Universität St.Gallen

«We all strive for the best…»

What effects have sanctions and treatments?

Martin Killias, Co-Chair, Campbell Crime & Justice Group

19th Council of Europe Conference of the Directors of Prison and Probation Services
"SHARED GOALS, SHARED VALUES IN PRISON AND PROBATION"
Helsinki, 17-18 June, 2014

Campbell Collaboration: A way to improve our knowledge on effects of interventions

The origins: The search for reliable knowledge

- The *Campbell Collaboration* (named after the famous specialists on evaluation Donald T. Campbell) has been established in England (and not in the USA!) in the late 1990'ies.
- The model was the *Cochrane Collaboration* (named after the British epidemiologist Archie Cochrane) that rapidly became a point of reference in medicine
- Cochrane had started as an informal network of a few pioneers looking for reliable evidence on obstetrics (treatment of pregnancies and births)
- After a short time, consensus was reached that the same principles could be applied to medical interventions in other areas. Dr. Cochrane wondered about the wide-spread ignorance of clinical trials and that nobody took stock of this evidence
- Established in 1993, the Cochrane Collaboration collects, disseminates and updates the lessons from trials conducted across the World on effects of medical interventions in all fields
- Today, it is the uncontested World leader as a source of medical knowledge across all sorts of treatments and all known diseases

From Cochrane to Campbell

- Cochrane had extended its focus to some issues of social interventions, but it remained essentially a medical initiative
- Therefore, it was felt that, after the successes of Cochrane, a similar initiative might be needed in the social sciences that would address the effects of social interventions
- The start came, once more, from the UK
- In 2000, a few leading scholar from different fields of the social sciences met in Philadelphia to establish the “Campbell Collaboration”
- By spring 2014, some 100 systematic reviews have been produced under the Campbell umbrella and published on the Campbell website

- www.campbellcollaboration.org

Internal Organization of Campbell Collaboration

- Campbell is organised in 6 sub-sections. Each is directed by a Steering Committee.

1. Crime and Justice
2. Education
3. International Development
4. Social Welfare
5. Methods Group
6. Users' Group

- The *Crime and Justice* and the *Education Groups* have published so far 36 and 17 "Systematic Reviews"
- They are, thus, the most productive sub-sections, and those that produce most of the studies that are relevant to policy-makers in the fields of crime prevention and corrections
- Systematic reviews focus on experimental or quasi-experimental studies meeting high methodological standards
- As the number of references documents, Campbell becomes more and more the standard in social sciences.

Working principles: What is a systematic review?

- Campbell is a network of independent researchers
- In most cases, researchers contact the Steering Committee of the relevant sub-section and submit a theme (i.e. a title) for a systematic review
- In most cases, the authors have no connection with the Steering Committee
- The proposal of a title is circulated among the Steering Committee, after a short peer-review procedure
- Once approved, the authors are invited to submit a review protocol that will be thoroughly scrutinized through peer-reviews
- The protocol has to indicate what kind of studies will be eligible, how they will be selected and analysed and what categories of studies will be excluded
- Once approved, the researchers start the work that will be peer-reviewed again. During this stage, the respect of the exclusion/inclusion criteria as well as the search strategies will receive primary attention
- Typical exclusion criteria are “out of topic”, missing data or insufficient quality
- Once completed, the SC decides to publish the report on the Campbell website
- Ideally, systematic reviews are due for update every five years

