



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI

**Eidgenössisches Büro für die
Gleichstellung von Frau und Mann EBG**

Fachbereich Häusliche Gewalt FHG

Die Wirksamkeit von Täterprogrammen zur Prävention von Wiederholungstaten bei häuslicher Gewalt:

Eine narrative Übersicht

Im Auftrag des
Eidgenössischen Büros für die Gleichstellung von Frau und Mann EBG

Erstellt durch :

Véronique Jaquier Erard, Dr. crim. lic. psych.
Centre romand de recherche en criminologie, Universität Neuchâtel

Bern, November 2016

Impressum

Die Wirksamkeit von Täterprogrammen zur Prävention von Wiederholungstaten bei häuslicher Gewalt: Eine narrative Übersicht

Autorin

Véronique Jaquier Erard, Dr. crim. lic. psych.

Centre romand de recherche en criminologie, Universität Neuchâtel

In Zusammenarbeit mit:

Géraldine Bugnon, Dr. soc.

Nicole Egli Anthonioz, Dr. sc. fo

Centre romand de recherche en criminologie, Universität Neuchâtel

Dieser Bericht wurde in französischer Sprache im Auftrag des EBG verfasst. Die darin enthaltenen Einschätzungen und Interpretationen entsprechen nicht zwingend der Sicht der Auftraggeberin.

Bern, 9. September 2016

Der Bericht kann wie folgt zitiert werden:

EBG (Hrsg.): Die Wirksamkeit von Täterprogrammen zur Prävention von Wiederholungstaten bei häuslicher Gewalt: Eine narrative Übersicht. Bern, 2016.

Ausgangslage

Dieser Bericht bietet eine deskriptive Bestandesaufnahme zur Beantwortung der folgenden Fragen:

1. Welche methodische Qualität weisen die Evaluationsstudien im Bereich der Programme für Täter von häuslicher Gewalt auf?
2. Welches sind die wichtigsten Schlussfolgerungen hinsichtlich der Wirksamkeit dieser Programme auf die Verhinderung von häuslicher Gewalt und welche Kriterien wurden angewandt?
3. Welche bestimmten Programm- oder Tätermerkmale können den Erfolg eines Programms positiv oder negativ beeinflussen?

Dieser Bericht ist keine systematische Literaturrecherche. Er umfasst eine Reihe von Studien, die aufgrund ihres methodischen Ansatzes oder ihrer spezifischen Fragestellung ausgewählt wurden. Die Studien von besonderer Bedeutung werden im Anhang

zusammengefasst, und alle beigezogenen Studien sind in der Bibliografie aufgeführt.

Dazu kommen Beispiele aus europäischen und schweizerischen Studien, auch wenn deren methodischer Ansatz für diese Literaturrecherche als ungenügend eingestuft wurde. Angesichts des aktuellen Stands der Evaluationsforschung im Bereich der Täterprogramme befasst sich dieser Bericht ausschliesslich mit den Programmen für Männer, die Gewalt gegen ihre Partnerinnen ausüben. Nicht behandelt wird die Art der Dienstleistungen für die Opfer; es wird jedoch darauf hingewiesen, falls diese Bestandteil der Täterprogramme sind.

Die Schlussfolgerungen dieses Berichts wurden auf der Grundlage der beigezogenen Studien und der verschiedenen evaluierten Programme erstellt. Es ist daher möglich, dass die Ergebnisse nicht mit der aktuellen Situation in der Schweiz übereinstimmen.

Inhaltsverzeichnis

1 Die Täterprogramme	5	Beobachtungsdauer und unregelmässige Bewertungsintervalle	
1.1 Interventionskontexte	6		
1.2 Theoretische Verankerung und Modalitäten der Täterarbeit	7	3.2d Unbekannte externe Validität	20
1.2a Feministisch orientierte psychoedukative Ansätze	7	4 Der Erfolg der Täterprogramme	22
1.2b Kognitive Verhaltenstherapie bei häuslicher Gewalt	7	4.1 Das Kriterium der Wiederholung häuslicher Gewalt	22
1.2c Interventionen in Form von Paartherapie	8	4.1a Ist das Glas halb voll oder halb leer?	22
1.2d Sogenannte individualisierte Interventionen	8	4.1b Worin besteht die Täterarbeit?	23
2 Die Evaluationsforschung im Bereich der Kriminalprävention	10	4.1c Wie viel Zeit soll der Zeit eingeräumt werden?	25
2.1 Evidenzbasierte Strategien und Praktiken	10	4.1d Massgeschneiderte Programme?	25
2.2 Die Designs der Evaluationsstudien und ihre methodischen Qualitätskriterien	11	4.1e Ein wenig Zuckerbrot und viel Peitsche	28
2.3 Die Synthesestudien und die Evaluationsstudien	13	4.1f Kann ein Täter zur Teilnahme verpflichtet werden?	29
3 Die Evaluationsforschung zu den Täterprogrammen	15	4.1g Und wenn die Tätern nicht motiviert sind für eine Verhaltensänderung?	30
3.1 Studiendesigns von unterschiedlicher methodischer Qualität	15	4.1h Ist es zu zweit wirklich besser?	32
3.1a Randomisierte experimentelle Studien	15	4.2 Verhaltens- und Einstellungsänderungen sowie psychologische Veränderungen	36
3.1b Quasi-experimentelle Studien	16	4.3 Eine verzerrte Definition des Erfolgs	37
3.1c Systematische Übersichtsarbeiten und Metaanalysen	17	4.3a Das englische Projekt MIRABAL, eine neue Ära von Evaluationsstudien?	37
3.2 Die wichtigsten methodischen Probleme in Evaluationen von Täterprogrammen	18	5 Schlussfolgerungen	40
3.2a Eine hohe Zahl vorzeitiger Programmabbrüche	18	6 Anhang	43
3.2b Die Wiederholung häuslicher Gewalt als wichtigster Indikator der Veränderung	19	7 Quellenverzeichnis	56
3.2c Eine zu kurze	20		

1 Die Täterprogramme

Jedes Mal, wenn sich ein Mann der Verantwortung zu entziehen schien und mangelnde Empathie zeigte, löste dies verständlicherweise Wutgefühle aus. Wenn hingegen ein Mann aufrichtige Einsicht zeigte, machte sich unter den Frauen der [Fokus-]Gruppe eine ergreifende Mischung aus Hoffnung und Traurigkeit breit: Die Hoffnung, dass Veränderung möglich ist, jedoch nicht in ihrer eigenen Beziehung.¹

Unabhängig vom Kontext kann die Arbeit mit Tätern häuslicher Gewalt nicht losgelöst von der Frage der Sicherheit der Opfer betrachtet werden. Diese Verbindung ist schon deswegen offensichtlich, weil insbesondere in Europa bei der Mehrheit der Programme der Opferschutz zu den Hauptzielen gehört.²

In den angelsächsischen Ländern und insbesondere in Nordamerika resultierten die Täterprogramme aus den Forderungen feministischer Organisationen und von Opferrechtsgruppen. Diese setzten die Regierungen unter Druck, die Polizei- und Gerichtspraxis im Bereich der häuslichen Gewalt durch die Einführung von strengeren Sanktionen gegen die Täter zu ändern. Der Opferschutz blieb zwar das vorrangige Anliegen, doch es zeigte sich rasch, dass ohne Verpflichtung der Täter, die Verantwortung zu übernehmen und das Verhalten zu verändern, keine echten Fortschritte möglich waren.

In mehreren europäischen Ländern haben auch Selbsthilfeprojekte von Männern für Männer eine Rolle gespielt. Sowohl in Deutschland als auch in der Deutschschweiz gibt es einige Beispiele von individuellen oder gruppen-basierten Angeboten mit dem Ziel, Männern, die in ihren Partnerschaften oder Familien gewalttätig sind, Raum für Diskussionen und konkrete Hilfe zu bieten. Mit zunehmender Professionalisierung der

Täterarbeit wurden diese Angebote auch durch psychiatrische (forensische) Kreise unterstützt.³

Täterprogramme gehören heute zwar zu den staatlichen und gerichtlichen Massnahmen zur Bekämpfung häuslicher Gewalt, ihre Ausgestaltung kann aber dennoch sehr unterschiedlich sein. Je nachdem, ob es sich um eine Bedingung für eine Entlassung, eine Ersatzmassnahme, eine strafrechtliche Sanktion oder eine Begleitmassnahme für eine bedingte Entlassung handelt, werden diese Programme zum Teil in sehr unterschiedlichen Kontexten entwickelt. Sie können Teil der Strafjustiz, einer koordinierten Reaktion der Gemeinschaft oder ganz unabhängig sein. Sie beruhen also auf einem ätiologisch unterschiedlichen Verständnis von häuslicher Gewalt.

An dieser Stelle sei an die Istanbul-Konvention⁴ und die darin formulierte Forderung erinnert: «Die Vertragsparteien treffen die erforderlichen gesetzgeberischen oder sonstigen Massnahmen, um Programme einzurichten oder zu unterstützen, die darauf abzielen, Täter und Täterinnen häuslicher Gewalt zu lehren, in zwischenmenschlichen Beziehungen ein gewaltfreies Verhalten anzunehmen, um weitere Gewalt zu verhüten und von Gewalt geprägte Verhaltensmuster zu verändern» (Art. 16 Abs. 1). Massnahmen, deren vorrangiges Anliegen die Sicherheit und die Unterstützung der Opfer ist, müssen folglich in Koordination mit den Institutionen, die Dienstleistungen für Opfer anbieten, umgesetzt werden (Art. 16 Abs. 3).

Die Täterprogramme weisen somit die Besonderheit auf, eine Intervention sowohl auf individueller als auch auf gemeinschaftlicher und struktureller Ebene zu sein, da sie gegen

¹ Elinoff Acker (2013: 185) «Any time an abusive man seemed to sidestep full responsibility or showed a lack of empathy, it triggered understandable feelings of anger. In contrast, when the men demonstrated genuine insight, it brought forward a poignant mixture of hope and sadness for the women in the [focus] group : that change is possible, though not for their own relationships.»

² Kelly & Westmarland (2015b).

³ Für die Evaluationen in Nordamerika siehe Chung (2015) und Shepard & Pence (1999); für die Evaluationen in Europa siehe Hester, Lilley, O'Prey & Budde (2014) und spezifisch für die Schweiz siehe Egger (2008).

⁴ Übereinkommen Nr. 210 des Europarats zur Verhütung und Bekämpfung von Gewalt gegen Frauen und häuslicher Gewalt, Istanbul, 11. Mai 2011. Von der Schweiz am 11. September 2013 unterzeichnet, derzeit ist der Ratifikationsprozess im Gang.

die Straffreiheit von häuslicher Gewalt stehen. In der Praxis zeigt sich dies in der – mehr oder weniger ausgeprägten – Pflicht, mit betroffenen Personen zu arbeiten und häusliche Gewalt als strukturelles Problem wahrzunehmen.⁵

1.1 Interventionskontexte

Auf europäischer Ebene lassen sich die Kontexte, in denen Täterprogramme durchgeführt werden, nicht immer klar festlegen. Es gibt Programme, die im Gefängnis stattfinden, oder solche, die von den Bewährungsdiensten für verurteilte Täter angeboten werden. Als Voraussetzung für eine bedingte Entlassung oder als Ersatzmassnahme werden diese Programme im Allgemeinen von sozialmedizinischen Institutionen oder unabhängigen Vereinen durchgeführt. Ordnet eine Behörde das Programm an, so wird diese über die regelmässige Anwesenheit der zugewiesenen Teilnehmer informiert. Bei einem Abbruch oder bei Verstössen gegen die Regeln des Programms sind grundsätzlich Sanktionen vorgesehen. In der Praxis zeigt sich jedoch, dass diese nicht immer durchgesetzt werden. Schliesslich weisen gewisse Programme eine Verbindung zur Strafjustiz auf, die beispielsweise einen Täter als Bedingung für die Einstellung des Strafverfahrens einem Programm zuweisen kann, während andere Programme völlig systemunabhängig sind.⁶ Die Klienten nehmen freiwillig teil, können aber

auch durch andere Organisationen oder Stellen zugewiesen werden. Oft werden die Programme von Gesundheits- oder Sozialhilfebehörden oder durch Nichtregierungsorganisationen finanziert. Daher kommt es häufig vor, dass die Teilnehmer derselben Gruppe nicht alle die gleichen Zugangsvoraussetzungen aufweisen.

Diesen verschiedenen Kontexten stehen verschiedene Traditionen der Täterarbeit gegenüber. Das «Duluth-Modell», das Mitte der 1980er-Jahre in Minnesota entwickelt wurde, ist eine koordinierte Reaktion der Gemeinschaft auf häusliche Gewalt.⁷ Die Täterprogramme, die den Hauptbestandteil dieses Modells bilden, umfassen eine Reihe von Massnahmen zur Verhütung von Wiederholungstaten und zur Gewährleistung der Sicherheit der Opfer sowie weitere Massnahmen wie die Verhaftung des Täters, strafrechtliche Sanktionen bei Nichteinhaltung der Programmregeln, Massnahmen zur Unterstützung der Opfer und Zusammenarbeit der verschiedenen Institutionen.⁸ Die Grundkonzepte mehrerer Interventionsprogramme in der Deutschschweiz basieren direkt auf dem Duluth-Modell.⁹ Dies ist der Fall beim Projekt «Halt Gewalt» von den Interventionsstellen beider Basel, welches von Bern übernommen wurde.¹⁰

In der Deutschschweiz wurden verschiedene Interventionsprogramme ausgehend vom Hamburger-Modell (Gewaltberatung/ Gewaltpädagogik GHM) mit dem Selbsthilfeansatz entwickelt. Die GHM-Beratungszentren boten zunächst offene und zeitlich unbefristete Gruppentrainings an. Daraus entwickelte sich

⁵ Chung (2015).

⁶ Hester & Lilley (2014) für eine Zusammenfassung und Hester *et al.* (2014) für detaillierte Beispiele. In der Schweiz können die Täterprogramme (1) von der Justiz angeordnet werden und zwar in Form: (a) einer Ersatzmassnahme anstelle der Untersuchungs- oder der Sicherheitshaft (Art. 237 StPO); (b) einer Bedingung für eine bedingte Strafe in Form einer Weisung (Art. 44 Abs. 2 und 94 StGB); (c) einer Voraussetzung für eine bedingte Entlassung in Form einer Weisung (Art. 87 Abs. 2 StGB); oder (d) von Weisungen, die im Rahmen der Ausgestaltung des Besuchsrechts von der Justiz- oder der Kindes- und Erwachsenenschutzbehörde ausgesprochen werden (Art. 307 ZGB). Sie können (2) von der Justiz vorgeschlagen werden im Rahmen (a) eines Vergleichs im Sinne von Art. 316 StPO; oder (b) einer Einstellung des Verfahrens im Sinne von Art. 55a StGB. Siehe insb. Moreillon & Druey (2012) und Mösch Payot (2012).

⁷ Domestic Abuse Intervention Project, www.theduluthmodel.org.

⁸ Es besteht ein gewisses Missverständnis, da sich viele Programme auf das Duluth-Modell berufen, obwohl sie nur das Lernprogramm, nicht aber sämtliche Grundsätze anwenden, die der koordinierten Reaktion der Gemeinschaft auf häusliche Gewalt zugrunde liegen.

