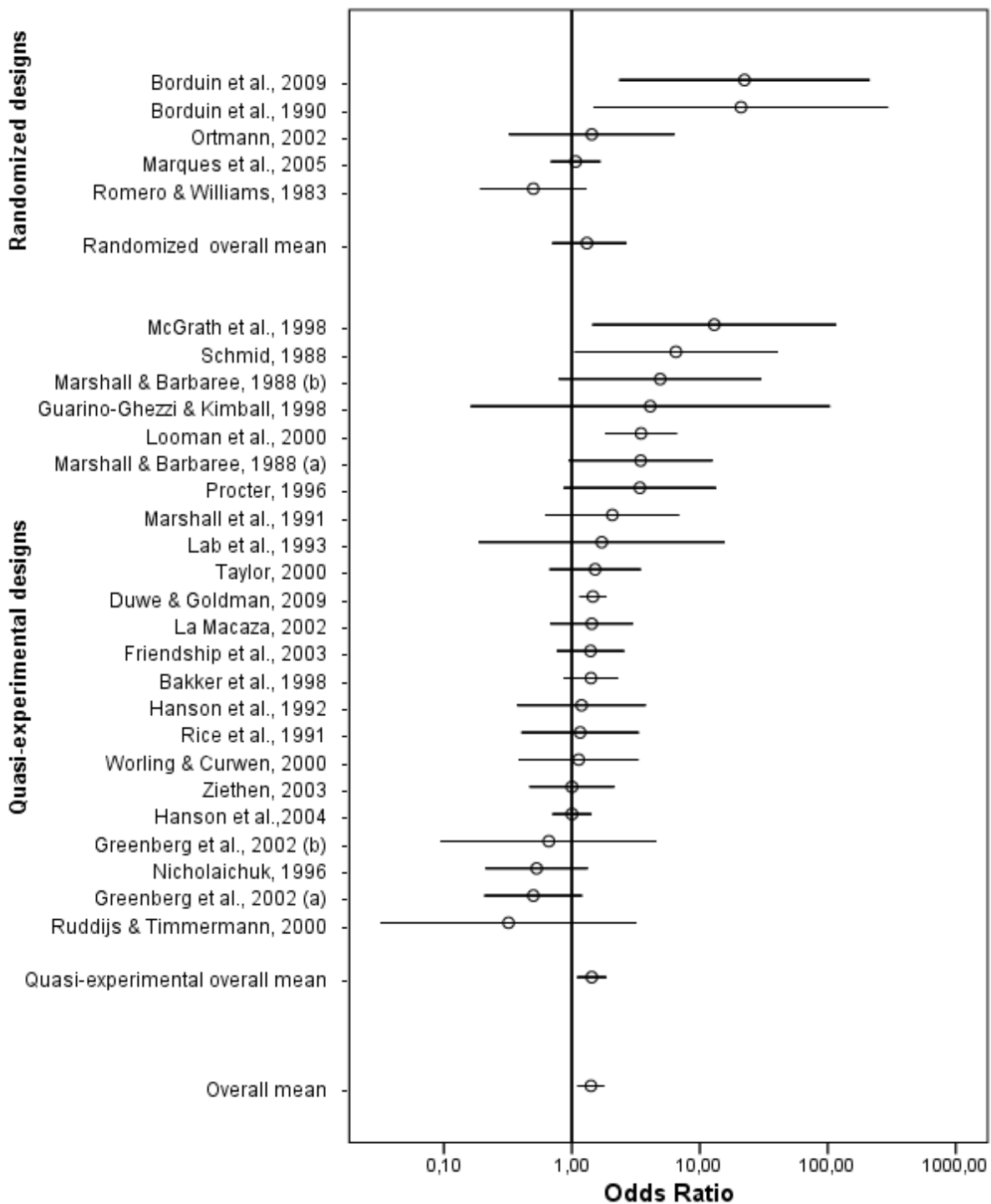
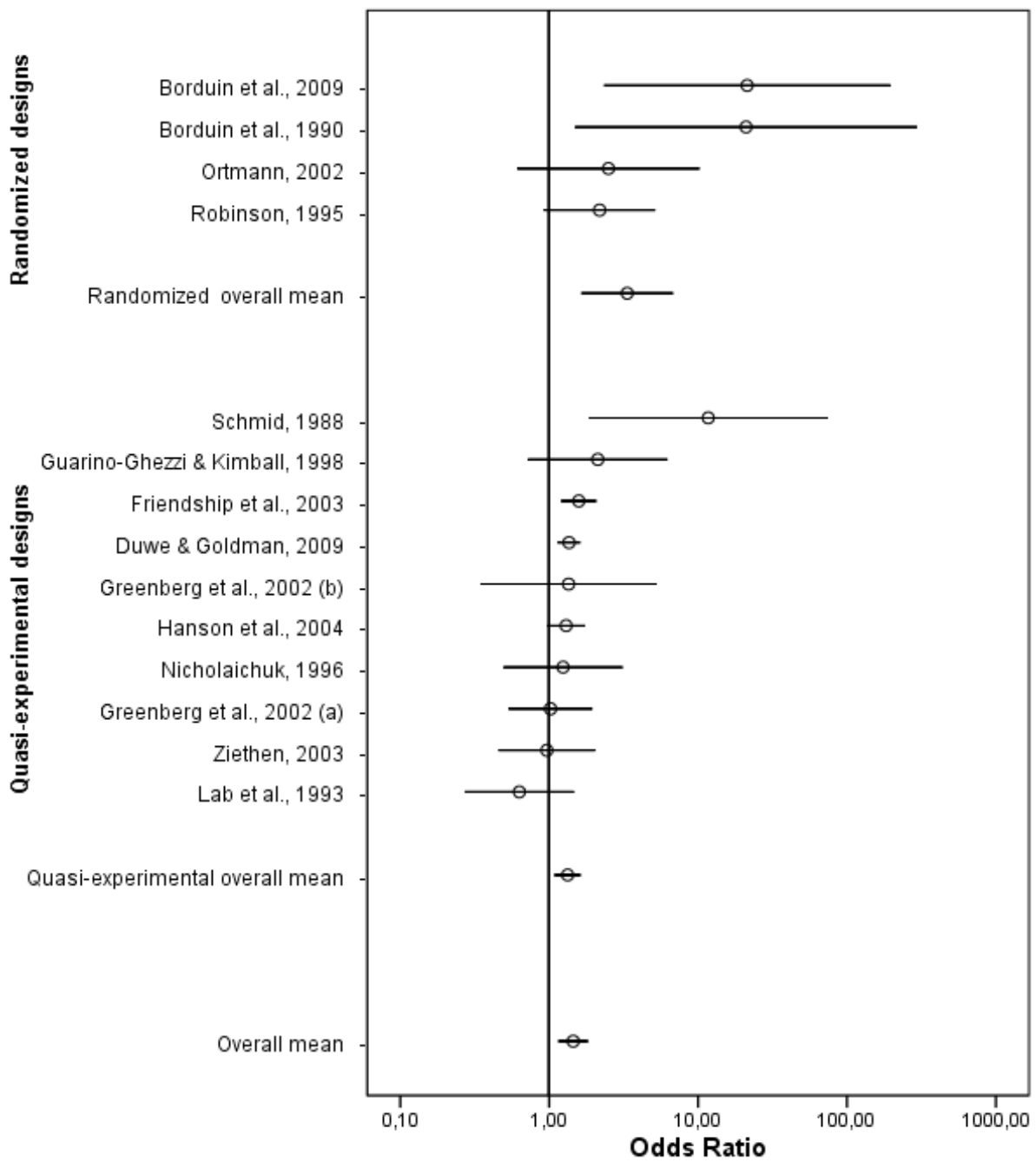


