Evaluation of sex offender treatment: International findings and perspectives for policy and practice

Friedrich Lösel

International Symposium “Sex Offenders in and out of Crime”, Flemish Royal Academy of Science, 8.3.2018, Brussels, Belgium

Conflict of interest?

- No own treatment program or current delivery of sex offender treatment
- Principal investigator of a project on the evaluation of treatment for sexual offenders (Bavarian Ministry of Justice)

- Advisory roles:
  - Correctional Services Accreditation and Advise Panel, England and Wales
  - Prevention of sexual abuse (Bavarian Ministry of Justice)
  - Investigation of sexual abuse in the Roman Catholic Church in Germany
Background

- Sex offending a highly emotional topic
- Media reports on the most serious cases as a regular 'emotional fuel'
- Over-estimation of crime figures in the general population
- Over-estimation of persistence
- Over-estimation of specialisation
- Increased punitivity/incapacitation and treatment as two parallel paths of policy
- International controversies about the effectiveness of sex offender treatment

Controversies on SOTP between experts

- Seto et al. (2008): No clear evidence that SOTP is effective; more Randomized Controlled Trials (RCTs needed; similar: Hanson (2010), Rice (2010)
- Marshall & Marshall (2010): Critique is wrong; effectiveness of sex offender treatment is sufficiently demonstrated insisting on RCTs is not adequate
- Ho & Ross (2012): Cognitive behavior therapy for sex offenders: Too good to be true?
- Mann et al. (2012): Ho & Ross are wrong. There is sound evidence that SOTP can reduce reoffending
- Ho (2015) wrote again that SOTP does not work; cited Lösel & Schmucker (2005) as a ‘proof’
- Koehler & Lösel (2015) replied to BMJ: argued for a more differentiated perspective
Policy making and evidence

- Crime policy often not based on research
- Example: Penal law reform in Germany in 1998: mandatory treatment of serious sex offenders
- A meta-analysis of SOTP in the German-speaking countries at this time (Lösel, 2000):
  - Only methodological weak treatment evaluations available
  - No significant effect on sexual recidivism; clear effect on non-sexual offending (were programs too unspecific?)
  - Windows of opportunity for scientific influence?
- A sad example from Bavaria
- Establishment of a criminological research unit of the MoJ

The British SOTP 'Scandal' in the Media (Summer 2017)

The scandal of the sex crime 'cure' hubs:
How minister buried report into £100million prison programme to treat paedophiles and rapists that INCREASED reoffending rates

Sex Offender Treatment Programme is psychological group-therapy course Believed to have cost taxpayers over £100 million since it was set up in 1991
Prisoners who take the rehabilitation courses are at least 25 per cent more likely to be convicted of further sex crimes than those who do not. Paedophiles convicted of attacking children are especially likely to offend again

By David Rose for The Mail on Sunday
Published: 22:00 BST, 24 June 2017 | Updated: 02:15 BST, 25 June 2017
The British Study (Mews et al., 2017)

- A large quasi-experimental study of the core SOTP
- 2,562 convicted sex offenders who started SOTP in prison between 2000 and 2012 in England and Wales
- Propensity score matching (PSM) to 13,219 comparison sex offenders using 87 matching factors (but no data on paraphilia)
- Binary reoffending rate for sexual offences low (7.5%), higher for all offences (38.3%)
- Sexual reoffending: Treated sex offenders 10.0%; control group 8.0%
- Child image re-offences: TG = 4.4%, CG = 2.9%
- In other outcomes the matched TG and CG had similar reoffending rates
- However, one study never tells the (full) truth
- Evidence comes by replication

Meta-analyses on SOTP for sex offenders: mean effects (d)
More homogeneous mean effects (d) in meta-analyses on the treatment of general/violent offenders

Why much heterogeneity in SOTP evaluation?