⁹ Für eine erste Anpassung siehe Logar, Rösemann & Zürcher (2002), zitiert nach Egger (2008).

¹⁰ Erfasst von Egger (2008); für die Evaluation siehe Gloor & Meier (2002, 2003).

ein strukturiertes Lernprogramm, das auf einer (eher) psychotherapeutischen Basis aufbaut.¹¹

Schliesslich macht ein Blick auf das Vereinigte Königreich die Rolle deutlich, welche die Bewährungsdienste in gewissen Rechtsordnungen bei der Entwicklung von Täterprogrammen spielten. Diese Dienste arbeiteten zunächst mit Gemeinschaftsinstitutionen zusammen, indem sie die Programme dieser Institutionen für freiwillige und von der Justiz zugewiesene Teilnehmer finanzierten. Die Bewährungsdienste in England und Wales beschlossen in der Folge, ihre eigenen obligatorischen Programme zu erarbeiten und distanzieren sich dabei von der Opferzentrierung, was nicht ohne gewisse Spannungen über die Bühne ging.¹² Es überrascht daher nicht, dass die nationalen Akkreditierungsstandards heute von den akkreditierten Organisationen verlangen, dass sie parallel zu den Dienstleitungen für die Täter häuslicher Gewalt auch Dienstleistungen für die Opfer anbieten.¹³ Die von den Bewährungsdiensten entwickelten Programme, wie z.B. das Zürcher Lernprogramm¹⁴, orientieren sich oft an den in England und Nordamerika (vor allem Kanada) angewandten evidenzbasierten Ansätzen.¹⁵

1.2 Theoretische Verankerung und Modalitäten der Täterarbeit

Gibt es gleich viele Täterprogramme wie Kategorien, denen sie zugeordnet werden? Bei einem Blick auf die Vielzahl der Programme, welche die ebenso zahlreichen Erklärungen zu häuslicher Gewalt und deren Präventionsmöglichkeiten zu widerspiegeln scheinen, ist dieser Eindruck nicht aussergewöhnlich. In diesem Bericht wird insbesondere zwischen

vier Ansatzgruppen mit verschiedenen theoretischen Grundlagen unterschieden.¹⁶

1.2a Feministisch orientierte psychoedukative Ansätze

Die psychoedukativen Ansätze im Bereich der häuslichen Gewalt werden als feministisch bezeichnet, wenn sie sich auf eine gesellschaftspolitische Analyse der häuslichen Gewalt und der ihr zugrunde liegenden strukturellen Mechanismen stützen, das heisst auf die Geschlechterungleichheiten und die patriarchale Ideologie. Die von den Männern ausgeübte Gewalt wird als eine Reihe von Taktiken verstanden, die dazu dienen, Kontrolle über die Frauen auszuüben.

Die auf feministisch orientierten psychoedukativen Ansätzen beruhenden Programme gründen im Duluth-Modell.¹⁷ Sie zielen auf eine Auseinandersetzung mit den männlichen Wertvorstellungen, Glaubenssätzen und Verhaltensweisen, die häusliche Gewalt begünstigen, verharmlosen oder entschuldigen, insbesondere indem gleichberechtigte und respektvolle Beziehungen modelliert werden. Im Zentrum der Gruppenarbeit steht die Verantwortungsübernahme der Täter, dem Schutz und der Sicherheit der Opfer (und ihrer Kinder) wird zentrale Bedeutung beigemessen.¹⁸

1.2b Kognitive Verhaltenstherapie bei häuslicher Gewalt

Einer der bekanntesten Ansätze, die kognitive Verhaltenstherapie, ist jedoch nicht der einzige psychotherapeutische Ansatz, der häusliche Gewalt als Ausdruck verschiedener Dysfunktionen des Täters auf psychischer und emotionaler Ebene betrachtet. Ihre Beliebtheit

¹¹ Oelemann & Lempert (1995), zitiert nach Egger (2008).

¹² RESPECT (2012).

¹³ Phillips, Kelly & Westmarland (2013).

¹⁴ Nach Egger (2008); für die Evaluation siehe Bächli-Biétry (2006).

¹⁵ Für eine ausführlichere Beschreibung siehe Egger (2008).

¹⁶ Siehe insb. Chung (2015) und Gondolf (2011) sowie Mackay, Gibson, Lam & Beecham (2015: 11-3) für eine hervorragende Zusammenfassung. Für einen Überblick über das Angebot in der Schweiz siehe Egger (2008) und bis zu einem gewissen Grad Hester *et al.* (2014); für Europa siehe Hamilton, Koehler & Lösel (2013) und Hester *et al.* (2014).

¹⁷ Siehe Fussnote 7.

¹⁸ Flood & Pease (2009).

ist sicher auch der geringeren Bekanntheit der anderen Ansätze zu verdanken.

Als Therapieform oder als Bestandteil einer umfassenderen Intervention¹⁹ versteht die bei häuslicher Gewalt angewandte kognitive Verhaltenstherapie das gewalttätige Verhalten als erlerntes Verhalten, das im Rahmen eines besonderen Prozesses «verlernt» werden kann. Sie will die Teilnehmer dazu bringen, die zur Gewalttätigkeit beitragenden Gedanken, Glaubenssätze und Gefühle zu identifizieren, und ihnen helfen, neue Strategien zur Verhaltensänderung zu erlernen. Die auf kognitiver Verhaltenstherapie basierenden Programme legen den Schwerpunkt auf die wiederholte Anwendung neuer Kompetenzen oder Fähigkeiten, denn die Täter sollen durch die feststellbaren positiven Auswirkungen auf ihr Umfeld zu einem nicht gewalttätigen Verhalten motiviert werden (positive Verstärkung).²⁰

1.2c Interventionen in Form von Paartherapie

Die Vertreterinnen und Vertreter der Paar- oder Familientherapie sehen in diesen Ansätzen Alternativen, die in bestimmten Kontexten häuslicher Gewalt angemessen sind, insbesondere wenn sich ein Paar nicht trennen will oder bei sogenannter «situativer» Gewalt, um es mit den Worten von Michael Johnson zu sagen.²¹

Diesen Interventionen für Paare, die sich auf verschiedene theoretische Ansätze berufen, ist gemeinsam, dass die häusliche Gewalt als Folge einer nicht funktionierenden Beziehung und die therapeutische Arbeit als ein Mittel zur Behandlung der Ursachen betrachtet werden.²²

Einige Fachleute stehen diesen Interventionen allerdings sehr kritisch gegenüber, da diese ihrer Ansicht nach das Risiko für die Opfer

erhöhen, der für häusliche Gewalt charakteristischen Ambivalenz nicht Rechnung tragen oder die Verantwortung für die vom Täter begangenen Gewalttaten gar teilweise auf das Opfer übertragen.²³

Viele Rechtsordnungen in Nordamerika und Europa raten daher davon ab oder lehnen Paartherapieprogramme als staatlich finanzierte Präventionsmassnahmen ganz ab.²⁴

1.2d Sogenannte individualisierte Interventionen

Schliesslich kann eine Reihe von Ansätzen unter der Kategorie der sogenannten «individualisierten» Interventionen zusammengefasst werden. Dabei wird die häusliche Gewalt auf eine Vielzahl von Erklärungsfaktoren sowohl psychologischer und psychiatrischer als auch psychologischer und soziologischer Art zurückgeführt. Es werden Interventionsprogramme empfohlen, die auf die spezifischen Bedürfnisse der Täter ausgerichtet sind. Dieser Grundsatz der Individualisierung ist bei den Vertreterinnen und Vertretern der traditionelleren Interventionsmodelle nicht sehr beliebt, da sie deren Ansicht nach die strukturellen Mechanismen, die der häuslichen Gewalt zugrunde liegen, in Abrede zu stellen.²⁵

Zu dieser Kategorie zählen die Programme für Täter, die aufgrund ihrer besonderen Verhaltensweisen, Einstellungen oder Persönlichkeitsmerkmale verschiedenen Tätertypologien zugeordnet werden. Das sind Programme, die auf typologischen Studien wie jene von Amy Holtzworth-Munroe oder Michael Johnson basieren.²⁶

Andere Programme berufen sich auf das in den 1980er-Jahren entwickelte kanadische

¹⁹ Mackay et al. (2015).

²⁰ Babcock, Green & Robie (2004) und Eckhardt, Murphy, Black & Suhr (2006).

²¹ Johnson (2008).

²² McCollum & Stith (2008) und Stith et al. (2004, 2011); siehe auch weiter unten S. 32 ff.

²³ Babcock et al. (2004) und Gondolf (2012: 147-52).

²⁴ Stith et al. (2004). In Spanien beispielsweise werden systemische Interventionen ebenfalls nicht empfohlen oder sind gar verboten (Carbajosa, Boira Sarto, 2013).

²⁵ Day, Chung, O'Leary & Carson (2009a).

²⁶ Z. B. Holtzworth-Munroe & Meehan (2004) und Johnson (2008).

Modell «Risk-Need-Responsivity» (RNR), das seither in der Rehabilitation von Straftätern breit eingesetzt wird.²⁷ Das auf den Grundsätzen Risiko, Bedarf und Lernfähigkeit beruhende Modell von Don Andrews und James Bonta empfiehlt, die Intensität des Programms auf das Rückfallrisiko des Täters, die Art des Programms auf den Rehabilitationsbedarf des Täters und den Programmablauf auf die Lernkompetenzen und -fähigkeiten des Täters auszurichten. Angesichts der empirischen – und politischen – Unterstützung, die das RNR-Modell genießt, sehen einige darin einen idealen Rahmen, um die Täterprogramme zu «verwissenschaftlichen».²⁸

Schliesslich beinhalten mehrere Programme die Voraussetzungen für Arbeiten gemäss dem Transtheoretischen Modell, das aus der Gesundheitspsychologie stammt und davon ausgeht, dass bei jeder Intervention die Veränderungsbereitschaft des Täters zu berücksichtigen ist.²⁹ Sie empfehlen, nicht nur den Inhalt, sondern auch den Programmablauf an das jeweilige Bereitschaftsstadium anzupassen, um die Wirksamkeit der Intervention zu erhöhen.³⁰ Daneben gibt es verschiedene Ansätze, welche die Motivierende Gesprächsführung (Motivational Interviewing) beinhalten. Dabei handelt es sich um eine im Suchtbereich entwickelte therapeutische Strategie. Die mit dem Transtheoretischen Modell vergleichbare Motivierende Gesprächsführung geht davon aus, dass die Betroffenen individualisierte Unterstützung benötigen, die auf das Stadium der Veränderungsbereitschaft, in dem sie sich befinden, ausgerichtet ist.³¹ All diese Ansätze messen der Qualität der Beziehung, die sich zwischen Patient/Klient und Therapeutin oder Therapeut aufbauen muss, einen besonderen Stellenwert bei.

Von Nordamerika über den europäischen Kontinent bis nach Australien kombinieren heute die meisten Täterprogramme verschiedene Ansätze und Techniken.³² Während bei den psychoedukativen Ansätzen nach dem Duluth-Modell bewährte therapeutische Techniken zur Veränderung der Faktoren gewaltverstärkenden Verhaltens zur Anwendung gelangen, beinhalten die kognitiv-verhaltenstherapeutischen Programme Aktivitäten, welche die patriarchalen Haltungen zur Rechtfertigung von Gewalt gegen Frauen in Frage stellen.³³

Es existieren unendlich viele Kombinationen, so dass es manchmal schwierig ist, zwischen den verschiedenen Ansätzen zu unterscheiden und sie den entsprechenden Kategorien zuzuordnen.³⁴ Zudem können die Struktur und der Ablauf eines Programms unabhängig vom geltend gemachten Ansatz und der Bezeichnung des Programms variieren. Dies wird dann problematisch, wenn es um die Evaluation der Wirksamkeit der Programme geht, denn während das Angebot an Täterprogrammen kontinuierlich ansteigt, ist im Bereich der Evaluationsstudien keine vergleichbare Entwicklung zu erkennen.

²⁷ Andrews & Bonta (2010).

²⁸ Radatz & Wright (2015).

²⁹ Prochaska & DiClemente (1992); siehe weiter unten S. 30 ff.

³⁰ Begun, Shelley, Strodthoff & Short (2001) und Day *et al.* (2009a).

³¹ Mcmurran (2009).

³² Babcock *et al.* (2004) und Mackay *et al.* (2015).

³³ Für eine interessante Diskussion zum Eklektizismus im Bereich der Therapie von Tätern häuslicher Gewalt siehe den Text von Vlais (2014).

³⁴ Z. B. das Programm «Creating a Process of Change for Men Who Batter», das im Rahmen des Duluth-Modells entwickelt wurde, wird von den einen als psychoedukativer Ansatz (Dutton, Corvo, 2006; Eckhardt, Murphy, Whitaker, Sprunger, Dykstra, Woodard, 2013b) und von den anderen als kognitiv-verhaltenstherapeutischer Ansatz, der die Gender-Aspekte beinhaltet, angesehen (Gondolf, 2012).

2 Die Evaluationsforschung im Bereich der Kriminalprävention

Diese Wissenschaftlerinnen und Wissenschaftler lehnten die professionelle Ideologie des «Nichts funktioniert» ab und verfolgten stattdessen einen strengen wissenschaftlichen Ansatz, um zu zeigen, dass die gängigen Strafinerventionen wirkungslos sind, die Kriminellen nicht hoffnungslos verloren sind und die Behandlungsprogramme, die auf kriminologischem Wissen basieren, Rückfälle signifikant senken können.³⁵

Während das Angebot an Täterprogrammen stetig zugenommen hat, lässt sich im Bereich der Evaluationsforschung keine vergleichbare Zunahme der Studien beobachten. Fast alle Evaluationsstudien befassen sich mit gerichtlich angeordneten Programmen oder mit solchen, an denen sowohl zugewiesene als auch freiwillige Täter teilnehmen. Die Studien wurden mehrheitlich in den USA durchgeführt. Die Evaluation der Einzeltherapie oder der Paartherapie ist bisher eher anekdotenhaft und erfolgt meistens in Form von Evaluationsstudien, die diese Interventionen mit Gruppentherapien vergleichen.

Die politischen Herausforderungen im Bereich der Kriminal- und Rückfallprävention im Allgemeinen und im Bereich der häuslichen Gewalt im Besonderen erfordern das Verständnis des aktuellen Kontexts der gemeinschaftlichen und gerichtlichen Praktiken – ein Kontext, in dem immer häufiger das evidenzbasierte Evaluationskonzept zur Anwendung gelangt – damit die Wirksamkeit der Täterprogramme ermittelt werden kann.

Aufgrund der Vielfalt der Täterprogramme und der grossen Variabilität der methodischen Designs ihrer Evaluationen braucht es für die Feststellung dessen, «was funktioniert», zwingend eine strenge Überprüfung der

methodischen Qualität der verschiedenen Studien. Einleitend werden in der Folge, deshalb einige Leitlinien im Bereich der Evaluationsforschung skizziert, um die Analyseschlüssel zur Feststellung der methodischen Qualität der verschiedenen Studien bereitzustellen, mit denen die Wirksamkeit der Täterprogramme evaluiert wurde.