FIGURE 3: FOREST PLOT OF GENERAL RECIDIVISM OUTCOMES

($k = 14$; Mean odds ratio and 95%-confidence interval)



Appendix

APPENDIX 1: RELEVANT STUDIES THAT APPEARED MORE RECENTLY

The following is a commented list of more recent evaluations of SOTP⁵. Some of these studies might have been eligible for the current meta-analysis, but we did not fully code them because we know of two rather large evaluations in progress that require a further update of our meta-analysis in the near future. In the meanwhile, the following list serves to check whether recent studies are in accordance with the findings of the present meta-analysis:

Abracen, Looman, Ferguson, Harkins, & Mailloux (2011)

This study compared a TG of 64 sex offenders from an institutional SOTP in Ontario with a CG of 55 untreated offenders from the region's correctional service. The groups were at high risk/need and matched for age at index offence, offender type, psychopathy scores, and risk of recidivism. The treatment was both individual and group based, applied a cognitive-behavioral relapse prevention concept and incorporated the Good Lives Model (e.g. Ward & Maruna, 2007). Outcome was measured by official reoffending with mean follow-up periods of 9.4 (TG) and 11.2 (CG) years. The base rate of sexual reoffending was low (ca. 10%) in both groups. Although the comparison between actual and predicted reoffending was somewhat more favorable in the TG, there was no significant difference in recidivism between TG and CG. It has to be noted that only a handful of offenders reoffended, both groups received other programs beyond SOTP, and there were more offenders with higher risk scores and mental health problems in the TG than the CG.