Examples of Campbell Crime & Justice Reviews

1. Formal System Processing of Juveniles: Effects on Delinquency, by Anthony Petrosino, Sarah Guckenburg and Carolyn Turpin-Petrosino (published 10.03.2010)
2. Interventions for Children, Youth, and Parents to Prevent and Reduce Cyber Abuse, by Faye Mishna, Charlene Cook, Robert MacFadden, Michael Saini, Meng-Jia Wu (published 05.06.2009)
3. Effects of Early Family/Parent Training Programs on Antisocial Behavior and Delinquency: A Systematic Review, by Alex R Piquero, David Farrington, Wesley G. Jennings, Richard Tremblay, Alex Piquero, Brandon Welsh (published 27.08.2008)
4. Court-Mandated Interventions for Individuals Convicted of Domestic Violence , by Lynette Feder, Sabrina Austin, David Wilson (30.08.08)
5. Parental Imprisonment: A systematic review of its effects on child antisocial behavior, crime and mental health, by Joseph Murray, David P. Farrington, Ivana Sekol, Rikke F. Olson (published 31.12.2009)
6. Effects of drug substitution programs on offending among drug-addicts , by Egli N, Pina M, Skovbo Christensen P, Aebi MF, Killias M. (published 27.08.09)
7. Systematic Review of Non-Custodial Employment Programs: Impact on Recidivism Rates of Ex-Offenders, Christy A. Visher, Mark B. Coggeshall, Laura Winterfield (03.07.06)
8. Mentoring Interventions to Affect Juvenile Delinquency and Associated Problems, by Patrick Tolan, David Henry, Michael Schoeny, Arin Bass, Peter Lovegrove, Emily Nichols (02.09.2013)

Examples of Campbell Reviews in the Field of Education

1. Indicated Truancy Interventions: Effects on School Attendance Among Chronic Truant Students, by Brandy R Maynard, Katherine Tyson McCrea, Michael S. Kelly (published 05.07.2012)
2. The Effects of Teachers' Classroom Management Practices on Disruptive, or Aggressive Student Behaviour: A Systematic Review, by Regina Oliver, Daniel Reschly, Joseph Wehby (published 24.06.2011)
3. School-Based Education Programmes for the Prevention of Child Sexual Abuse: A Systematic Review, by Karen Zwi, Tracey O`Brien, Paul Tait, Danielle Wheeler, Katrina Williams, Sue Woolfenden (published 19.07.2007)
4. Effectiveness of Adult Employment Assistance Services for Persons with Autism Spectrum Disorders, by John D. Westbrook, Chad Nye, Carlton J. Fong (published 09.03.2012)
5. Post-Basic Technical and Vocational Education and Training (TVET) Interventions to Improve Employability and Employment of TVET Graduates in Low- and Middle-Income Countries: A Systematic Review by Janice Tripney, Jorge Garcia Hombrados, Mark Newman, Kimberly Hovish, Chris Brown, Katarzyna T. Steinka-Fry, Eric Wilkey (published 02.09.2013)
6. Dropout Prevention and Intervention Programs: Effects on School Completion and Dropout Among School-Aged Children and Youth, by Sndra Jo Wilson, Mark Lipsey, Emily Tanner-Smith, Chiungjung Huang Huang, Katarzyna T. Steinka-Fry (published 14.04.2011).

A recent example: *Effects of custodial vs non-custodial sanctions on re-offending*

Example of a systematic review:

Effects on re-offending of custodial vs. non-custodial sanctions

- «Prison is damaging – it negatively affects social bonds (employment, family) and favors, therefore, re-offending» (Bonneville de Marsangy)
- Countless studies have – seemingly – confirmed this statement ever since. Everywhere and at all times, re-offending is more frequent among those sent to prison compared to those who received a non-custodial sanction
- The pitfall: under all systems, defendants with the worst perspectives have a far higher probability to go to prison

➢ Is the worse outcome after prison the result of “treatment” or selection?

➢ www.campbellcollaboration.org (crime & justice, published reviews, 2006)

➢ update currently under review (we shall give a short account here)

Method: Inventory of relevant studies (2006, 2014)

- During the first inventory (Villettaz, Killias, Zoder, 2006), more than 3,000 abstracts of potentially relevant studies published between 1961 and 2002 were located
- ➢ Among these, >300 studies were found potentially relevant and abstracted (coded)
- ➢ For the update (Villettaz, Gilliéron, Killias, 2014), an additional 100 studies published between 2003 and 2013 have been identified
- To be included, a study had to meet the following criteria:

1. It had to include at least two distinct groups: a custodial sanction group and a non-custodial sanction group
2. The sanctions to be compared were imposed following a conviction for a criminal offense
3. There was at least one outcome measure of recidivism (new arrests, re-convictions, re-incarceration or self-report data)
4. The study was completed after 1960 and 2002, and between 2003 and 2013 for the update
5. No restriction about type of publication, geographical area, language, type of delinquency, age, or gender has been applied
6. It had to meet certain methodological standards