- Fewer controlled studies
- Different eligibility criteria of MAs
- Often small samples & low methodological quality
- Heterogeneous index offences & offender types
- Influence of comorbidities
- Heterogeneity between and within treatment modes (e.g. 'CBT' programs target different risk factors and needs)
- Different outcome criteria & follow-ups
- Due to long follow-up periods often evaluation of 'old' programs

• However, there is also consistency
Meta-analysis of Lösel & Schmucker (2005), JOEX

- 80 comparisons between treatment group (TG) and control group (CG)
- Mean significant Odds Ratio = 1.7 for sexual recidivism
- 6 percentage points (about 35%) reduction of recidivism in TGs
- Cognitive-behavioral treatment (CBT) significant
- Hormonal medication rather positive, but lower methodological quality
- Largest effect: surgical castration, but very weak studies

---

Replicated criminogenic factors that are typically addressed in CBT (Marshall, 2015)*

**Sexual factors**
- Sexual preoccupation
- Sexual preferences for children
- Sexualized violence

**Cognitive factors**
- Emotional congruence with children
- Hostility towards women
- Lack of concern for others
- Offence supportive attitudes

**Relationship problems**
- Lack of intimacy
- Insecure attachment
- Emotional loneliness

**Self-regulation issues**
- Emotional dysregulation

**Low self-esteem/shame**

*See also: Hanson et al. (2013), SONAR etc.*
Other findings of L & S (2005)

- Largest part (45%) of outcome variance due to methodological characteristics (e.g. type of outcome measure, sample size, dropout rate)
- Only few RCTs (with inconsistent results)
- Most studies had non-equivalent CGs (Level 2 of the Maryland Scale)
- No clear relationship between overall design quality and effect size

MA on methodologically better SOTP evaluations:
Schmucker & Lösel (2015, JOEX), (2017, Campbell Collaboration)

- Similar eligibility criteria as in L & S
- Only Maryland Scale Level 5 (RCT), 4 (matching) or 3 (indicators of equivalence)
- Only official recidivism as outcome
- No language restrictions
- Published & unpublished reports
- Scan of more than 3,000 documents
- 27 eligible studies; 29 independent TG-CG comparisons
- N = 10,387 offenders (4,939 treated)
- Mainly cognitive-behavioral programs (n=21)
- No study on hormonal treatment fulfilled design criteria
Time of publication

Comparisons (k)

<table>
<thead>
<tr>
<th>Time of Publication</th>
<th>Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>before 1980s</td>
<td>2</td>
</tr>
<tr>
<td>1980s</td>
<td>4</td>
</tr>
<tr>
<td>1990s</td>
<td>12</td>
</tr>
<tr>
<td>since 2000</td>
<td>14</td>
</tr>
</tbody>
</table>

Maryland Scale of Methodological Rigor

Comparisons (k)

<table>
<thead>
<tr>
<th>Maryland Scale Rating</th>
<th>Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
Follow-up period (months)

Forrest Plot of Effect Sizes for Sexual Recidivism

O = Mean Effect
 --- = Confidence Interval

- Borduin et al., 2001
- Borduin et al., 1990
- McGrath et al., 1998
- Schmid, 1998
- Marshall & Barbaree, 1988 (b)
- Guarino-Ghezzi & Kimball, 1998
- Looman et al., 2000
- Marshall & Barbaree, 1988 (a)
- Proctor, 1996
- Marshall et al., 1997
- Lab et al., 1993
- Taylor, 2000
- Oronn, 2002
- La Mascaza, 2002
- Friendship et al., 2003
- Bakker et al., 1998
- Hanson et al., 1992
- Rice et al., 1991
- Worling & Corwin, 2000
- Marques et al., 2005
- Zeitham, 2003
- Hanson et al., 2004
- Greenberg et al., 2002 (b)
- Nicholaschi, 1996
- Greenberg et al., 2002 (a)
- Romero & Williams, 1993
- Ruddy & Timmermann, 2000
Mean effects (OR)

- Sexual Recidivism: Odds Ratio $k = 28$, OR = 1.41
- Any Recidivism: Odds Ratio $k = 14$, OR = 1.45

'CBT': OR = 1.38 for sexual recidivism

Recidivism rates

(Length of follow up: $M = 5.7$ years)

- Treatment group: 26% reduction
- Control group: 26% reduction

<table>
<thead>
<tr>
<th>Type of reoffending</th>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexual</td>
<td>10.1</td>
<td>13.7</td>
</tr>
<tr>
<td>All offences</td>
<td>32.6</td>
<td>41.2</td>
</tr>
</tbody>
</table>

26% reduction for sexual and all offences.
Maryland Scale rating and effect size

Odds Ratio

Maryland Scale

Effects on sexual recidivism:
Randomized controlled trials only

Odds Ratio

Borduin et al., 2009
Borduin et al., 1990
Ortmann, 2002
Marques et al., 2005
Romero & Williams, 1983
Studies on MST for young offenders