Comment: This study has the strength of evaluating SOTP in daily practice and with a rather long follow-up period. The findings of overall low rates of sexual recidivism are consistent with our meta-analysis. They also agree with the non-significant effect we observed for custodial treatment.

Grady, Edwards, Pettus-Davis, & Abramson (2012)

This study mainly investigated whether volunteering for treatment has an impact on reoffending. However, the analyses included actual treatment participation as a variable in a Cox regression model that also controlled for a number of variables (e.g. Static-99 risk score, volunteering for treatment, type of sexual offence, pedophilia diagnosis) and thus allowed to estimate a treatment effect, too. 161 sex offenders volunteered and participated in a mainly cognitive behavioral, group based institutional program in North Carolina. Non-participants had either volunteered for treatment but were not selected ($n = 282$) or were eligible but had not volunteered ($n = 443$). The study participants had a moderate-low recidivism risk. Official recidivism was assessed after five

⁵ Abbreviations: SOTP = Sexual offender treatment program; TG = Treatment group, CG = Control/comparison group

years for all study participants. The results indicate a significant reduction in sexual, violent, and non-violent reoffending in the TG.

Comment: The study is not designed to evaluate a treatment effect in the first place and, thus, selection bias is not controlled for, although the analyses applied incorporate a number of relevant control variables. The results differ somewhat from our meta-analytic findings as there is a significant effect in a custodial setting in a rather low risk group.

Letourneau, Henggeler, McCart, Borduin, Schewe, & Armstrong (2013)

This study is a further U.S. based evaluation of MST for young sex offenders; for others see Borduin et al., (1990; 2009) in our MA. It is asked whether positive results in efficacy trials could be replicated and sustained after two years in an implementation in a community mental health center. In a blockwise RCT on juvenile sex offenders (mean age 14.7 years) a TG of 66 young male offenders was compared with a CG of 58 offenders who received ‘treatment as usual’, i.e. mainly group-based CBT interventions. The study reports on a 2-year follow-up for a number of outcomes including official recidivism (rearrests), but differences in sexual reoffending could not be analyzed because of a very low base rate. There was also no significant decrease in rearrests when analyses were controlled for baseline status.

Comment: The randomized design is a clear strength of this study. However as the study does not provide enough ‘hard’ recidivism data it would not influence our results.

Olver, Nicholaichuk, Gu, & Wong (2012)

This study compared a TG of 625 incarcerated sex offenders in Canadian institutions with a CG of 107 sex offenders who did not receive the respective treatment. All programs based on the Canadian standards of the Risk-Need-Responsivity Model. There were some pre-treatment differences between TG and CG (e.g. less singlehood, more unrelated victims, lower risk scores and higher age at release). The authors used a brief actuarial risk scale to assess and control for group differences. A Cox regression controlling for risk found a significant effect on violent reoffending but only a smaller and not significant effect on sexual recidivism. In further analyses treated and untreated offenders were stratified for risk level. These showed that only for the high risk group there was a significant treatment effect on sexual recidivism. In addition, in the TG the time to new sexual offences was longer for treated offenders and the offences committed were somewhat less harmful.

Comment: This is a relatively large study with particular strengths in risk-oriented analyses and differentiated outcome measurement. The overall nonsignificant effect is consistent with our above findings on custodial treatment and large sample sizes. The significant effects for offenders at higher risk are also in accordance with our results.

Smallbone & McHugh (2010)

This study evaluates prison-based treatment in Queensland, Australia. The Queensland prison system offers different treatments according to the risk (medium vs. high) and cultural background of sexual offenders. In total, 158 sexual offenders had attended a treatment program and were compared to 251 untreated sexual offenders with regard to official recidivism (police records) after an average of 29 months. The two groups differed on a number of variables (including risk measures). Treated offenders mostly had moderate-low risk while untreated offenders were at higher risk according to Static-99. Analyses controlling for risk only found a small and non-

significant treatment effect for sexual recidivism and a marginally significant effect for any recidivism.

Comment: The results are basically in line with our findings in that a prison-based treatment of mainly low to moderate risk sexual offenders showed a weak effect only.