Meeting methodological standards: The most important criterion of eligibility

- In the area of sanctions, all studies suffer from an immanent and strong bias: the worse cases go to prison…, better ones get alternative sanctions
- ➢ *The crucial question is how well a study has controlled for these selection factors*
- ➢ *Randomization and natural experiments are the best way to control for this bias*
- ➢ *We have been able to locate 5 RCTs (randomized controlled trials) and 2 natural experiments*

- "Matched-pair" studies (using propensity scores) are another method often applied today. We have located 8 such studies.
- The problem with matched-pair studies is that matching can only be done based on *known factors* of selection (e.g. age, criminal history etc.). Unknown factors (history of alcohol or drug abuse, family and employment record etc.) remain uncontrolled!

- ➢ In order to avoid that the few strong studies are being submerged in a mass of weaker studies, we have performed the meta-analysis separately for

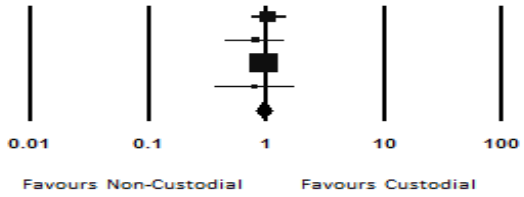
1. RCTs and natural experiments
2. Quasi-Experimental studies using propensity score matching

The remaining studies were not meta-analyzed (lack of interest, the outcome being forseeable given strong bias, great investment in analyzing hundreds of weak studies)

Meta-Analysis (odds ratios, fixed effect), RCTs/NE only

Model	Study name	Subgroup within study	Time point	Statistics for each study					Odds ratio and 95% CI					Weight (Fixed)
				Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00	Relative weight
	Barton 1990	Blank	Blank	1.034	0.733	1.457	0.189	0.850						19.87
	Schneider 1986	Blank	Blank	0.801	0.445	1.443	-0.739	0.460						6.75
	Van der Werff 1979	Combined	Blank	0.947	0.788	1.137	-0.588	0.557						69.61
	Killias 2010	Blank	Combined	0.795	0.361	1.747	-0.572	0.567						3.77
Fixed				0.946	0.812	1.103	-0.709	0.478						

Study name	Subgroup within study	Time point	Statistics for each study					Odds ratio and 95% CI
			Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	
Barton 1990	Blank	Blank	1.034	0.733	1.457	0.189	0.850	
Schneider 1986	Blank	Blank	0.801	0.445	1.443	-0.739	0.460	
Van der Werff 1979	Combined	Blank	0.947	0.788	1.137	-0.588	0.557	
Killias 2010	Blank	Combined	0.795	0.361	1.747	-0.572	0.567	
			0.946	0.812	1.103	-0.709	0.478	



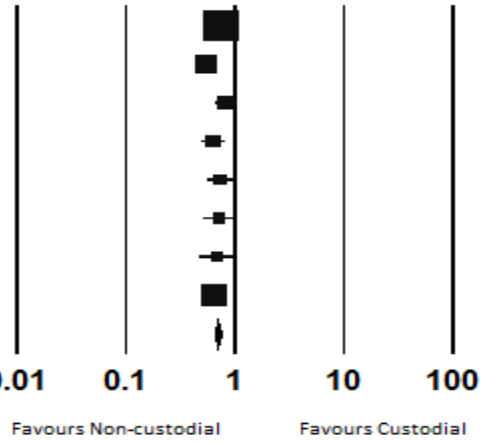
Meta Analysis

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	4	0.946	0.812	1.103	-0.709	0.478	0.752	3	0.861	0.000	0.000	0.032	0.001	0.000
Random	4	0.946	0.812	1.103	-0.709	0.478								