2.1 Evidenzbasierte Strategien und Praktiken

Im Bereich der Programme zur Kriminal- und Rückfallprävention war in den 1990er-Jahren eine schrittweise Aufgabe der pessimistischen Haltung des «nothing works» der 1970er-Jahre zu beobachten³⁶, die durch einen differenzierteren Ansatz abgelöst wurde. Doch mit der wachsenden Vielfalt der Programme wurde rasch klar, dass ihre Wirksamkeit einem Vergleich unterzogen werden musste. Vor diesem Hintergrund ist der Erfolg des «evidence-based programming» zu verstehen, das aus der biomedizinischen Forschung zur Ermittlung der Wirksamkeit und der Sicherheit einer medizinischen Intervention hervorgegangen ist. Ohne eine vergleichende Evaluation der Gruppe, die in den Genuss einer Intervention kommt (z. B. medikamentöse Behandlung, therapeutisches Programm) mit einer Gruppe, die keine Intervention erhält, und ohne zufällige Einteilung in diese Gruppen (Randomisierung) kann nicht mit Sicherheit bestimmt werden, ob die Intervention bei den Personen, bei denen sie angewandt wurde, eine signifikante Veränderung bewirkt hat.

Die Verfolgung eines evidenzbasierten programmatischen Ansatzes bedeutet, dass die erarbeiteten und umgesetzten Programme auf der systematischen Evaluation ihrer Wirkung in verschiedenen Situationen und bei verschiedenen Personen beruhen.³⁷ Die politischen Behörden und die Kostenträger

³⁵ Cullen (2005: 1): «These scholars rejected the 'nothing works' professional ideology and instead used rigorous science to show that popular punitive interventions were ineffective, that offenders were not beyond redemption, and that treatment programs rooted in criminological knowledge were capable of meaningfully reducing recidivism.»

³⁶ Gemäss Lipton, Martinson & Wilks (1975); für eine Analyse der Verdinglichung des «nothing works» siehe Cullen (2005).

³⁷ Für weitere Einzelheiten siehe insb. Coalition for Evidence-Based Policy Working Group (2006) oder Robert (2009).

suchen vor allem angemessene Interventionen, die an den lokalen Kontext angepasst werden können und wirksam, effizient und nachhaltig sind. Dies ist insbesondere der Fall in einem angespannten sozioökonomischen Klima, das eine vernünftige Verwendung der beschränkten Ressourcen erforderlich macht.

Viele Interventionsprogramme müssen sich daher «bewähren». Einige wurden jedoch nie oder nur anekdotenhaft evaluiert, und andere wurden von denjenigen Personen evaluiert, die sie erarbeitet hatten, wodurch die Unabhängigkeit nicht gewährleistet war. Zweifelsohne liegen vielen Programmen lobenswerte Absichten zugrunde, doch solange sie nicht evaluiert wurden, kann ihre Wirksamkeit nicht als gesichert gelten. Zudem kann es im Bereich der Kriminalprävention vorkommen, dass Programme aus politischen oder ideologischen Gründen umgesetzt werden, was keine ausreichende Garantie für ihre Wirksamkeit ist.

Schliesslich ist zu bedenken, dass die Wirksamkeit einer Intervention oder eines Programms nicht als dichotomes Konzept zu verstehen ist. Vielmehr ist sie in einem Kontinuum zu sehen, und zwar in einer dynamischen Perspektive. Wenn also eine Evaluationsstudie zum Schluss kommt, dass ein Programm wirkungslos ist, bedeutet dies nicht, dass das für immer so bleibt. Und umgekehrt heisst das, dass die Programme wiederholt evaluiert werden müssen, zumal sie oft geändert werden, um an einen besonderen Kontext oder an besondere Umstände angepasst zu werden. Die Evaluationen dürfen daher nicht auf die «klassische» Version eines Programms beschränkt werden, sondern müssen auch die abgeänderten Versionen berücksichtigen.

Diese Bewegung in Richtung evidenzbasierte Evaluation lässt sich auch bei den Täterprogrammen erkennen und hat ihren Ursprung in den Evaluationsstudien über die Wirksamkeit der Programme zur Rückfallprävention im Allgemeinen. Auf europäischer Ebene wird diese Entwicklung

insbesondere durch zwei internationale Projekte illustriert, die durch das Daphne-Programm der Europäischen Kommission finanziert wurden: Das Projekt «WORK WITH PERPETRATORS» (WWP, 2006–08), das einen Austausch der bewährten Vorgehensweisen im Bereich der Interventionsprogramme zum Ziel hatte, und das Projekt «IMPACT» (2013–14), das die Entwicklung von methodischen Instrumenten und Standards zur Harmonisierung und Förderung der Evaluation von Täterprogrammen bezweckte.³⁸

2.2 Die Designs der Evaluationsstudien und ihre methodischen Qualitätskriterien

Angesichts der Ziele dieses Berichts unterscheiden wir drei Typen von Evaluationsstudien nach ihrem methodischen Design.

- a Die **Studien ohne Kontrollgruppe** messen zumeist bei den Programmteilnehmern vor Beginn des Programms und nach dessen Abschluss einen oder mehrere Indikatoren und vergleichen anschliessend die Werte vor mit jenen nach der Intervention, um allfällige Veränderungen festzustellen. Diese Studien sind für ein Evaluationsvorgehen kaum geeignet, da damit die Faktoren, die bei den Programmteilnehmern allfällige Veränderungen herbeigeführt haben, nicht bestimmt werden können, und der Erfolg nicht dem fraglichen Programm zugeschrieben werden kann.
- b Die **quasi-experimentellen Studien** vergleichen das Verhalten der Programmteilnehmer mit jenem von Personen, die nicht am Programm teilgenommen haben. Damit können zwar die Unterschiede zwischen den beiden Gruppen aufgezeigt werden, es kann aber keine Aussage zur Wirksamkeit des

³⁸ Für die verschiedenen im Rahmen dieser Programme entwickelten Dokumente siehe die Internetseite: www.work-with-perpetrators.eu.

Programms gemacht werden, da weder allfällige vorbestehende Unterschiede zwischen den Gruppen ausgeschlossen noch Drittvariablen überprüft werden können.

- c Die **randomisierten experimentellen Studien** gelten als Referenz im Bereich der Interventionsevaluation. Sie beinhalten die zufällige Zuteilung der Teilnehmer zur Experimentalgruppe (an der eine Intervention durchgeführt wird) und zur Kontrollgruppe (die keine Intervention erhält). Die Randomisierung maximiert die Wahrscheinlichkeit, dass die beiden Gruppen in Bezug auf die Variablen, die aufgrund ihres wahrscheinlichen Einflusses auf die mit dem Programm angestrebten Veränderungen festgelegt wurden, gleichwertig sind.

Das Forschungsdesign einer Studie ist fester Bestandteil ihrer methodischen Qualität. Die Ermittlung der Qualität einer Evaluationsstudie ist zentral, um die Aussagekraft der Rückschlüsse von Ursache und Wirkung zu bestimmen. Für diese Untersuchung müssen fünf verschiedene Validitätskriterien berücksichtigt werden.³⁹

- Die **Validität der statistischen Inferenz**, das heisst die Existenz einer statistischen Beziehung zwischen der Teilnahme am Programm und den beobachteten Veränderungen (z. B. Wiederholung der häuslichen Gewalt).
- Die **interne Validität**, das heisst die Tatsache, dass die Teilnahme am Programm – und nur sie – die beobachtete Veränderung bewirkt hat.
- Die **Konstruktvalidität**, das heisst die Übereinstimmung zwischen den Definitionen und den Operationalisierungen der theoretischen Konstrukte, die dem

Programm und den zu messenden Veränderungen zugrunde liegen.

- Die **externe Validität**, das heisst das Ausmass, in dem die in einer besonderen Studie beobachtete Wirkung eines Programms auf andere Kontexte und Personen übertragen werden kann.
- Die **deskriptive Validität**, das heisst die Übereinstimmung mit den in der Forschung für die Evaluation beschriebenen zentralen quantitativen und qualitativen Elementen.

Die (randomisierten und nicht randomisierten) experimentellen und quasi-experimentellen Evaluationsstudien weisen den höchsten Grad an statistischer und interner Validität sowie an Konstruktvalidität auf. Die Ermittlung der externen Validität beruht auf der Replikation dieser Studien in vielfältigen Kontexten.

Im Bereich der Evaluationsforschung zur Wirksamkeit der Täterprogramme sind folgende **Validitätsprobleme** zu beachten.

Die *statistische Validität* einer Evaluationsstudie ist nicht mehr gewährleistet, wenn ihre statistische Aussagekraft ungenügend ist (z. B. zu kleine Stichprobe) oder wenn sie auf ungeeigneten statistischen Techniken beruht (z. B. inadäquate statistische Kombination).

Zahlreiche Faktoren können die *interne Validität* einer Studie beeinträchtigen.⁴⁰ Bei den Täterprogrammen sind die häufigsten Probleme das Selektionsbias⁴¹, das die Vergleichbarkeit der Experimental- und der Kontrollgruppe erschwert; das Instrumentenbias, das die Art betrifft, wie die Veränderungen gemessen werden; und das Attritionsbias, das heisst die Tatsache, dass die Teilnehmer die Studie abbrechen oder die

³⁹ Nach Farrington (2003), der jedoch die ersten vier Kriterien Shadish, Cook & Campbell (2002) zuschreibt.

⁴⁰ Klassischerweise werden sieben Gefährdungen unterschieden: Selektion, History, Reifung, Instrumentierung, Testing, Regression auf den Durchschnitt und Attrition.

⁴¹ Ein Bias ist eine systematische Abweichung, die zu fehlerhaften Ergebnissen einer Untersuchung führt. Das Selektionsbias beispielsweise ist ein Fehler in der Auswahl (Aufbau) der zu untersuchenden Stichprobe oder Stichproben.

Abbruchquote zwischen den verschiedenen Gruppen variiert.

Die Probleme der *Konstruktvalidität* sind nicht einfach zu ermitteln und machen zahlreiche Indikatoren erforderlich. Bei den Täterprogrammen betreffen diese Probleme häufig die Wahl und die Qualität der für die Messung der Gewaltabnahme verwendeten Indikatoren (z. B. amtliche Daten), aber auch die Schwierigkeiten abzuschätzen, ob das Programm seine Wirkung auch wirklich mit den angestrebten Mechanismen entfaltet (z. B. Konfliktlösungskompetenzen). Die Konstruktvalidität ist eine Voraussetzung, um die Mechanismen, die der Wirksamkeit eines Programms zugrunde liegen, identifizieren zu können.

Die Faktoren der *deskriptiven Validität* schliesslich (z. B. Beschreibung der Intervention, der statistischen Kontrollen) setzen die Ergebnisse in den Kontext, erleichtern die Wiederholung der Studie und verbessern die Validität der Analysen. Sie haben einen beachtlichen Einfluss auf das Vertrauen in die Ergebnisse einer Studie. Ohne zu verallgemeinern, ist angesichts der Vielfalt von Interventionen darauf hinzuweisen, dass die Evaluationen der Täterprogramme häufig nicht genügend Informationen zur Art des Programms, den Indikatoren oder den verschiedenen Programmphasen aufweisen. Dies trifft insbesondere auf die unveröffentlichten Studien (graue Literatur) oder auf Evaluationen zu, die seitens Programmkonzeptionierung und/oder Programmleitung durchgeführt wurden.⁴²

2.3 Die Synthesestudien und die Evaluationsstudien

Die zu entwickelnden Praktiken und Strategien auf der Grundlage einer einzigen Evaluationsstudie festzulegen, so exakt diese auch erarbeitet worden sein mag, ist kein besonders empfehlenswertes Vorgehen. Üblicherweise werden daher Zusammen-

fassungen der Evaluationsstudien erstellt, die sich hinsichtlich ihrer methodischen Qualität ebenfalls unterscheiden. Wir unterscheiden vier Arten von Übersichtsarbeiten zur Evidenz.⁴³

1. Die **narrative Methode** besteht darin, die bestehenden Evaluationsstudien zu beschreiben, ohne die Identifikation der Quellen zu systematisieren. Dieser Ansatz leidet jedoch an einem Selektionsbias, da gewisse Studien vernachlässigt werden können und die methodische Qualität der verwendeten Quellen nicht unbedingt berücksichtigt wird.
2. Die etwas differenziertere **Vote-Counting-Methode** quantifiziert die bestehenden Evaluationsstudien, wobei diejenigen Studien bewertet werden, die einem Programm Wirksamkeit attestieren, sowie diejenigen, die zum Schluss kommen, dass das Programm wirkungslos ist. Diese Methode gewinnt an Qualität, wenn bei der Zuteilung der Punkte der methodischen Qualität der verwendeten Quellen Rechnung getragen wird.
3. Die **systematische Übersichtsarbeit** ist eine rigorose, umfassende und explizite Zusammenfassung der Literatur zu einer bestimmten Problematik. Sie basiert vor allem auf der Ausarbeitung einer expliziten Untersuchungsfrage, anhand derer transparente und objektive Kriterien festgelegt werden, mit denen sich Selektionsbiases einschränken lassen. Die Kriterien können auf bereits bestehenden Klassifikationssystemen basieren oder spezifisch im Rahmen der systematischen Übersichtsarbeit entwickelt werden und beinhalten immer methodische Mindeststandards. Beispielsweise kann bei einer systematischen Übersichtsarbeit

⁴² Farrington (2003).

⁴³ Angepasst nach Welsh (2007), der eine fünfte Art, die «single study review method» vorsieht, wenn eine bekannte Studie zu Generalisierungen hinsichtlich der Wirksamkeit eines Programms führt. Diese ist aber für das Ziel dieses Berichts nicht wirklich relevant. Für eine eingehendere Betrachtung der systematischen Übersichtsarbeiten siehe z. B. Sherman, Farrington, Welsh & Mackenzie (2006).

beschlossen werden, die quasi-experimentellen Studien einzuschliessen, während diese bei einer anderen ausgeschlossen werden.⁴⁴

4. Die **Metaanalyse** ist ein statistischer Ansatz, mit dem die Daten, die aus einer systematischen Übersichtsarbeit hervorgegangen sind, kombiniert werden können. Sie beruht auf einer quantitativen Zusammenfassung dieser Studien, die genügend Ähnlichkeiten aufweisen müssen, damit sie in Form eines gewichteten Wirkungswerts – das bedeutet, dass die Effektgrösse jeder Studie nach ihrer Stichprobe gewichtet wird – gemeinsam analysiert werden können. Die Metaanalyse ermöglicht die Untersuchung einer grossen Zahl von Studien, sofern diese genügend gemeinsame Merkmale aufweisen, um zusammengefasst werden zu können. Der Einsatz von Metaanalysen ist beschränkt, zum einen aufgrund des Mangels an Evaluationsstudien und zum anderen aufgrund der Heterogenität der Programme und ihrer Evaluationsprozesse.

Für diese Bestandesaufnahme haben wir uns in erster Linie mit den systematischen Übersichtsarbeiten und den Metaanalysen befasst. Zudem integrieren wir die Ergebnisse bestimmter experimenteller Qualitätsstudien, die interessante Perspektiven für die Praxis und Forschung aufzeigen, auch wenn es noch verfrüht ist, ihre Feststellungen zu verallgemeinern. Schliesslich haben wir aufgrund des erhaltenen Auftrags auch europäische und schweizerische Quellen eingeschlossen, weisen aber darauf hin, dass deren methodischer Ansatz oft unzureichend ist. Ihre Ergebnisse sind daher mit Vorsicht zu interpretieren.