- Two studies with very large effects of Multisystemic Therapy (MST, Henggeler et al., 2009)
- When we excluded the MST studies: overall SOTP effect still significant; OR = 1.38 (vs. 1.41)
- Most evaluations of MST have been carried out by the program developers & some contain methodological problems (Littell et al., 2006)
- Some reviews also suggest better effects of SOTP for young sex offenders (Reitzel & Carbonell, 2006; Walker, 2004), but not such strong effects as MST
- A 2017 review did not find a significant effect of treatment of young sex offenders (7 studies only!)

Sample size and effect size

- No linear relationship ($r = -.05$)
- But small samples ($n \leq 50$) had a significant effect ($p < .05$)
Further methodological moderators

- Larger effects in studies with high descriptive validity (Lösel & Köferl, 1989): transparency of the report etc.; $r = 0.48^*$
- Larger effect in studies with a higher base rate of recidivism: $r = 0.39^*$
- Smaller effects in studies with longer follow up when controlled for recidivism base rate and outliers: $r = -0.34^*$

Group- vs. individual sessions

<table>
<thead>
<tr>
<th>Treatment setting</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>only group</td>
<td>1.01</td>
</tr>
<tr>
<td>mainly group</td>
<td>1.38</td>
</tr>
<tr>
<td>mixed</td>
<td>1.87 *</td>
</tr>
<tr>
<td>mainly individual</td>
<td>1.82</td>
</tr>
<tr>
<td>only individual</td>
<td>3.21 *</td>
</tr>
</tbody>
</table>

$r = .41^*$
**Group vs. individual treatment**

- The superiority of more individualized treatment is not only due to MST: \( r = 0.41 \text{ with MST vs. } 0.31 \text{ without MST} \)
- Ware et al. (2009): an article suggesting advantages of the group format
- However, as the authors mentioned, mainly based on general therapeutic reflections and practical experience
- Our finding supports some caution re. group processes (also Seto et al., 2008)
- Sex offenders may hesitate to talk about intimate issues in group contexts (in a rolling format?)

**Demonstration vs. routine practice**

<table>
<thead>
<tr>
<th>Odds Ratio</th>
<th>Demonstration project</th>
<th>Routine practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.78</td>
<td></td>
<td>1.33</td>
</tr>
</tbody>
</table>

\( ns \)
Author affiliation/involvement

<table>
<thead>
<tr>
<th>Odds Ratio</th>
<th>Author affiliated to treatment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.71</td>
<td>yes</td>
</tr>
<tr>
<td>1.09</td>
<td>no</td>
</tr>
</tbody>
</table>

ns

Mandatory vs. voluntary treatment participation

<table>
<thead>
<tr>
<th>Odds Ratio</th>
<th>Treatment participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.33</td>
<td>voluntary</td>
</tr>
<tr>
<td>1.32</td>
<td>non-voluntary</td>
</tr>
</tbody>
</table>

ns
Effects for different age groups

- Adults only: Odds Ratio = 1.48
- Juveniles only: Odds Ratio = 2.97

Offenders' risk level & effect size

- Low risk: Odds Ratio = 1
- Medium risk: Odds Ratio = 1.33
- High risk: Odds Ratio = 3.95

*p < .01*
Potentially negative effects of imprisonment

- Contagion in criminal subcultures
- Learning more crime skills
- Adopting criminal values
- Defiance & resentment against society
- Loss of positive social bonds
- Resettlement problems
- Stigmatization & cumulative disadvantage

Research:
- No deterrent but a criminogenic effect of prison (Durlauf & Nagin, 2011)
- But no sign. difference in direct comparisons (Villettaz et al., 2015)
- Serious methodological problems
- Isolated view on 'imprisonment'; no attention to content issues such as treatment programs
Meta-analyses on SOTP with data on programs in custody & community (Lösel & Koehler, 2016)

Treatment of substance involved offenders (Lösel & Koehler, 2016)
A few examples of relatively (!) sound recent studies on SOTP in custody:

1. Canada: Olver et al. (2012)
   - Federal incarceration
   - Quasi-experimental design
   - SOTP: Canadian RNR standard
   - TG: n = 612
   - CG: n = 107 (control for actuarial risk)
   - Follow-up: M = 11.7 years
   - Some pre-treatment differences between TG & CG
   - No significant general effect on sexual recidivism, but significance for violent recidivism
   - Sign. effect on sexual recidivism in high risk offenders
   - TG: Longer time to reoffending & less harmful offences

2. Netherlands: Smid et al. (2014)
   - Custody; special therapeutic institutions
   - Quasi-experimental design
   - SOTP: RNR-based and social therapy
   - TG: n = 90
   - CG: n = 176 (not referred to any treatment)
   - Follow-up: M = 12.3 years
   - Some pre-treatment differences between TG & CG
   - Static-99, age and ethnicity as control variables
   - No significant main effect on sexual recidivism
   - Marginally significant effect for high-risk offenders (stronger for violent recidivism)
   - High-risk offenders in the CG recidivated faster
3. USA: Grady et al. (2015)

- State prison-based SOTP in North Carolina
- Quasi-experimental design (following CODC standards)
- SOTP: CBT (SO Accountability & Rehabilitation; SOAR)
- Mainly child-related sexual offences
- TG: n = 256 (propensity score matching; out of 297)
- CG: n = 256 (PSM; out of 3,568)
- Follow-up 4 – 14 years
- No significant difference in survival analyses for sexual and violent recidivism
- Effect on non-violent reoffending within 120 months
- The authors conclude: “The findings generated from this study raise more questions than answers” (p. 22).


- Social-therapeutic prisons in Bavaria (Germany)
- Quasi-experimental design
- Social therapy, group & individual SOTP treatment
- TG: n = 366; CG: n = 331 (no or unspecific treatment)
- Follow-up 4-9 years
- Static-99 as matching variable
- Low base-rate of official sexual recidivism (ca. 7%)
- No significant treatment effect on sexual recidivism
- Tendency (ns) of a small treatment effect in low-risk offenders, but negative tendency (CG<TG) for high-risk offenders
- Significant treatment effect on any reoffending
- Lösel, Link, Endres et al. (2017): Re-analysis using Propensity Score Matching (PSM):
  - Tendency of less sexual recidivism in TG than in CG
  - Methodology matters!
Methodological issues

- Our MAs as well as relatively (!) sound primary studies show both convergent and divergent findings
- Nearly all recent evaluations found very low base rates of (official!) sexual recidivism (ca. 10% +/-5%)
- Difficulty to get significant treatment effects even in larger samples (floor effect)
- A few cases can have a strong impact on results
- Base rates for other kinds of reoffending higher: more often effects
- Dichotomous outcome (‘yes/no’ recidivism) not sensitive
- Indicators such as delayed time of reoffending, frequency and harm more promising
- Many influences beyond program content

Variation of ES between evaluations of CBT programs

All Studies: Maryland Scale Level >= 3

Lösel & Schmucker (2014)
Factors that can have an influence on the effect of treatment evaluations (Lösel, 2012)

Program factors
- Type/details of content
- Quality of delivery/integrity
- Intensity/dosage
- Basic format, group/individual
- Details of delivery (e.g., rolling)
- Content of control condition

Treatment context
- Custody vs. community
- Institutional climate
- Staff competence & motivation
- Therapeutic relationship
- Continuity of support/rapport
- ‘Natural’ protective factors

Evaluation methods
- Quality of evaluation design
- Sample size
- Practice vs. demonstration
- (In)dependent evaluation
- Type of outcome measure
- Length of follow up

Offender factors
- Risk level
- Strengths
- ‘Types’ of offenders
- Personality, comorbidities
- Motivation/denial
- Age, e.g., juvenile, adult

MA on RNR in SOTP (Hanson et al., 2009)

Mean Effect Size (1-Odds Ratio)*

* ORs below 1 showed a positive effect
Perspectives for research, policy and practice

- The evidence of past research on sex offender treatment & evaluation should be the basis for the future
- No ‘silver bullet’ approach or ‘gold standard’ program
- Most promising: CBT approaches, in the community, in a more individualized format, in smaller samples etc.
- More sound evaluations (i.e. RCTs) needed
- However, legal, practical & ethical obstacles
- RCTs do not solve all methodological problems
- New treatment challenges through the ‘migration crisis’