Smid, Kamphuis, Wever, & van Beek (2014)

This study from the Netherlands applied a quasi-experimental design to evaluate inpatient treatment for high-intensity sex offenders. The sample consisted of 25% of all convicted Dutch sex offenders that were not referred to any kind of treatment between 1996 and 2002 (CG; $n = 176$) and all convicted Dutch sex offenders of the same time period who received an inpatient sex offender treatment (TG; $n = 90$). The treatment took place in special institutions that contain elements of social-therapeutic prisons and forensic hospitals. The concept is based on the Risk-Need-Responsivity model. The Static-99 was applied to control for nonequivalence in risk between the TG and CG. Outcome was measured by official data on sexual reconviction. The mean follow-up period was 12.33 years. In total, 15% had a sexual and 38% a violent (including sexual) offence. There were some differences in demographic and offence characteristics between the TG and CG. The results showed no overall significant treatment effect on sexual recidivism when regressions controlled for risk level, age and ethnicity. However, there was a marginally significant treatment effect for high-risk offenders. The latter was stronger for violent recidivism in general and untreated sex offenders at higher risk recidivated more frequently and faster.

Comment: This study has various strengths: It evaluated a complex institutional treatment facility outside North America, uses a long follow-up period, applies a risk-related analysis and investigates survival time curves. Although the Cox regressions may not fully control for baseline differences between TG and CG, the findings are in accordance with our MA: The mean recidivism rates were in a similar range, treatment in a custodial setting had no significant effect on sexual recidivism, and the outcome was more favorable for high-risk offenders.

Worling, Littlejohn, & Bookalam (2010)

This is an update of the Worling and Curwen (2000) study included in our meta-analysis. It is less detailed with regard to the subgroups studied and only compares treatment completers vs. a comparison group comprised of non-treated sex offenders as well as treatment dropouts. Thus we decided to retain the “older” study with 10 years follow-up but more differentiated reporting of subgroups that allowed a more sensible intent-to-treat estimate of treatment effects. The update that recurs on a 20-years follow-up shows that the results are virtually unchanged and there were only few additional offenders who recidivated in the 10 years after the first report.

Comment: While the very long follow-up period is a clear strength, the report does fail to meet stricter methodological criteria. It corroborates the results from the shorter follow-up that met the inclusion criteria of our meta-analysis.

APPENDIX 2: CODING SCHEME

In the following the coding variables are listed (in italics) across the different more general domains (general study characteristics, sample, study design, follow-up data, treatment characteristics, comparison group, outcome data).

In case of categorical variables the coding categories are added in brackets.

General study characteristics

- *Publication type* (journal article, book chapter, thesis, unpublished report, conference paper, online presentation, other)
- *Year of publication*
- *Country of origin*
- *Confounding/Dependency of evaluation*: Are the authors involved in the treatment program? (yes, no)
- *Descriptive validity*

Descriptive validity is a concept introduced by Lösel & Köferl (1989) as an extension to Cook & Campbell's (1979) formulation of different threats to validity. It refers to a report's quality in documenting the relevant details of an evaluation. We used 4-point-scales to judge descriptive validity in the areas of:

- Treatment concept
- Treatment realization
- Study design
- Presentation of results
- Overall transparency of report

Sample

- *Original sample size*
 - Treatment group
 - Comparison group
- *Attrition rate*
 - Treatment group
 - Comparison group
- *Mean age*
 - In treatment group
 - In comparison group
- *Age group* (adult, adolescent, mixed)
- *Homogeneity of age* (large, medium, low)
- *Offense types* (separately yes-no-coded whether present in TG: rape, child molestation, incest, exhibitionism, other: specify)
- *Offender risk* (low, medium, high)

Rated on the group level using the items of the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997):

- Earlier conviction for sex offenses: Mean number of previous convictions for a sexual offence (max. 3)
- Age below 25: (Estimated) relative frequency of offenders aged < 25
- Male victims: Relative frequency of offenders with male victims
- Any unrelated victims: Relative frequency of offenders with unrelated victims

The sum of the item scores was then categorized in low risk (≤ 1.5), medium risk (1.5 to 2.5) or high risk (≥ 2.5). If results of other risk instruments are reported these are used to categorize the mean offender risk accordingly if possible. If neither the RRASOR rating can

be applied nor detailed results of another risk instrument are reported but the authors present a risk classification for the offender group this is used to group the study in our risk categories.