⁴⁴ Im Bereich der Sozialwissenschaften hat der Ausschluss der quasi-experimentellen Studien oft zur Folge, dass die Zahl der wählbaren Studien auf ein inakzeptables Niveau reduziert wird.

3 Die Evaluationsforschung zu den Täterprogrammen

Nun... die Studien zu den Interventionsprogrammen für Täter häuslicher Gewalt zeigen: dass die Interventionsprogramme für Täter häuslicher Gewalt wirksam sind, dass die Interventionsprogramme für Täter häuslicher Gewalt nicht wirksam sind, und dass sich unmöglich sagen lässt, ob die Interventionsprogramme für Täter häuslicher Gewalt wirksam sind oder nicht!⁴⁵

Der Stand der Evaluationsforschung im Bereich der Täterprogramme kann kaum klarer zusammengefasst werden als mit den Worten von Robert Moyer.

Die oben skizzierten methodischen Kriterien lassen darauf schliessen, dass die methodische Qualität der Evaluationsstudien zu den Täterprogrammen tief ist, sowohl in Europa als auch anderswo.⁴⁶ Auch wenn Vertreterinnen und Vertreter aus Forschung und Politik den Einsatz solcher Programme als Massnahme zur Bekämpfung häuslicher Gewalt unterstützen⁴⁷, stärkt die Tatsache, dass die Wirkung der Programme mit zunehmender methodischer Qualität der Evaluationsstudien tendenziell abnimmt oder gar verschwindet, das Vertrauen in ihre Ergebnisse nicht.⁴⁸

Einige, wie beispielsweise Stuart, Temple & Moore (2007: 561), haben sogar «die

Überstürztheit der Justizbehörden bei der Umsetzung von Interventionsprogrammen» angeprangert. Diese Überstürztheit habe zur Folge, dass die Programme weitgehend verwendet würden, bevor ihre Wirksamkeit in einer Evaluation untersucht wurde. Andere, wie Dutton & Corvo (2007), haben gar von Praktiken gesprochen, die auf ideologischen Überlegungen und nicht auf wissenschaftlichen Ergebnissen beruhen.⁴⁹

Im Folgenden werden die wichtigsten Erkenntnisse zur Wirksamkeit der Täterprogramme zusammengefasst, wobei gleichzeitig auf die Lücken im Bereich der Evaluation und der Praxis hingewiesen wird.

3.1 Studiendesigns von unterschiedlicher methodischer Qualität

3.1a Randomisierte experimentelle Studien

Die in der Evaluationsforschung als Referenz geltenden randomisierten experimentellen Studien, die ihren Ursprung in der biomedizinischen Wissenschaft haben, werden im Bereich der Bekämpfung und Prävention von Kriminalität nach wie vor nur wenig verwendet. Spezifisch bei den Täterprogrammen sind die randomisierten Studien (individueller oder Cluster-Ansatz) in der Unterzahl; bei Eckhardt *et al.* (2013a, 2013b) sind es lediglich 40 Prozent. Im Bereich der Kriminologie und der Strafjustiz ist ihre Zahl zwar angestiegen, aber die Durchführung solcher Studien geht mit einer Reihe von ethischen, rechtlichen und praktischen Einschränkungen einher, die ihre Beliebtheit schmälern.⁵⁰

In der Zunahme der randomisierten Studien wird von einigen zwar die Antwort auf die Unsicherheiten gesehen, mit der die Schlussfolgerungen der Evaluationsstudien zu Täterprogrammen behaftet sind. Dennoch bleibt fraglich, inwieweit es möglich oder realistisch ist, solche Studiendesigns im

⁴⁵ Nach Moyer (2004: 4): «So...research on batterer intervention programs shows : That batterer intervention programs are effective, that batterer intervention programs are not effective, and that it is impossible to say whether batterer intervention programs are effective or not!»

⁴⁶ Vor rund fünfzehn Jahren gelangte man in der Forschung zur Rehabilitation von Straftätern zur gleichen Feststellung. Gleich wie heute bei den Täterprogrammen war auch damals das Problem nicht so sehr der Mangel an Evaluationsstudien, sondern ihre geringe methodische Aussagekraft (Feder, Wilson, 2005).

⁴⁷ Dobash, Dobash, Cavanagh & Lewis (2000), Gondolf (2012) und Mackay *et al.* (2015).

⁴⁸ Diese Feststellung stammt aus der Metaanalyse von Babcock *et al.* (2004), wurde seither aber in vielen Arbeiten verwendet, insb. bei Labriola, Rempel & Davis (2005), Lehmann & Simmons (2009) sowie in mehreren Arbeiten unter der Leitung von Eckhardt (2006; 2013a; 2013b) und bis zu einem gewissen Grad auch bei Smedslund, Dalsbø, Steiro, Winsvold & Clench-Aas (2011).

⁴⁹ Für eine Antwort siehe Gondolf (2007).

⁵⁰ Zu diesem Punkt siehe Farrington & Welsh (2005).

Kontext der von der Justiz angeordneten Programmteilnahme umzusetzen. Auch wenn eine Randomisierung vorgenommen wird, so unterliegt diese nach wie vor den Regeln der Strafjustiz. Da die Richterinnen und Richter immer die Möglichkeit haben, die zufällige Zuteilung eines Täters zur einen oder zur anderen Gruppe des Studiendesigns zu ändern⁵¹, ist die Randomisierung nie absolut.⁵² Zudem ist auch bei diesem Designtyp die Verbleibquote der Studienteilnehmer tief. Schliesslich scheint es im Bereich der Täterprogramme, im Gegensatz zu den psychotherapeutischen Interventionen im Allgemeinen, nur selten eine Zusammenarbeit zwischen Praxis und Forschung zu geben, was die Erstellung von qualitativ hochwertigen Evaluationsstudien einschränkt und folglich auch den Kenntnissgewinn über die Arbeit mit Tätern von häuslicher Gewalt behindert.⁵³

Aufgrund dieser Feststellungen stehen viele Forscherinnen und Forscher einer (übermässigen) Fokussierung auf die randomisierten Studien kritisch gegenüber. Da jedes Studiendesign im Kontext des zu evaluierenden Programms erarbeitet werden muss, zweifeln sie an der Relevanz der randomisierten Studien, in denen die Täter künstlich einem Programm zugeteilt werden. Im «echten Leben» würde die Justiz eine andere Option auf der Grundlage spezifischer Kriterien vorziehen.⁵⁴ Mit anderen Worten, die reelle Welt der Justiz funktioniert nicht gleich wie ein Experiment. Warum also versuchen, sie auf diese Art zu verstehen?⁵⁵

3.1b Quasi-experimentelle Studien

Bei den quasi-experimentellen Studien wird eine Kontrollgruppe eingesetzt, deren Ergebnisse mit jenen der Experimentalgruppe verglichen werden können, das heisst mit der Gruppe, bei der das zu evaluierende Programm angewendet wird. Die relative Validität von verschiedenen quasi-experimentellen Studien hängt hauptsächlich davon ab, wie sich diese Kontrollgruppe zusammensetzt. Bei den Täterprogrammen stützen sich die Evaluationsstudien vor allem auf vier Typen von Kontrollgruppen:

- eine Kontrollgruppe ohne Behandlung oder (strafrechtliche) Sanktion;
- eine aktive Kontrollgruppe, die an einer anderen Intervention als die Experimentalgruppe teilnimmt (z. B. ein anderes Programm, Einzel- oder Paartherapie);
- eine Kontrollgruppe, bei der kein «Programm» durchgeführt wird (z. B. inhaftierte Täter); und
- eine Kontrollgruppe, die aus Personen besteht, die nicht für das Programm ausgewählt wurden, die sich nicht für das Programm gemeldet haben und/oder die das Programm vorzeitig abgebrochen haben.

Bei der Arbeit mit Tätern häuslicher Gewalt wurden in Evaluationsstudien am häufigsten die Programmteilnehmer mit Personen verglichen, bei denen keine Intervention stattfand. Studien, die zwei Experimentalgruppen verglichen, das heisst zwei Programme oder ein Programm und eine andere Intervention, sind viel seltener.

Die Kritik an der Methodik von quasi-experimentellen Studien betrifft hauptsächlich jene Studien, bei denen die Kontrollgruppe aus Personen bestand, die nicht am Programm teilgenommen oder das Programm vorzeitig abgebrochen haben. Der Forschungsstand legt zwar nahe, dass die Wahrscheinlichkeit eines Rückfalls bei den Teilnehmern, die ein

⁵¹ Für ein Beispiel siehe Feder & Dugan (2002).

⁵² Gemäss Eckhardt *et al.* (2013a, 2013b) sollen diese Probleme bei Studien zu alternativen Interventionen, die sich häufiger an der biomedizinischen Forschung orientieren, seltener sein, aber in diesem Fall ist überhaupt nichts bekannt darüber, wie diese Programme in der Gemeinschaft und der Strafjustiz breiter eingeführt werden könnten.

⁵³ Zu diesem Punkt siehe insb. Eckhardt *et al.* (2006) und Stuart *et al.* (2007).

⁵⁴ Für diese Denkweise siehe insb. Dobash *et al.* (2000) oder Gondolf (2001).

⁵⁵ Van Voorhis, Cullen & Applegate (1995).

Programm beenden, kleiner ist als bei denjenigen, die es abbrechen. Aber die Art des Studiendesigns erlaubt es nicht, allfällige bereits vorgängig bestehende Unterschiede zwischen den beiden Teilnehmergruppen zu ermitteln. Es ist möglich, dass die Teilnehmer, die das Programm beenden, einfach diejenigen sind, bei denen die Rückfallwahrscheinlichkeit von Anfang an am kleinsten war. Oder dass sie besondere Merkmale aufweisen, die sowohl ihre Motivation, das Programm zu beenden, als auch ihre geringere Rückfälligkeit erklären.⁵⁶

Mit anderen Worten, wenn festgestellt werden soll, ob die Teilnehmer, die das Programm beendet haben, weniger häufig rückfällig werden als diejenigen, die es abgebrochen haben, beantwortet dieser Designtyp nicht die richtige Frage, denn dieser Ansatz klärt die Wirkung eines Programms und geht über die Frage zur Wirkung der gerichtlichen Anordnung hinaus. Dies ist eine andere Frage, wenn *in Erfahrung gebracht werden soll, ob die angeordnete Programmteilnahme die Rückfallwahrscheinlichkeit im Allgemeinen verringert*. Diese zweite Frage, für die sich die Politik interessiert, ist umfassender: Sie ist auf den Nutzen der angeordneten Programmteilnahme für die Gesellschaft, und weniger für die betroffene Einzelperson ausgerichtet.⁵⁷

Unabhängig von der Zusammensetzung der Kontrollgruppe scheint es angezeigt, Selektionsbiases zu minimieren, indem statistische Variablen aufgrund ihres wahrscheinlichen Einflusses auf das Evaluationsergebnis festgelegt werden.⁵⁸ Dies dürfte sich im Bereich der Strafjustiz jedoch als besonders schwierig erweisen, da es sich hier um einen Bereich handelt, in dem die Behandlung (hier die Sanktion) sowohl von der Situation als auch von den Merkmalen des

Straftäters (z. B. Motivation, Risiko) bestimmt wird.⁵⁹

3.1c Systematische Übersichtsarbeiten und Metaanalysen

Aufgrund der Heterogenität der Täterprogramme und der inhärenten methodischen Probleme der Studien, die versucht haben, die Wirksamkeit der Programme zu evaluieren, ist die Zahl der systematischen Übersichtsarbeiten und Metaanalysen nach wie vor gering, und die ihnen zugrunde liegenden Kriterien verschieden.

Bei den Metaanalysen variiert die Zahl der berücksichtigten Studien zwischen 7 und 19, genauer 2 experimentelle Studien und 5 quasi-experimentelle Studien bei Davis & Taylor (1999), 5 experimentelle Studien und 17 quasi-experimentelle Studien bei Babcock *et al.* (2004), 4 experimentelle Studien und 6 quasi-experimentelle Studien bei Feder & Wilson (2005), 9 (quasi-)experimentelle Studien bei Miller, Drake & Nafziger (2013), und 6 experimentelle Studien und 13 quasi-experimentelle Studien bei Arias, Arce & Vilariño (2013).

Die systematischen Übersichtsarbeiten hingegen umfassen eine grössere Zahl von Studien, da ihre Kriterien weniger strikt sind; es handelt sich um zwischen 6 und 65 Studien für die in diesem Bericht berücksichtigten Evaluationsstudien. Unter Festlegung von unspezifischen Zielen und unter Anwendung von flexiblen Kriterien wurden in der im Rahmen des europäischen Projekts IMPACT⁶⁰ durchgeführten Übersichtsarbeit 65 und in jener des englischen «National Institute of Clinical Excellence»⁶¹ 33 Studien untersucht. Unter Anwendung deutlich strikterer Kriterien haben Eckhardt *et al.* (2013a, 2013b) 20 Studien zu Programmen untersucht, die auf

⁵⁶ Eckhardt *et al.* (2013a, 2013b) weisen auch darauf hin, dass nicht genau festgestellt werden kann, was die Teilnehmer nach dem Programmabbruch getan haben und inwiefern dies einen Einfluss auf ihre Rückfälligkeit hat.

⁵⁷ Eckhardt *et al.* (2013a, 2013b).

⁵⁸ Kontrollen wie die Methode des Propensity Score-Matchings; für konkrete Beispiele zur Evaluation von Täterprogrammen siehe Jones, D'Agostino, Gondolf & Heckert (2004) oder Labriola *et al.* (2005).

⁵⁹ Farrington & Welsh (2005).

⁶⁰ Hester *et al.* (2014).

⁶¹ NICE (2014) und die Arbeitsdokumente, www.nice.org.uk/guidance/ph50 et www.nice.org.uk/advice/lgb20.

feministisch orientierten psychoedukativen Ansätzen oder kognitiver Verhaltenstherapie beruhen. Ferner haben sie 10 Studien zu alternativen Ansätzen verglichen. Akoensi, Koehler, Lösel & Humphreys (2013) haben 12 Evaluationen von in Europa umgesetzten Programmen untersucht. Smedslund *et al.* (2011) haben sehr strikte Kriterien angewandt und nur 6 Studien in ihre Übersichtsarbeit einschlossen, die ausschliesslich Programme mit einem kognitiv-verhaltenstherapeutischen Ansatz betrafen.

Bis auf wenige Ausnahmen⁶² stammen die in die Metaanalysen und die systematischen Übersichtsarbeiten eingeschlossenen Studien aus Nordamerika, die mehrheitlich von der Justiz angeordnete Programme untersuchten.⁶³ So betrafen 71–91 Prozent der von Eckhardt *et al.* (2013a, 2013b) untersuchten Studien solche Programme.