Meta-Analyse (odds ratios, fixed effect), quasi-experiments

Model	Study name	Statistics for each study					Odds ratio and 95% CI					Weight (Fixed)
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00	Relative weight
	Wermink	0.759	0.681	0.847	-4.936	0.000						38.36
	Nieuwbeerta	0.546	0.455	0.656	-6.449	0.000						13.56
	Loughran	0.827	0.654	1.047	-1.577	0.115						8.27
	Nirel	0.635	0.493	0.817	-3.525	0.000						7.16
	Apel	0.740	0.555	0.986	-2.055	0.040						5.54
	Weisburd	0.706	0.514	0.970	-2.151	0.031						4.55
	Bales	0.680	0.478	0.966	-2.154	0.031						3.71
	Lulham	0.640	0.548	0.748	-5.603	0.000						18.85
Fixed		0.693	0.648	0.742	-10.618	0.000						

Study name	Statistics for each study					Odds ratio and 95% CI
	Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	
Wermink	0.759	0.681	0.847	-4.936	0.000	
Nieuwbeerta	0.546	0.455	0.656	-6.449	0.000	
Loughran	0.827	0.654	1.047	-1.577	0.115	
Nirel	0.635	0.493	0.817	-3.525	0.000	
Apel	0.740	0.555	0.986	-2.055	0.040	
Weisburd	0.706	0.514	0.970	-2.151	0.031	
Bales	0.680	0.478	0.966	-2.154	0.031	
Lulham	0.640	0.548	0.748	-5.603	0.000	
	0.693	0.648	0.742	-10.618	0.000	



Meta Analysis

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	8	0.693	0.648	0.742	-10.618	0.000	12.993	7	0.072	46.124	0.009	0.011	0.000	0.096
Random	8	0.684	0.618	0.758	-7.276	0.000								

What do the results show? What do they mean?

- The Meta-Analysis based on RCTs and NE shows a zero-effect
- The Meta-Analysis based on quasi-experiments (using propensity score matching) shows significant, but rather weak negative effects of imprisonment (with 3 exceptions in 9 studies, one not included here)
- Why these contradictory outcomes? Possible explanations:

➢ *Quasi-experiments may not control for a few relevant selection factors*

- Propensity score matching being based on *known* variables (age, gender, criminal history, marital status etc.), it does not take into account for history of alcohol or drug abuse, marital life, work record, behavior in court etc.
- Such variables are likely to affect judicial sentencing decisions. They are likely to be related also to the probability of re-offending

➢ Conclusion:

- Studies that control better such additional factors tend to show a zero-effect
- The less such factors have been controlled, the more likely will the study outcomes show “damaging effects” of imprisonment

➢ *The safe conclusion is that prison does not prevent re-offending better than its alternatives, but it does not increase either the odds of re-offending*

Therapy is not the only purpose of criminal law

- In the medical field, we often find outcomes that do not favor certain interventions (e.g. surgery) over other treatments (e.g. «conservative» therapies)
- In such cases, the zero-effect means that doctors/patients have a margin of decision
- That prison does not do much good, but neither much bad to prisoners could mean that policy-makers keep a margin to consider other priorities but rehabilitation

- Here are a few examples further goals of criminal sanctions:

(1) doing justice to victims and society at large

(2) stabilizing social order

(3) general deterrence

(4) protection of the public (and prevention of future victimizations)

Research offers support to all these goals (including general deterrence, e.g. the brilliant natural experiment by Drago, Vertone & Galbiati in Italy.

Conclusions: a few «Caveats»

- Any systematic review is only as good as the included eligible studies
- Our systematic review suffers most of all from the low number of RCTs
- Even in the absence of RCTs, quasi-experiments should be improved to include, e.g., observational data from the courtroom beyond better biographical data
- None of the studies comparing custodial and non-custodial sanctions has considered the number of offences prevented through incapacitation. These offences should be added to the “zero-effect” in favor of prisons
- Interestingly, few studies have looked at post-sanction family life and work records. In one of the few RCTs that looked at this, former prisoners were slightly better off
- There are some indications that the post-release experience is more important than what has been done (or tried) with prisoners during their incarceration. Moving to other geographic areas seems to be particularly promising

➢ *Future research should look more carefully at post-release experiences*.

➢ Finally, many studies have observed that offending rates decrease during post-sanction periods – independently of the type of sanction.

➢ We have not looked sufficiently into what occurs later on, and why!