- *Referral to treatment* (voluntary, compulsory/both)
- *Reporting of results for treatment dropouts* (reported separately, counted as part of treated group, counted as part of comparison group, partly counted as treated partly as untreated, dropouts not included in analysis, no dropout participants)
- *Percentage of treatment dropout*

Study design

- *Type of Comparison group* (volunteered for treatment, no treatment available at that time or at that place/region, not suitable for treatment/no treatment order, refused treatment, other)
- *Assignment to treatment* (random without matching or stratification, random after individual matching or stratification, nonrandom with individual post-hoc matching, available samples)
- *Group differences* (negligible differences, some relatively unimportant differences, some differences judged to be important, not tested; relevant characteristics such as age, prior offending, type of offenses, psychopathology/personality disorders, risk assessment scores etc.)
- *Statistical control of relevant potential group differences* (yes, no)
- *Internal validity of overall study design according to the Maryland Scale of Methodological Rigor Rating:*
 - (1) No control or comparison group present with regard to recidivism outcomes.
 - (2) Nonequivalent comparison group: Differences on relevant variables effecting recidivism are reported or are to be expected because of assignment strategy (e.g., treatment dropouts or subjects who refuse treatment as CG).
 - (3) Incidental assignment but no serious doubts that assignment resulted in equivalent groups, or sound statistical control of potential differences. The assignment strategy is not related to relevant risk variables (e.g. CG did not receive treatment because it was not offered in that region or at the time). This should be controlled and demonstrated by comparing TG and CG on relevant variables relating to risk of reoffending. There might be small differences in relevant variables but they should not be statistically significant or otherwise substantial. Even if not statistically significant, the magnitude of the differences has to be checked and evaluated for its magnitude, e.g. by taking the variable's SD into account. If there is indication of actual or potential group differences the statistical analyses have to adequately take care for this (e.g. regression methods including relevant variables as control).
 - (4) Matching procedures. Systematic strategy to attain equivalence of the control group (e.g. theoretically sound matching or propensity score techniques). The variables used for matching have to be relevant with regard to differences that actually or potentially arise from treatment assignment for the program under evaluation.
 - (5) Random assignment of treated and untreated subjects. If there is attrition with regard to the recidivism data the study has to be downgraded or even excluded depending on its severity: TG and CG have to remain reasonably well comparable despite the (potential) effects of selective attrition (see Level 3)

Follow up

- *Start of follow-up period* (beginning of treatment, end of treatment, time at risk)
- *Follow-up time* (average in months)
 - Treatment group
 - Control group

Treatment

- *Main treatment approach* (Cognitive/Behavioral/Relapse Prevention, Insight-oriented, Systemic, Therapeutic community, Psycho-educational, Hormonal or otherwise medicinal, Surgical castration)
- *Sex offender specific treatment* (yes, no)
- *Setting* (prison, hospital in-patient, outpatient/ambulatory, mixed)
- *Treatment length as planned (in weeks)*: Estimated by factual average duration of treatment if planned treatment length is not reported
- *Treatment dosage* (5-point scale integrating frequency, length of sessions and therapist/patient ratio; e.g. a short group meeting of an hour any two weeks would be regarded as low dosage)
- *Treatment intensity* (5-point scale rating integrating treatment length and dosage)
- *Format of treatment* (5-point scale: individual only, mainly individual, mixed, mainly group, group only)
- *Aftercare* (obligatory, optional, not offered/not reported)
- *Time of treatment*: Year in which the treatment program was initially established
- *Integrity of treatment implementation* (positive, negative): Will be judged positively if either the description of actual treatment conforms to the treatment concept or if measures like supervision/specialized training of therapists have been taken to ensure treatment integrity; will be judged negatively if problems like insufficient monitoring of therapists, administrative problems, or resistance to treatment by correctional officers are reported
- *Researcher-monitored model project (vs. routine practice)* (yes, no)

Comparison/control group (CG)

- *Treatment of CG* (none, judged as ineffective/placebo, unspecific, psycho-educational, psychotherapeutic, medicinal, other)
- *Intensity/adequacy of CG treatment* (5-point scale; no to full treatment with partial treatment in between, i.e. only restricted treatment or only some CG members treated)

Outcome data (separately coded for each eligible outcome variable reported)

- *Type of recidivism* (sexual, violent, any sexual or violent, nonsexual, neither sexual nor violent, any recidivism, parole violations)
- *Definition of recidivism* (rearrest, reconviction, new charge)
- *Sample sizes for effect size (ES) computation*
 - Treatment group
 - Comparison group
- *Effect size measure*
 - *ES calculation based on* (proportions/frequencies, means and standard deviations, test statistics, p values, other)
 - *Confidence in ES calculation* (5-point scale: highly estimated, moderate, some, slight, no estimation)