Schliesslich betreffen sowohl die systematischen Übersichtsarbeiten als auch die Metaanalysen hauptsächlich Programme, die sich am Duluth-Modell (feministisch orientierter psychoedukativer Ansatz) und am kognitiv-verhaltenstherapeutischen Ansatz orientieren; bei den von Eckhardt *et al.* (2013a, 2013b) untersuchten Studien war das bei zwei Dritteln beziehungsweise einem Drittel der Fall.⁶⁴ Die alternativen Modelle (z. B. Motivierende Gesprächsführung, Paartherapie) wurden weniger häufig evaluiert.⁶⁵

3.2 Die wichtigsten methodischen Probleme bei Evaluationen von Täterprogrammen

Unabhängig von ihrem Studiendesign sind die Evaluationsstudien mit wiederkehrenden methodischen Problemen konfrontiert, die beachtet werden müssen. Die Hauptprobleme bei der Evaluation von Täterprogrammen sind: (1) die Attrition oder die differentielle Attrition; (2) die Operationalisierung der Wiederholung häuslicher Gewalt; und (3) die Beobachtungsdauer und -intervalle sowie die externe Validität der Studien.

3.2a Eine hohe Zahl vorzeitiger Programmabbrüche

Die Attrition bezieht sich hier auf den Programmabbruch durch einen Teilnehmer, entweder (a) zwischen dem ersten Kontakt und der ersten Programmsitzung oder (b) nach Beginn des Programms. Bei nahezu allen Evaluationsstudien lässt sich eine hohe Attrition feststellen: Offenbar beenden die Hälfte der Teilnehmer das Programm nicht.⁶⁶ Die Forschung hat keine besonderen Techniken oder Strategien identifiziert, mit denen sich dieses Problem mindern liesse, auch wenn oft davon ausgegangen wird, dass das Ausbleiben von Sanktionen vonseiten der Justiz Programmabbrüche fördert. Gleichzeitig ist die Tatsache, dass eine Person im Programm bleibt, nicht unbedingt eine Garantie dafür, dass sie die Verantwortung für ihr Handeln übernimmt oder sich verändert hat.⁶⁷

Die Attrition führt zudem zu methodischen Schwierigkeiten, da es zu entscheiden gilt, ob die weitere Entwicklung der Teilnehmer, die das Programm abgebrochen haben, berücksichtigt werden soll oder nicht. Beide

⁶² Akoensi *et al.* (2013) und Hester *et al.* (2014).

⁶³ Die Studien von besonderer Bedeutung werden im Anhang zusammengefasst, und alle beigezogenen Studien sind in der Bibliografie aufgeführt.

⁶⁴ Siehe oben, Titel 1.2.

⁶⁵ Diese alternativen Ansätze sind bei Eckhardt *et al.* (2013a, 2013b) Teil einer spezifischen Untersuchung, es ist jedoch schwierig, aus den Analysen allgemeine Schlüsse zu ziehen, da unter dieser Bezeichnung sehr unterschiedliche Interventionen zusammengefasst werden; siehe weiter unten, S. 30 ff.

⁶⁶ Gemäss Edleson (2012). Nach Auffassung von Mullender & Burton (2001) beträgt die Abbruchquote bei den englischen Programmen zwischen 30 und 90 Prozent und ist bei den freiwilligen Teilnehmern höher. Die schweizerischen Daten weisen ebenfalls darauf hin, dass die Hälfte der Teilnehmer das Programm nicht beenden; siehe z. B. Gloor & Meier (2002, 2003) oder Jaquier, Stein & Van Daele (2016).

⁶⁷ Day, O'Leary, Chung & Justo (2009b) und Gondolf (2004).

Ansätze führen zu Verzerrungen. Eine Evaluation nach dem Intent-to-treat-Prinzip führt zu sehr konservativen Schätzungen, da die Behandlungsgruppe auch Personen umfasst, die am Programm nur teilweise oder überhaupt nicht teilgenommen haben. In einer Evaluation nach dem Per-Protocol-Prinzip («dosage») werden hingegen die Personen, die das Programm nicht beendet haben, ausgeschlossen, was die signifikanten Wirkungen der Behandlung begünstigt, da nur die «guten» Teilnehmer berücksichtigt werden.⁶⁸

3.2b Die Wiederholung häuslicher Gewalt als wichtigster Indikator der Veränderung

Angesichts des ursprünglichen Ziels der Täterprogramme erstaunt es nicht, dass ihre Evaluationen fast ausschliesslich Indikatoren der Veränderung auf individueller Ebene, hauptsächlich die Wiederholung häuslicher Gewalt, messen.⁶⁹ Dazu stützen sich die Studien auf drei Indikatoren: die *amtlichen Daten* (polizeilich registrierte Straftaten, Verurteilungen), *selbstberichtete Daten*, das heisst Informationen der Teilnehmer, und *Daten zur Viktimisierung*, das heisst Informationen der Partnerinnen. Mehrheitlich werden die verschiedenen Indikatoren in den Studien kombiniert.⁷⁰

Die Betrachtung von ausschliesslich amtlichen Daten geht mit einer offensichtlichen Verzerrung einher, da den Behörden nur eine kleine Zahl der Fälle häuslicher Gewalt gemeldet wird.⁷¹ Zudem basieren diese Daten

auch auf dem Willen des Opfers, neue Gewalttaten anzuzeigen. Diese Motivation kann von einer Vielzahl von Faktoren abhängen, namentlich vom Wunsch, dass der Partner weiterhin am Programm teilnimmt, weil die Partnerin an dessen Nutzen glaubt, oder vom Misstrauen gegenüber den Behörden, weil die sie über die staatliche Intervention enttäuscht oder verärgert ist.⁷² Zudem werden mit den amtlichen Daten nur die offiziellen Straftaten erfasst. Die selbstberichteten Daten oder jene zur Viktimisierung ergeben daher ein umfassenderes Bild des Verhaltens des Täters.

Die Integration der von den Opfern angegebenen Daten in ein Evaluationsvorgehen ist jedoch mit gewissen Problemen verbunden, und das Risiko einer Unterschätzung der Wiederholungstaten bleibt bestehen. Nach Feder *et al.* (2008) ist der Anteil der Opfer, die während der Beobachtungsperiode an der Studie teilnehmen, mit 30–80 Prozent der Opfer tief. Es kann jedoch nicht ausgeschlossen werden, dass sich die Opfer, die nicht mehr erreichbar sind, von denjenigen, die antworten, unterscheiden, wodurch eine Verzerrung entsteht. Dies nicht zuletzt, weil die Forschung nahelegt, dass bei gewissen Opfern die Wahrscheinlichkeit in einer Evaluationsstudie unberücksichtigt zu bleiben grösser ist als bei anderen, vor allem bei denjenigen, die randständig(er) und Opfer von schwereren und häufigeren Gewalttaten sind.⁷³

Umgekehrt wurde auch gefolgert, dass die Opfer dank der Unterstützung, die sie im Rahmen gewisser Täterprogramme erhalten, oder dem neu gefundenen Vertrauen in die

⁶⁸ Für eine Diskussion siehe ins. Gondolf (2001).

⁶⁹ Die meisten Evaluationen untersuchen nur die Wiederholungstaten im Bereich der häuslichen Gewalt und nicht alle neuen Straftaten, die möglicherweise begangen wurden (Eckhardt *et al.*, 2013a, 2013b).

⁷⁰ Im Projekt IMPACT beispielsweise beinhalteten von insgesamt 65 berücksichtigten Evaluationen 15 amtliche Daten zur Wiederholung häuslicher Gewalt, 11 selbstberichtete Daten, und nur 4 Daten zur Viktimisierung (Hester *et al.*, 2014: 26–33). Von den 20 von Eckhardt *et al.* (2013a, 2013b) untersuchten Evaluationen beinhalteten 11 selbstberichtete oder von den Opfern übermittelte Daten, während nur 9 amtliche Daten berücksichtigten.

⁷¹ Je nach Studie zwischen einem Viertel und einem Drittel; aus Nordamerika: (Johnson, Dawson, 2010: 66, 94; Tjaden, Thoennes, 2000: 49–50), aus Europa: (Jaspard, Brown, Condon, Firdion, Fougeyrollas-Schwebel, Houel, Lhomond,

Maillochon, Saurel-Cubizolles, Schiltz, 2002: 275–8) und aus der Schweiz: (Gillioz, De Puy, Ducret, 1997: 101; Killias, Simonin, De Puy, 2005: 83–6).

⁷² Insb. Eckhardt *et al.* (2013a, 2013b) und Feder, Wilson und Austin (2008).

⁷³ Siehe Sullivan, Rumpitz, Campbell, Eby & Davidson (1996) zur Art, die Attrition in den Studien zu reduzieren, in denen die Frauen als Opfer häuslicher Gewalt mit berücksichtigt sind. Andere Studien kommen zum Schluss, dass der Konformismus der Teilnehmer die Wahrscheinlichkeit erhöht, dass sie das Programm abschliessen und nicht rückfällig werden (Feder *et al.*, 2008); siehe auch S. 29 ff.

Behörden eher in der Lage sind, die erlittenen Gewalttaten anzuzeigen. Ein Zusammenhang mit der Wirkung der Intervention bei den Männern besteht dabei nicht.⁷⁴

Schliesslich wird neben den Verzerrungen, die mit den Indikatoren der Wiederholung von Gewalttaten verbunden sind, auch die Tatsache kritisiert, dass die Gewaltwiederholung überhaupt als Massstab für die Wirksamkeit der Täterprogramme genommen wird. Ohne ins Detail zu gehen, da diese Frage anschliessend spezifisch erörtert wird,⁷⁵ sei hier darauf hingewiesen, dass in jüngster Zeit die Tendenz besteht, andere «Erfolgs»-Indikatoren zu berücksichtigen, wie Verhaltens- oder Einstellungsänderungen und psychologische Veränderungen beim Täter sowie quantitative und qualitative Informationen, die bei der Partnerin (oder Ex-Partnerin) erhoben werden.⁷⁶

3.2c Eine zu kurze Beobachtungsdauer und unregelmässige Bewertungsintervalle

Es ist nicht überraschend, dass sich die Studiendesigns hinsichtlich der Beobachtungsdauer und der Bewertungsintervalle unterscheiden. Durch diese Unterschiede werden die Validität der Vergleiche zwischen den Studien und die Möglichkeit der Durchführung von Metaanalysen eingeschränkt. Eckhardt *et al.* (2013a, 2013b) haben beispielsweise festgestellt, dass die Beobachtungsdauer nach der Intervention zwischen 0 und 54 Monate umfassen konnte, was einer Mediandauer von 18 Monaten entspricht.⁷⁷ Sowohl die Beobachtungsdauer

als auch die Bewertungsintervalle sind abhängig von den Indikatoren der Veränderungen, die gemessen werden sollen.⁷⁸ Die Studien, welche die Daten ihrer Indikatoren unmittelbar nach Beendigung des Programms messen, dürften daher im Bereich der Verminderung von Rückfällen positivere Ergebnisse aufweisen als diejenigen, die diese Ergebnisse nach einem längeren Zeitintervall messen.⁷⁹

Unabhängig von der Beobachtungsdauer und dem Bewertungsintervall bleiben die Ergebnisse abhängig von der Quote des Verbleibs der Teilnehmer in der Studie, das heisst von der Möglichkeit, mit den Tätern oder ihren Partnerinnen für die Messung der ausgewählten Indikatoren wieder Kontakt aufzunehmen. Bei den von Eckhardt *et al.* (2013a, 2013b) untersuchten Evaluationen der feministisch orientierten psychoedukativen und kognitiv-verhaltenstherapeutischen Programme beispielsweise konnten nach dem letzten Kontakt mit dem Teilnehmer in 61 Prozent der Fälle Daten für die selbstberichteten Indikatoren erhoben werden (Medianwert; Intervall: 25–100 Prozent), während bei Feder & Dugan (2002) nur 30 Prozent der Partnerinnen an der Evaluation teilnahmen, die 6 Monate nach Beendigung des Programms durchgeführt wurde, nach 12 Monaten waren es nur noch 22 Prozent.

3.2d Unbekannte externe Validität

Schliesslich können auch die Studien mit den striktesten Kriterien die Verzerrungen der externen Validität nicht ausgleichen, wenn sie nicht mit einer allgemeinen Tätergruppe durchgeführt werden, was selten der Fall ist.

Bei den meisten Evaluationsstudien erfolgt eine Selektion der Teilnehmer, beispielsweise indem freiwillige Teilnehmer akzeptiert werden oder nicht oder indem die Teilnehmer aufgrund

⁷⁴ Klein & Crowe (2008).

⁷⁵ Auf diese Frage wird weiter unten eingegangen, S. 36 ff.

⁷⁶ Insb. Day *et al.* (2009b: 233), Kelly & Westmarland (2015b) und Westmarland & Kelly (2013). Für einen Überblick über die verschiedenen Indikatoren, die im Rahmen der Täterarbeit erhoben werden können, siehe die verschiedenen Dokumente und Leitlinien, die im Rahmen der europäischen Projekte WWP und IMPACT erstellt wurden (Fussnote 38); siehe auch unten S. 37 ff.

⁷⁷ Die Evaluationen von traditionellen Programmen umfassen in der Regel sechs Monate nach der Intervention eine Bewertung, während die alternativen Programme, namentlich die psychodynamischen Interventionsprogramme, für die betrachteten Veränderungsindikatoren kürzere Intervalle beinhalten können (z. B. ein Monat nach der Intervention).

⁷⁸ Als Schlussfolgerung ihrer Metaanalyse empfehlen Davis & Taylor (1999) beispielsweise eine Beobachtung mit sechsmonatigen Intervallen über einen Zeitraum von mindestens einem, idealerweise zwei Jahren.

⁷⁹ Dunford (2000) und Feder *et al.* (2008).

ihrer kriminellen Vorgeschichte oder ihrer Gefährlichkeit ausgewählt werden. Viele dieser Kriterien sind aber nicht unbedingt explizit erwähnt. Dazu kommt, dass auch die Unterschiede zwischen den Teilnehmern, beispielsweise hinsichtlich ihrer soziodemografischen, psychologischen und motivationalen Merkmale, einen Einfluss auf die Generalisierbarkeit der Beobachtungen haben können.⁸⁰

Schliesslich muss angesichts der Bedeutung der rechtlichen und soziokulturellen Rahmenbedingungen im Bereich der Bekämpfung von häuslicher Gewalt und im Hinblick auf die Wirksamkeit der Täterprogramme auch der Kontext berücksichtigt werden, in dem ein Programm umgesetzt wird – insbesondere die Verbindungen zu anderen Dienstleistungen, der Bezug zur Justiz und das Bestehen von Sanktionen bei Nichteinhaltung der Regeln – damit die Schlussfolgerungen verallgemeinert werden können.⁸¹

Es ist zentral, die Kriterien zu identifizieren, die die Verallgemeinerung der beobachteten Ergebnisse auf alle Täter von häuslicher Gewalt beeinflussen können.

⁸⁰ Einige wie Eckhardt *et al.* (2013a, 2013b) haben zwar die Hypothese aufgestellt, dass die besten Ergebnisse eines Programms mit bestimmten Tätergruppen auf deren ursprüngliche Motivation zurückzuführen sei, aber die Ergebnisse der Voruntersuchungen in diesem Bereich sind inkonsistent. Davis & Taylor (1999) haben hingegen bei zwei Gruppen mit vergleichbarer Motivation unterschiedliche Ergebnisse beobachtet. Zum jetzigen Zeitpunkt scheint eine Schlussfolgerung schwierig; siehe auch unten S. 30 ff.

⁸¹ Z. B. Day *et al.* (2009a) und Gondolf (2004); siehe auch Titel 1.1.