More research and practice on ambulatory treatment options

- Often sex offender treatment in prisons
- When justified for legal reasons and based on thorough risk assessment ambulatory interventions seem to be more promising
- Sex offenders can practice new competences, self control etc. in real life contexts
- Ambulatory treatment & relapse prevention after release important
- Probation or parole; community treatment centers or networks
- Slightly positive results on recidivism in Germany (Kessler & Rettenberger, 2017):
  - Public prosecution data: sex offences: TG=18.6% vs. CG=25.4%; violent offences: TG=14.2% vs. CG=26.9%
  - Reconviction data (more valid): sex offences: TG=9.2% vs. CG=9.3%, violent offences: TG=10.7% vs. CG=13.2%
- More structured concepts & sound evaluations of supervision needed
- Evidence-based combination with technology (electronic monitoring?)
Systems-orientation instead of 'silo' approaches to treatment

- Many offenders have accumulated risk factors, multiple life problems & comorbidities
- Combinations of interventions that address different pathways to reoffending
- E.g. CBT plus accommodation, education, employment
- Evaluation of 'packages' of interventions is more complicated, but practically relevant (see clinical pharmacology)
- Requires theoretical concepts of accumulations and interactions between (causal) risk factors

Differentiation & individualisation

- Often a tendency of 'one size fits all'
- However, heterogeneity of sex offender groups: index offence, offence history, needs, personality
- More attention to specific responsivity
- Use of sex offender typologies based on dynamic risk factors; e.g. Martinez-Čatena et al. (2016)
- Often very intensive assessments, but not much 'translation' into differentiated treatment
- More comparative evaluations of approaches
- Conservation of the strengths of manualized programs, but flexibility with regard to individual needs and circumstances
Relationships and staff skills

- Therapeutic relation as important as the type of treatment (Orlinsky et al., 1998)
- Correlations between staff relationship skills and ES (Gendreau et al., 2005)
- Staff qualification: e.g. ’Skills for Effective Engagement and Development’ (SEED) in UK; Strategic Training Initiative in Community Supervision (STICS) in Canada
- Therapeutic alliances could benefit from more large-scale research on the Good Lives Model, desistance research and some psychodynamic ideas (that have been too much stereotyped)

Process evaluation & implementation

- Often few details on program delivery
- Influence of mixed groups of rapists and child molesters?
- Denial and motivation?
- Impact of a rolling vs. fixed format?
- Optimal degree of manualization?
- Adequate program length?
- Best sequence of treatment in long prison sentences?
- Valid intermediate goals (beyond psychometrics)?
- Descriptive validity & ES: $r=0.48$; Schmucker & Lösel (2017)
- Staff training, supervision, assessment, support; Fixsen et al. (2009)
Institutional context

- Importance of prison climate (already Moos, 1974)
- Regime dimensions are relevant beyond programs (Liebling & Auty, 2015)
- Needed: More knowledge about their impact on programs
- How to reduce negative influences of incarceration?
- If custody is necessary: detailed planning and evaluation of how to ensure transfer of treatment gains to the world outside (aftercare and relapse prevention)
- Institutional factors also relevant in community settings
More integration of neurobiology

- Connectivity between subcortical and cortical regions: emotional input & decision making (Raine, 2013)
- More neurobiological research on sex offenders and pharmacological treatment
- Neurophysiological correlates of sex offending: Mainly small studies, mixed findings (Sanchez & Lösel, 2016)
- SSRIs: very strong deviant fantasies and impulses
- Anti-androgenes: high impulsivity/aggression, psychopathy, sadistic paraphilia?
- SSRIs side effects: anxious, depressive, compulsive symptoms?
- Medication not the main approach in SOTP
- As in treatment of depression, combination of CBT and medication for appropriate cases
- New approaches of neuro-feedback

Evidence-driven development of 'what works' instead of too general controversies

- Too much polarization between different treatment approaches and 'paradigms'
- Too general controversies about the effects of SOTP
- Realistic expectations on the potential of treatment
- Strategies to cope with occasional failure
- Adequate information of policy makers and the general public about facts on sexual offending
- Intervention concepts for special challenges; e.g. migrants with language problems and offence-supportive attitudes
- Key question: What works with whom, in what contexts, under what conditions, with regard to what outcomes, and also why?
References


Thank you!

Contact: fal23@cam.ac.uk