4 Der Erfolg der Täterprogramme

Da weder ein Behandlungsmodell noch eine Behandlungsmodalität eine den anderen überlegene Wirksamkeit zeigte, ist es für die Staaten noch zu früh, Standards zur Einschränkung des verfügbaren Behandlungsangebots für die Täter von häuslicher Gewalt zu erlassen. Die Interventionsdienste im Bereich der häuslichen Gewalt können ihre Dienstleistungen vielmehr dadurch verbessern, dass sie Elemente hinzufügen oder die Behandlungen auf eine spezifische Klientel zuschneiden, als dass sie stur an einem Programm festhalten, solange es keine empirische Evidenz für dessen überlegene Wirksamkeit gibt. (...) Die Programmanbieter sollten alternative Ansätze entwickeln und mit der Forschung zusammenarbeiten, um deren Wirksamkeit zu evaluieren und so die Entwicklung einer evidenzbasierten Praxis voranzubringen. Dazu muss die Forschung zu einem integralen Bestandteil der koordinierten Antwort der Gesellschaft auf häusliche Gewalt werden.⁸²

Ein Literaturüberblick zeigt, dass mehr über die Programme bekannt ist, die nicht funktionieren – oder nicht zu funktionieren scheinen – als über die Interventionen, die einen gewissen Erfolg erzielt haben. Zudem muss angesichts der Kritik an der methodischen Qualität der Evaluationsstudien jede Stellungnahme vorsichtig formuliert werden und auf einer genauen Betrachtung der methodischen Qualität der Studien, auf die sie sich bezieht, beruhen.

Die Evaluationsstudien zu den Täterprogrammen zeichnen sich in erster Linie durch Inkonsistenz aus. Auf eine erste Reihe von Studien, die den Programmen in etwas

undifferenzierter Art eine Abnahme der Wiederholung der häuslichen Gewalt attestierten, folgten Evaluationsforschungen, die wissenschaftlich fundierter waren, aber oft zu widersprüchlichen Ergebnissen gelangten.

Die Widersprüche sind zurückzuführen auf: (a) die unterschiedlichen Ansprüche an die methodische Qualität dieser Studien;⁸³ (b) die gewählten Erfolgskriterien und ihre Indikatoren; (c) die Beobachtungsdauer; und (d) die unterschiedlichen Grade der Programm-umsetzungen.

4.1 Das Kriterium der Wiederholung häuslicher Gewalt

4.1a Ist das Glas halb voll oder halb leer?

Die Auseinandersetzung mit der Frage der Wirksamkeit der Täterprogramme basiert auf einer kleinen Zahl methodisch strikter experimenteller und quasi-experimenteller Studien und einer grösseren Zahl nicht experimenteller Studien. Die grosse Mehrheit der zweiten Studien kommt zum Schluss, dass die Teilnahme an einem Programm das spätere Risiko von häuslicher Gewalt reduziert, während die ersten ergaben, dass die positive Wirkung der Täterprogramme gering und in gewissen Fällen gar gleich null ist.

Bestimmte Forschungs- und Praxiskreise rufen dazu auf, die Strafjustiz solle häufiger Täter in standardisierte Programme zuweisen, um die Opfer von häuslicher Gewalt zu schützen, auch wenn noch kein definitiver «Nachweis» für die Wirksamkeit dieser Programme vorliegt. Andere hingegen sind zurückhaltender und weisen darauf hin, dass es zurzeit keine solide empirische Evidenz gibt zur signifikanten Reduktion von Wiederholungstaten häuslicher Gewalt durch diese Programme noch zur

⁸² Nach Babcock *et al.* (2004: 1048–9): «Because no one treatment model or modality has demonstrated superiority over the others, it is premature for states to issue mandates limiting the range of treatment options for batterers. Battering intervention agencies are more likely to improve their services by adding components or tailoring their treatments to specific clientele, than by rigidly adhering to any one curriculum in the absence of empirical evidence of its superior efficacy. (...) Treatment providers should develop alternative techniques and collaborate with researchers to evaluate their efficacy in an effort to develop evidence-based practice. To this end, researchers need to become an integral part of the coordinated community response to domestic violence.»

⁸³ Dies veranlasst zur Hypothese, dass je strikter die Evaluationen, umso kleiner die Wahrscheinlichkeit, dass sie den Täterprogrammen eine Wirkung auf die Abnahme der Wiederholung häuslicher Gewalt attestieren; siehe Fussnote 48.

Bevorzugung des einen oder anderen Programms.

In dieser Diskussion Stellung zu beziehen bedeutet in einem gewissen Mass zu entscheiden, ob das Glas halb voll oder halb leer ist. Dies trifft wortwörtlich auf die systematische Übersichtsarbeit von Eckhardt *et al.* (2013a, 2013b) zu, gelangt doch die Hälfte der untersuchten Evaluationen zum Schluss, dass das Programm wirksam ist, während die andere Hälfte den gegenteiligen Schluss zieht.

Insgesamt stellen die Metaanalysen bei den Männern, die ein Programm abschliessen, eine leichte, aber signifikante Abnahme der Wiederholung häuslicher Gewalt fest im Vergleich zu den Männern, die das Programm abbrechen oder die an gar keinem teilgenommen haben.⁸⁴ Dieser Rückgang wird vor allem in den Studien beobachtet, die amtliche Daten als Indikatoren der Wiederholung verwenden;⁸⁵ es kann daher nicht festgestellt werden, ob sich diese Schlussfolgerungen auf alle Täter von häuslicher Gewalt übertragen lassen.⁸⁶ Schliessen die Evaluationen die von den Opfern übermittelten Viktimisierungsdaten ein, nehmen die beobachteten Unterschiede zwischen der Experimental- und der Kontrollgruppe tendenziell ab oder verschwinden ganz.⁸⁷

⁸⁴ Davis & Taylor (1999) kamen auf der Grundlage von 7 Studien zum Schluss, dass die Interventionsprogramme eine signifikante Wirkung auf den Rückgang der Wiederholung häuslicher Gewalt haben. Babcock *et al.* (2004) attestierten diesen Programmen gestützt auf 22 Studien eine sehr geringe Wirkung. In einer vorsichtigeren Untersuchung von Feder *et al.* (2008) mit lediglich 10 Studien wurde auf der Grundlage von amtlichen Polizeidaten eine geringe Wirkung der Programme und gestützt auf die Daten der Opfer gar keine Wirkung festgestellt. Andere Autorinnen und Autoren waren in der Wahl der Kriterien noch strenger, so Smedslund *et al.* (2011), die zum Schluss gelangten, dass sich mit den bestehenden Daten überhaupt keine Wirksamkeit der Programme auf die Abnahme der Wiederholung häuslicher Gewalt feststellen lasse. Feder & Wilson (2005) schliesslich stellten fest, dass die Interventionsprogramme wirkungslos seien oder gar negative Auswirkungen hätten.

⁸⁵ Babcock *et al.* (2004), Edleson (2012) und Feder & Wilson (2005).

⁸⁶ Feder *et al.* (2008).

⁸⁷ Edleson (2012) und Eckhardt *et al.* (2013a, 2013b).

Nur eine der 12 von Akoensi *et al.* (2013) ausgewählten europäischen Studien, die erst noch von schlechter Qualität war, beinhaltete eine Vergleichsgruppe. Das Projekt IMPACT weist zwar die Besonderheit auf, Studien aus verschiedenen rechtlichen und soziokulturellen Kontexten zu berücksichtigen, aber bei der Formulierung der Schlussfolgerungen wurde weder zwischen den Programmen noch zwischen den Studiendesigns unterschieden. Zudem wurde die Zusammenfassung auf der Grundlage von zahlreichen verhaltens- und einstellungsbezogenen sowie psychologischen und motivationalen Veränderungsindikatoren oder in Anlehnung an andere Risikofaktoren der Wiederholung häuslicher Gewalt erstellt, ohne dabei irgendeine Unterscheidung vorzunehmen. So gelangten 13 der 65 verglichenen Studien zum Schluss, dass die Interventionsprogramme eine statistisch signifikante Wirkung auf verschiedene Veränderungsindikatoren haben, während 37 Studien den Programmen eine geringe Wirkung attestieren, ohne jedoch auf statistisch signifikante Ergebnisse zu verweisen. Leserinnen und Leser, die sich spezifischer informieren wollen, werden auf die Zusammenfassung der einzelnen Studien verwiesen.⁸⁸

4.1b Worin besteht die Täterarbeit?

Während die Täterprogramme tatsächlich eine geringfügige Abnahme der Wiederholung häuslicher Gewalt bewirken, liefert die Forschung nur wenige Anhaltspunkte zu den Mechanismen, die diesen beobachteten Veränderungen zugrunde liegen.

Hinsichtlich der Modalitäten der Betreuung, der wirksamsten Aktivitäten oder gewisser besonderer Merkmale der Gruppenarbeit scheinen sich keine Regelmässigkeiten abzuzeichnen, dies weder in jüngeren⁸⁹ noch in älteren Studien.⁹⁰ Daher lässt sich nicht

⁸⁸ Hester *et al.* (2014).

⁸⁹ Babcock *et al.* (2004), Eckhardt *et al.* (2013a, 2013b), Edleson (2012) und Gondolf (2004).

⁹⁰ Z. B. ein Vergleich zwischen einem psychoedukativen Programm und einem Programm für Paare (Dunford, 2000); ein Vergleich zwischen einem psychoedukativen Programm

ausschliessen, dass die Kontrolle, die durch die regelmässige Teilnahme in einer Gruppe entsteht, für die beobachteten Veränderungen verantwortlich ist, ungeachtet der Modalitäten der verschiedenen Programme oder ihrer charakteristischen Aktivitäten.

Gesamthaft betrachtet scheint bei den Tätern von häuslicher Gewalt kein Interventionsmodell besonders wirksam zu sein. So beobachteten Eckhardt *et al.* (2013a, 2013b) zum einen eine ähnliche Wirksamkeit der traditionellen und der alternativen Interventionen auf die Abnahme von häuslicher Gewalt und zum anderen keinen Unterschied zwischen den Modellen mit einem profeministischen Ansatz und jenen mit einer kognitiv-verhaltenstherapeutischen Ausrichtung.⁹¹ Dazu ist jedoch zu erwähnen, dass das Aufzeigen der Auswirkungen eines besonderen theoretischen Modells oder von spezifischen Aktivitäten durch den Eklektizismus⁹² der Ansätze eingeschränkt, wenn nicht sogar verhindert wird.⁹³

Es ist theoretisch anerkannt, dass jedes Programm auf der Grundlage einer kohärenten und empirisch soliden Theorie erarbeitet werden muss, die sowohl die gewalttätigen Verhaltensmechanismen als auch die angestrebten Veränderungsmechanismen erklärt. Die Logik der verschiedenen Programme wird jedoch nur selten formuliert, was sich in einer oft unvollständigen Programmdurchführung äussert.⁹⁴ Zudem

existiert kein Konsens darüber, welches die Kriterien für die Vollständigkeit eines Programms sind.⁹⁵

Schliesslich scheint auch die Tatsache, dass gewisse Programme strukturierter sind als andere, keinen Einfluss auf ihre jeweilige Wirksamkeit zu haben. Zwar legen ältere Studien nahe, dass die Wirksamkeit von sehr strukturierten Programmen bei allen Tätern⁹⁶ oder nur bei bestimmten Täterkategorien⁹⁷ grösser ist, aber keine hat sich hierzu kategorisch geäussert.

Der Frage der Strukturierung scheint viel grössere Bedeutung zuzukommen, wenn sie in Bezug auf die Vollständigkeit der verschiedenen Interventionsprogramme, und nicht nur in Bezug auf ihren jeweiligen Inhalt, analysiert wird. So wurde die bereits in den 1990er-Jahren formulierte Erkenntnis, dass die koordinierten Gemeinschaftsansätze die Wirksamkeit der Täterprogramme erhöhen, im Laufe der Zeit untermauert, insbesondere durch den Ausdruck «system matters», der durch die Arbeiten von Gondolf populär geworden ist.⁹⁸ Die Wirksamkeit eines Programms beruhe dabei zum Teil auf der Fähigkeit des Systems, die Legitimität der Interventionen zu verankern, ihre Umsetzung zu begleiten, ihren Ablauf zu überwachen und jeglichen Verstoß gegen die Regeln zu sanktionieren.

4.1c Wie viel Zeit soll der Zeit eingeräumt werden?

Einige Studien haben die Gesamtdauer der Täterprogramme verglichen⁹⁹, ohne jedoch

und Selbsthilfegruppen (Edleson, Syers, 1990); oder ein Vergleich zwischen einem psychodynamischen Ansatz und einem feministisch orientierten kognitiv-verhaltenstherapeutischen Ansatz (Saunders, 1996).

⁹¹ Während die Wirksamkeit der kognitiv-verhaltenstherapeutischen Ansätze im Bereich der Rückfallprävention in der Regel anerkannt ist, scheint dies im Bereich der häuslichen Gewalt nicht besonders der Fall zu sein (Smedslund *et al.*, 2011). Dies deckt sich mit der Feststellung von Day *et al.* (2009a), die nahelegen, dass die Täterprogramme insgesamt weniger wirksam sind, um die Wiederholung von häuslicher Gewalt zu vermindern, als die Präventionsprogramme für andere Kategorien von Straftätern.

⁹² Zur Frage des Eklektizismus im Bereich der Arbeit mit Tätern von häuslicher Gewalt siehe die ausgezeichnete Zusammenfassung von Vlais (2014).

⁹³ Zur Verwechslung der Ansätze siehe Titel 1.2.

⁹⁴ Day *et al.* (2009a).

⁹⁵ Kelly & Westmarland (2015a: 36).

⁹⁶ Edleson & Syers (1990).

⁹⁷ Saunders (1996).

⁹⁸ Gondolf (2004).

⁹⁹ In der Multi-Site-Prüfung von Gondolf (1999) wurde 15 Monate nach Beendigung des Programms kein signifikanter Unterschied zwischen den Gruppen mit einer dreimonatigen beziehungsweise einer neunmonatigen Programmdauer festgestellt. Edleson & Syers (1990) haben 6 Monate nach Beendigung des Programms ebenfalls keinen Unterschied zwischen den Teilnehmern festgestellt, die entweder einer Intensivversion (32 Sitzungen, zweimal wöchentlich) oder einer Standardversion (12 Wochensitzungen) zugeteilt waren.

hinsichtlich der Wiederholung von häuslicher Gewalt signifikante Unterschiede festzustellen: Es scheint keine Regel für eine allfällige ideale Interventionsdauer zu geben.¹⁰⁰

Auch wenn die Zahl der erforderlichen Sitzungen und deren Dauer in den Evaluationen nicht immer angegeben sind, weisen die verfügbaren Daten auf eine hohe Variabilität hin. Eckhardt *et al.* (2013a, 2013b) stellten fest, dass die Zahl der Sitzungen zwischen 8 und 52 beträgt, das sind im Durchschnitt 22,2 Sitzungen, die jeweils zwischen 75 und 180 Minuten, also durchschnittlich 118 Minuten dauerten. Dabei ist jedoch zu beachten, dass die Wahl eines Programms mit einer besonderen Dauer auch aus anderen Gründen (z. B. symbolische Dimension, logistische Faktoren) als nur wegen der Auswirkungen auf die Verhaltensveränderung erfolgen kann. Was den Grundsatz angeht, von der Dauer auf die Ergebnisse zu schliessen, so scheint die Mehrheit der Programme (deutlich) kürzer zu sein als der Zeitraum, in dem von einer Rückfallgefahr ausgegangen wird,¹⁰¹ der bei der Mehrheit der Täter häuslicher Gewalt rund fünfzehn Monate beträgt.¹⁰²

Auf jeden Fall scheint es wichtig, dass die Evaluationsstudien die Wirksamkeitsindikatoren in unterschiedlichen zeitlichen Abständen messen, und dabei sowohl die unmittelbaren oder kurzfristigen Veränderungen – die vorübergehender Natur sein dürften – als auch die längerfristigen Veränderungen berücksichtigen¹⁰³.

4.1d Massgeschneiderte Programme?

Die Täter häuslicher Gewalt sind nicht alle gleich. Es wird nach wie vor über die Notwendigkeit diskutiert, Programme zu erarbeiten, die sich je nach Täterprofil oder Gewaltschema unterscheiden, wobei hier insbesondere die Verfechterinnen und Verfechter der (feministisch orientierten) psychoedukativen Ansätze und jene der psychodynamischen Ansätze geteilter Meinung sind.

Aus praktischer Sicht sei jedoch zum einen auf die Vielfalt der Evaluationen vor dem Programm und zum anderen auf die Tatsache hingewiesen, dass viele Rechtsordnungen oft nicht verschiedene Interventionsprogramme anbieten.

Zu vier besonderen Aspekten wurden vertiefte Studien durchgeführt: (1) die Persönlichkeitsmerkmale und die Tätertypen; (2) ihre soziodemografischen Merkmale; (3) ihre soziokulturelle Zugehörigkeit; und (4) bei einigen von ihnen das Bestehen von Problemen der psychischen Gesundheit oder von Suchtproblemen, die mit der Anwendung von Gewalt einhergehen. Die Forschung versuchte, den Einfluss dieser verschiedenen Parameter auf die Teilnahme in der Gruppe (Abbruchquote) und die Wiederholung von Gewalttaten systematisch zu untersuchen, um Argumente für eine «individualisierte» Betreuung zu identifizieren.

Persönlichkeitsmerkmale und Tätertypen

Die Vorstellung, dass eine bessere Differenzierung der Täter nötig ist, um sie entsprechend ihren Bedürfnissen spezifischen Programmen zuzuweisen, hat ihren Ursprung in den Arbeiten zur Typologie der Täter häuslicher Gewalt.¹⁰⁴ Dieses Anliegen wurde

¹⁰⁰ Eckhardt *et al.* (2013a, 2013b), Edleson (2012) und Gondolf (2004).

¹⁰¹ Was zeigt, dass eine Beobachtung nach Beendigung des Programms wichtig ist, siehe z. B. Edleson (2012).

¹⁰² Gondolf (2004) und Weisz (2001). Dabei wird nicht berücksichtigt, dass diese Rückfälle in der Regel die Tat einer begrenzten Anzahl von Einzelpersonen sind, bei denen sich die Anwendung von Gewalt zurzeit nicht vermindern lässt. Bei Gondolf (2004) beispielsweise sind 20 Prozent der Männer verantwortlich für 80 Prozent der neuen Delikte, die nach der Aufnahme in das Programm begangen wurden.

¹⁰³ Das Projekt IMPACT sieht daher ein Beobachtungsschema mit fünf Zeitebenen vor; (T₀) der Täter wird dem Programm zugewiesen; (T₁) die Einzelberatung oder das Gruppenprogramm beginnt; (T₂) während der Einzelberatung oder des Gruppenprogramms; (T₃) am Ende der

Einzelberatung oder des Gruppenprogramms; und (T₄) nach einer Dauer von beispielsweise 6 Monaten nach Beendigung der Einzelberatung oder des Gruppenprogramms (Hester *et al.* 2014).

¹⁰⁴ Zum einen siehe z. B. Holtzworth-Munroe & Meehan (2004) oder Cavanaugh & Gelles (2005) für die Typologien, die in Anlehnung an die Persönlichkeitstypen konstruiert wurden, sowie Capaldi & Kim (2007) für eine Einbettung in den

durch eine Reihe von Studien gefördert, die eine unterschiedliche Wirksamkeit bestimmter Interventionsmodelle auf die Abbruchquoten und die Wiederholung häuslicher Gewalt feststellen.¹⁰⁵

Diese Auffassung ist jedoch nicht unumstritten. So stützen verschiedene Vertreter aus Forschung und Praxis wie White & Gondolf (2000: 486) den Interventionsgrundsatz «one size fits most». Sie sind der Ansicht, dass sich die Männer, ungeachtet ihres typologischen Profils, in den Programmen mehr oder weniger identisch verhalten und dass der Persönlichkeitstyp keinen Prädiktor für den Erfolg eines Programms darstellt.¹⁰⁶ Chung (2015) weist darauf hin, dass der Bezug auf typologische Kategorien auch mit einem Risiko einhergeht. Da die Kategorien oft mit besonderen Risikoniveaus in Verbindung gebracht werden, könnte ihre Anwendung, wenn sie prädiktiv erfolgt, zu einer Unterschätzung der Gefahr für das Opfer führen.

Soziodemografische Merkmale

Die Studien, die sich mit den Abbruchquoten befassen, legen ebenfalls nahe, dass bestimmte soziodemografische und biografische Merkmale der Täter einen Einfluss auf ihre Bereitschaft haben, das Programm zu Ende zu führen. Dabei handelt es sich hauptsächlich um Evaluationen von nordamerikanischen Programmen psycho-educativer oder kognitiv-verhaltens-therapeutischer Art. Diese kommen zum Schluss, dass junge, unverheiratete Teilnehmer mit tiefem beruflich-sozialem Status und geringem Einkommen, die einer ethnischen Minderheit angehören, eher Gefahr

laufen, das Programm abubrechen.¹⁰⁷ Während eine kriminelle Vorgeschichte eine ungünstige Voraussetzung für die Beendigung des Programms zu sein scheint, ist der aktuelle Kenntnisstand über einen möglichen Einfluss anderer gewaltbezogener Faktoren (z. B. Schwere der Gewalttaten, Missbrauchserfahrung in der Kindheit) auf die Abbruchquoten nicht schlüssig.¹⁰⁸

Diese Ergebnisse legen bis zu einem gewissen Grad nahe, dass die Bedeutung, welche die Täter den familiären und sozialen Bindungen und der Verbundenheit mit den sozialen Institutionen einerseits und der Einhaltung der Gemeinschaftsnormen andererseits beimessen, einen Einfluss auf die regelmässige Teilnahme haben.

Auch falls sich diese statistischen Assoziationen bestätigen würden, so ginge es nicht darum, besondere Programme oder Inhalte zu erarbeiten, sondern die Akteure müssten die Ansätze identifizieren, mit denen sich Täter mit schwachen familiären und sozialen Bindungen oder einem weniger konformistischen Profil zur Teilnahme am Programm bewegen lassen.

Ethnische Zugehörigkeit

Die sogenannten kulturspezifischen oder kultursensiblen Ansätze¹⁰⁹ beziehen sich zum einen auf die ethnischen Minderheiten in den Vereinigten Staaten¹¹⁰ und zum anderen auf

Kontext. Zum anderen siehe Johnson (2008) für eine Typologie, die in Anlehnung an die Schwere der Gewalttaten und das Bestehen von Kontrollverhalten entwickelt wurde.

¹⁰⁵ Z. B. Corvo, Dutton & Chen (2008), Eckhardt, Holtzworth-Munroe, Norlander, Sibley & Cahill (2008) und Stith *et al.* (2004). Auch in Europa; in Spanien beispielsweise werden sowohl die Dauer als auch die Inhalte des Programms «Galicia» an das Täterprofil angepasst (Arce, Fariña, 2010).

¹⁰⁶ Gondolf (2012: 127–66).

¹⁰⁷ Nach Jewell & Wormith (2010); siehe auch die amerikanischen Studien von Bennett, Hsieh & Stoops (2010) und Feder & Dugan (2002), aber auch Echeburúa (2013), der im spanischen Kontext auch einen Zusammenhang zwischen Randständigkeit und Programmabbruch sieht.

¹⁰⁸ Jewell & Wormith (2010). Echeburúa & Fernández-Montalvo (2009) stellten in Spanien hingegen eine Zunahme der Wahrscheinlichkeit eines Programmabbruchs bei Tätern fest, die eine lange kriminelle Vorgeschichte haben.

¹⁰⁹ Gondolf & Williams (2001) beispielsweise unterscheiden 3 Arten von Programmen für afroamerikanische Gewalttäter: die «color blind» Programme, die diesen Dimensionen keine Bedeutung beimessen, die «culturally-specific» Programme, die gemeinsame Kulturelemente der verschiedenen ethnischen Gruppen, die unter den Teilnehmern vertreten sind, einschliessen, und die «culturally centered» Programme, die auf eine bestimmte ethnische Gruppe ausgerichtet sind.

¹¹⁰ Afroamerikanische oder hispanische Täter; siehe Gondolf & Williams (2001) und Gondolf (2012: 155–6).

die indigenen Bevölkerungsgruppen in Australien und Kanada¹¹¹ und am Rande auf die zugewanderten Bevölkerungsgruppen in einem westlichen Kontext.¹¹²

Einige wenige Studien verweisen auf die Relevanz von Modellen, die besondere soziokulturelle Inhalte einschliessen, oder auf besonderen Formen der Programmdurchführung oder auf spezifischen Aktivitäten basieren. Diese Anpassungen bei den herkömmlichen Programmen würden zur Reduktion sowohl der Abbruchquoten als auch der Wiederholung häuslicher Gewalt beitragen.¹¹³ In der Praxis wird auch auf die Relevanz von kulturellen Kompetenzen bei der Arbeit vor Ort verwiesen.¹¹⁴

Suchtverhalten und begleitende psychische Gesundheitsprobleme

Die Forschung hat bei den Tätern häuslicher Gewalt eine hohe Suchtprävalenz¹¹⁵ sowie begleitende Störungen der psychischen Gesundheit¹¹⁶ nachgewiesen. Männer, die Interventionsprogrammen zugewiesen werden, weisen nicht nur häufiger psychische Störungen und/oder Suchtprobleme auf als Männer in der Gesamtbevölkerung, sondern auch als Täter häuslicher Gewalt, die keinem Programm zugewiesen wurden. Werden diese Fragen angesprochen, so stösst das bei bestimmten Akteuren auf Widerstand, da sie darin einen Versuch sehen, die Gewalttäter von ihrer Verantwortung zu entbinden oder sie gar zu entschuldigen. Doch auch wenn diese Komorbiditäten nicht als Ursache für

gewalttätiges Verhalten angesehen werden, muss erkannt werden, dass durch sie eine Intervention erschwert wird.

Im Zusammenhang mit Suchtproblemen im Allgemeinen und mit Alkoholmissbrauch im Speziellen scheinen spezifische Ansätze notwendig, die eine gleichzeitige Behandlung des Alkoholkonsums und der Ausübung von Gewalt zum Ziel haben.¹¹⁷ Angesichts gewisser grundlegender Unterschiede zwischen den Programmen für Täter häuslicher Gewalt und denjenigen zur Behandlung von Suchtproblemen erweist sich die Herangehensweise als komplex, insbesondere wenn es um die Entscheidung geht, in welcher Reihenfolge welche Intervention durchgeführt werden soll. In der Praxis heisst dies, dass die Akteure über vielseitige klinische und soziale Kompetenzen verfügen müssen, und in der Politik erfordert dies eine Durchlässigkeit der Institutionen.¹¹⁸

Angesichts der ermutigenden Ergebnisse eines kombinierten Ansatzes¹¹⁹ scheint die Weiterführung der theoretischen und praktischen Auseinandersetzung mit der Frage wichtig, wie sich diese beiden Interventionsmodalitäten am besten integrieren lassen.

Ähnliche Diskussionen fanden auch zum gleichzeitigen Auftreten von häuslicher Gewalt

¹¹¹ Mackay *et al.* (2015: 14–7).

¹¹² In Spanien kamen Echauri, Fernández-Montalvo, Martínez & Azkarate (2013) zum Schluss, dass ihr Interventionsprogramm bei Immigranten und Einheimischen gleich wirksam war.

¹¹³ So haben Gondolf & Williams (2001) auf der Grundlage von Opfermeldungen beobachtet, dass ein kleinerer Teil der afroamerikanischen Täter das Programm beendeten (52 % gegenüber 82 %) und dass bei ihnen die Wahrscheinlichkeit einer Verhaftung zwar zweimal grösser (13 % gegenüber 5 %), jene einer Wiederholungstat hingegen kleiner war (32 % gegenüber 39 %).

¹¹⁴ Baobaid, Kovacs, MacDiarmid & Tremblay (2015).

¹¹⁵ Braaf (2012) und Galvani (2010).

¹¹⁶ Cavanaugh & Gelles (2005) und Holtzworth-Munroe & Meehan (2004).

¹¹⁷ Insb. weil Alkohol manchmal ein Faktor für einen Ausschluss aus den Interventionsgruppen ist und weil das Risiko wiederholter Gewalttaten durch Alkoholkonsum stark erhöht wird (Gondolf, 2012).

¹¹⁸ Für eine Illustration im Bereich der Bekämpfung von häuslicher Gewalt und von Alkoholmissbrauch siehe das englische Projekt Stella, «Stella Toolkit» (2007) und «Complicated Matters» (2013); dieses Projekt wurde kürzlich in der Deutschschweiz von Aeberhard & Frei (2015) angepasst und wird bald auch in der Romandie zur Verfügung stehen.

¹¹⁹ In einer amerikanischen Studie war die Rückfallwahrscheinlichkeit bei Männern, die an einem kombinierten Programm teilnahmen, zweimal kleiner als bei jenen der Kontrollgruppe (6 % bzw. 14 %), siehe Goldkamp (1996); für jüngere Studien siehe Gilchrist, Blazquez *et al.* (2015) und Gilchrist, Radcliff *et al.* (2015). Die Suchtbehandlung im Kontext von häuslicher Gewalt ist ebenfalls ein zentrales Anliegen der Vertreterinnen und Vertreter von Programmen für Täter häuslicher Gewalt in Europa, insbesondere in Spanien, auch wenn die Projekte dort eher im Rahmen von parallelen als von integrierten Behandlungen erfolgen (Arce, Fariña, 2009; Carbajosa, Boira Sarto, 2013).

und Problemen psychischer Gesundheit statt (z. B. Depression, Angst, Borderline-Persönlichkeitsstörung, Bindungsstörung), wobei insbesondere eine Neuausrichtung der Programme auf psychodynamische Interventionen empfohlen wird. Doch die unterschiedliche Qualität der Untersuchungen vor Programmbeginn beziehungsweise deren fehlende Systematik schränkt die Aussagekraft der Schlussfolgerungen zur Relevanz – und Machbarkeit – solcher Programme ein.¹²⁰

Mehrere spanische Studien weisen beispielsweise darauf hin, dass der Zusammenhang zwischen der Abbruchwahrscheinlichkeit und dem Täterprofil, insbesondere was die psychiatrische Geschichte und die Suchterfahrungen angeht, grösser ist als der Einfluss der Art und der Schwere der Gewalttaten.¹²¹ Die Abbrüche würden vorwiegend in den ersten Wochen erfolgen. Die Hälfte der Abbrüche fänden sogar während den ersten beiden Sitzungen statt¹²² und seien zahlreicher unter den Tätern mit einer einstellungsbedingten (z. B. Machismus) oder psycho(patho)logischen Auffälligkeit.¹²³

Es scheint noch verfrüht, Stellung zu beziehen. Wir teilen jedoch die Einschätzung von Day *et al.* (2009b), dass es möglich ist, die vielfältigen Bedürfnisse der Täter anzuerkennen, ohne dass ihnen damit die Verantwortung für die begangenen Taten abgenommen wird.

4.1e Ein wenig Zuckerbrot und viel Peitsche

Mit den ersten Studien in diesem Bereich kam der Gedanke auf, dass das System, in welches das Täterprogramm integriert wird, für dessen Wirksamkeit von zentraler Bedeutung ist. Hinter dieser Idee verbergen sich sehr

unterschiedliche Realitäten: das Duluth-Modell, ein koordinierter Gemeinschaftsansatz, Verfolgung von Amtes wegen oder Case Management.¹²⁴

Kurz gesagt, die Männer würden das Programm weniger häufig abbrechen und auch weniger häufig rückfällig werden in Systemen, die: (1) eine sehr rasche Integration in ein Programm ermöglichen (2–3 Wochen); (2) die regelmässige Teilnahme der Täter fortlaufend überprüfen (Compliance); und (3) die Personen, die die Regeln nicht einhalten oder das Programm abbrechen, sanktionieren.¹²⁵ Eine Verbesserung des Systems könnten beispielsweise eine Zuweisung zu einem Programm bereits in der Phase vor der Verurteilung («pretrial») sowie ein beschleunigtes Verfahren («fast track») bieten, welches von auf häusliche Gewalt spezialisierten Gerichten durchgeführt wird. Die entsprechenden Evaluationsdaten sind jedoch vorläufiger Art, und es liegen nicht genügend Informationen zu ihren allfälligen negativen Konsequenzen vor.¹²⁶

Für jedes dieser Modelle ist eine umfangreiche Koordination unter den Justizbehörden und den Institutionen, die die Täterprogramme durchführen, erforderlich. Zudem stellen sich Probleme hinsichtlich der Vertraulichkeit: Welche Informationen müssen – oder können – weitergegeben werden und an wen? Die Beteiligten fühlen sich oft nicht wohl dabei, wenn sie weitere Informationen als nur die Teilnahme an dieser oder jener Sitzung bekannt geben müssen, beispielsweise ob der Täter frühere oder neue Gewalttaten zugegeben hat. Dazu kommt, dass gewisse Forschungs- und Praxiskreise einer Evaluation der Verhaltensänderungen, die ausschliesslich

¹²⁰ Gondolf (2011).

¹²¹ Boira Sarto & Jodrá Esteban (2010). Im Gefängnisumfeld habe jedoch eine lange kriminelle Vorgeschichte negative Auswirkungen auf die Beendigung eines Programms (Fernández-Montalvo, 2009).

¹²² Z. B. Boira Sarto & Jodrá Esteban (2010) und Echeburúa, Sarasua, Zubizarreta & De Corral (2009). Die Abbruchquote in spanischen Programmen, die von der Justiz angeordnet werden, sei hingegen tiefer (10 %).

¹²³ Echeburúa (2013).

¹²⁴ Im Sinne von Case Management in englischen Texten. Im Zusammenhang mit diesem Begriff findet sich manchmal auch die Idee, dass eine individualisierte Evaluation mit der Formulierung von besonderen Zielen eine Ergänzung zur Gruppenarbeit sein könnte (Vlais, 2014).

¹²⁵ Nach Gondolf (2004), siehe auch Day *et al.* (2009a).

¹²⁶ Eckhardt *et al.* (2013a, 2013b).

auf dem Verhalten in der Gruppe beruht, kritisch gegenüberstehen.¹²⁷

Wenn von Case Management gesprochen wird, ist damit oft auch die Praxis bestimmter Institutionen oder Organisationen gemeint, die Opfer (Partnerinnen oder Ex-Partnerinnen) während des Programms zu kontaktieren.¹²⁸ Das Anliegen, mit den Partnerinnen in Kontakt zu bleiben, erscheint umso wichtiger, als diese differenziertere – oder gar abweichende – Angaben zur Situation und zu den «Fortschritten» der Täter machen können.¹²⁹

4.1f Kann ein Täter zur Teilnahme verpflichtet werden?

Diese Frage verweist auf den Unterschied zwischen freiwilliger und angeordneter Teilnahme an einem Täterprogramm.

Einige Studien legen nahe, dass diejenigen Täter, die ein Programm zu Ende führen, auch diejenigen mit der grössten Motivation für eine Veränderung sind. Ob ein Programm beendet wird, ist aber auch abhängig von der Art der Teilnahme des Täters (z. B. davon, wer ihn geschickt hat und unter welchen Umständen) und von den Vorteilen einer Teilnahme. Demnach wäre die Behauptung, dass die freiwilligen Teilnehmer zwangsläufig motivierter sind als die von der Justiz zugewiesenen, ein Vorurteil. Die Tatsache, dass die zugewiesenen Teilnehmer überwacht und bei Verstössen gegen die Regeln sanktioniert werden, ist für einige ebenfalls eine starke Motivation. Doch auch bei strengen Rahmenbedingungen einer Intervention bleibt deren Erfolg teilweise abhängig vom Motivationsgrad der Täter.¹³⁰ Auf jeden Fall

betrifft die Zuweisung zu einem Programm nur einen Teil der Täter häuslicher Gewalt.

Täter, die freiwillig an einem Programm teilnehmen um die elterliche Sorge behalten zu können (extrinsische Motivation), verspüren möglicherweise ebenfalls nicht den Wunsch nach einer realen Veränderung. Die motivationalen Faktoren für die Teilnahme an einem Programm verändern sich jedoch mit der Zeit. Vor allem steigt mit fortschreitender Teilnahme die Wahrscheinlichkeit, dass die Täter eine intrinsische Motivation zur Veränderung entwickeln.¹³¹

Um eine möglichst grosse Zahl von Tätern zu erreichen, empfehlen Fachpersonen die Errichtung von Programmen für Täter häuslicher Gewalt, zu denen Betroffene von der Justiz zugewiesen oder von Institutionen vermittelt werden oder auch freiwillig daran teilnehmen können.¹³² Grund dafür ist, dass die Unterscheidung zwischen angeordneter und freiwilliger Teilnahme manchmal unscharf ist.¹³³

Wichtig ist zudem, besser zu verstehen, wie eine gerichtliche Anordnung die Teilnahme am Programm einerseits und dessen Erfolg andererseits beeinflusst. Dies ist nicht zuletzt von Bedeutung, weil die verschiedenen Formen der gerichtlichen Anordnung wahrscheinlich unterschiedliche Auswirkungen auf die Abbruchquoten haben.¹³⁴

¹²⁷ Eckhardt *et al.* (2013a, 2013b); für eine Diskussion im schweizerischen Kontext siehe auch Gloor & Meier (2002).

¹²⁸ Für Praxisbeispiele siehe den Bericht 2013 des Workshops zum Daphne-Programm WWP, Fussnote 76.

¹²⁹ Zu diesem Punkt siehe hauptsächlich Kelly & Westmarland (2015a), aber auch Eckhardt *et al.* (2013a, 2013b); siehe auch unten S. 36 ff. für einige Erwägungen zum englischen Programm MIRABAL.

¹³⁰ Echeburúa (2013) für ein Gefängnisprogramm und Lila, Catalá, Conchell, García, Lorenzo, Pedrón & Terreros (2010) für ein Strafersatzprogramm.

¹³¹ Für eine Erläuterung siehe z. B. Stanley, Borthwick, Graham-Kevan & Chamberlain (2011). Weiter hat eine in der Schweiz durchgeführte explorative Studie gezeigt, dass eine gerichtliche Anordnung eine starke Wirkung auf den Eintritt in ein Programm hatte, sich diese Wirkung aber abschwächte, wenn die Teilnehmer eine gewisse Zeit am Programm teilgenommen hatten (Jaquier *et al.*, 2016).

¹³² Hester & Lilley (2014) und Kelly & Dubois (2008). Dazu kommt die Notwendigkeit, die Partner, die Organisationen und die amtlichen Institutionen noch besser über den Zweck der Täterprogramme zu informieren, für ein Beispiel siehe Gloor & Meier (2002: 25).

¹³³ Hagemann-White, Kavemann, Beckmann, Brandfass, Grieger, Leopold, Rabe & Katenbrink (2004).

¹³⁴ Gemäss z. B. den Arbeiten von Barz, Helfferich, Kavemann, Aksu-Yagci, Breunig & Sommer (2006) in Deutschland.

4.1g Und wenn die Täter nicht motiviert sind für eine Verhaltensänderung?

Neben den Massnahmen zur Erhöhung der Zahl der Sitzungen und des Verbleibs der Teilnehmer in den verschiedenen verfügbaren Interventionsprogrammen sollte denjenigen, die eine nachhaltige Veränderung bei den Teilnehmern begünstigen, besondere Aufmerksamkeit beigemessen werden. Ein besseres Verständnis der Motivationsarten und -quellen und ihrer Auswirkungen auf eine erfolgreiche Teilnahme und auf die Einstellungs- und Verhaltensänderungen stellt einen Meilenstein in der Entwicklung von Täterprogrammen dar.

Die Massnahmen zur Steigerung der Motivation der Täter, sich zu verändern, finden seit einiger Zeit grosse Beachtung.¹³⁵ Eine standardisierte Evaluation des anfänglichen Grads der Veränderungsmotivation, eine Erhöhung der Anzahl Sitzungen, um die Teilnehmer ständig neu zu motivieren, und eine Begleitung der Täter in einem schrittweisen Prozess der Verantwortungsübernahme: All diese Massnahmen sind Teil einer therapeutischen Arbeit, die auf Veränderung ausgerichtet ist.¹³⁶ Die Faktoren, die eine Einstellungs- oder Verhaltensänderung auslösen und dazu beitragen, diese über die Zeit aufrechtzuerhalten, sind jedoch noch nicht identifiziert.¹³⁷

Das Transtheoretische Modell, das Prochaska & DiClemente im Suchtbereich entwickelten (1984), fand im Bereich der häuslichen Gewalt zunehmend Beachtung, insbesondere bei den Vertreterinnen und Vertretern von psychodynamischen Programmen. Dieses Modell versteht die Motivation einer Person zur Veränderung als einen stufenweisen

Prozess.¹³⁸ Die Voraussetzung für jede Veränderung ist, dass die betroffene Person versteht, dass ihr Verhalten problematisch ist und geändert werden muss. Dieses Modell trägt auch der Tatsache Rechnung, dass sich Personen nicht linear entwickeln. Es kann sogar vorkommen, dass einige Personen mehrmals alle Stufen des Programms durchlaufen, bevor eine tatsächliche Veränderung stattfindet. Es weist zudem darauf hin, dass verschiedene externe Faktoren die Motivation der Täter zur Veränderung beeinflussen können, namentlich wechselnde Wohnsituationen, Interaktionen mit der Justiz, ein allfälliger Bruch mit der Partnerin. All diese Elemente sind zu berücksichtigen, wenn eine Veränderung begünstigt werden soll.

Die Evaluationsstudien zu den Programmen, welche die Grundsätze des Transtheoretischen Modells mit einbezogen haben, zeigen vielversprechende Ergebnisse, angefangen bei einer Abnahme von Angst- und Depressions-symptomen der Teilnehmer bis hin zu einer besseren Kontrolle von Wutreaktionen.¹³⁹ Der Einsatz von Methoden oder Techniken zur Motivationssteigerung für eine Verhaltensveränderung der Männer reduziert sowohl die Programmabbruchrate als auch die Gewaltrückfälle.¹⁴⁰

Die eingesetzten Techniken stammen hauptsächlich aus der Motivierenden Gesprächsführung. Diese hat zum Ziel, die Bedenken der Patienten/Teilnehmer abzubauen und sie zum Denken anzuregen,

¹³⁵ In gewissen Programmen sei ein tiefer Motivationsgrad ein Ausschlusskriterium (Barz *et al.*, 2006; Kraus, 2013).

¹³⁶ Stanley *et al.* (2011).

¹³⁷ Edleson (2012) und Mackay *et al.* (2015). Zum Konzept der Veränderung siehe auch die qualitativen Daten, die im Rahmen des englischen Projekts MIRABAL erhoben wurden (Kelly, Westmarland, 2015a).

¹³⁸ Die Stadien in den am weitesten entwickelten Versionen des Modells sind: Absichtslosigkeit («Precontemplation», keine Anerkennung des Problems / keine Veränderungsabsicht); Absichtsbildung («Contemplation», Absicht, das Verhalten in den kommenden 6 Monaten zu verändern); Vorbereitung («Preparation», Absicht, innerhalb 1 Monats Massnahmen zu ergreifen); Handlung («Action», spezifische, sichtbare Verhaltensänderungen in den vergangenen 6 Monaten); Aufrechterhaltung («Maintenance», Konsolidierung, Rückfallprävention); und Abschluss («Termination», der Veränderungsprozess ist abgeschlossen); nach Casey, Day & Howells (2005: 159).

¹³⁹ Alexander & Morris (2008), Day *et al.* (2009a), Eckhardt (2004) oder Eckhardt *et al.* (2008), Murphy & Baxter (1997) oder Murphy & Maiuro (2009b) sowie Musser, Semiatin, Taft & Murphy (2008).

¹⁴⁰ Babcock *et al.* (2004).

indem sie eine nicht wertende Rückmeldung zu ihrem problematischen Verhalten bekommen.¹⁴¹

So haben Eckhardt *et al.* (2013a, 2013b) 10 Studien untersucht, die sich auf die Motivierende Gesprächsführung und/oder das Transtheoretische Modell berufen. Sie stellen eine grosse Variabilität der Interventionen, der Aktivitäten und deren Intensität fest: von einer oder zwei vorgängigen Motivationssitzungen bis zum Einbezug aller Leitsätze des Transtheoretischen Modells in ein vier- bis sechsmonatiges Programm. Weiter weisen sie auch auf Unterschiede bei der Art der Programmdurchführung hin: aus Einzelpersonen bestehende Gruppen, gemeinsame Betreuung von Paaren oder auch Paartherapie. Trotz dieser Unterschiede seien die ersten Ergebnisse ermutigend und vor allem solid, da sie oft im Rahmen von randomisierten Studien erzielt wurden. Es sei ein leichter Rückgang der Wiederholung häuslicher Gewalt zu beobachten sowie positive Auswirkungen auf die Teilnahme und die Verantwortungsübernahme.¹⁴² Die Schlussfolgerungen zum Verbleib der Teilnehmer in einem Programm waren hingegen weniger klar.¹⁴³ Schliesslich wurden unmittelbar nach Beendigung des Programms gewisse Veränderungen beobachtet, ihre Nachhaltigkeit ist aber nach wie vor unzureichend belegt.¹⁴⁴

Die gegen diese «neuen» Ansätze vorgebrachte Kritik betrifft zurzeit genau die Schwierigkeit, den Grad der Veränderungsbereitschaft zu messen und diesem die entsprechende Handhabung

zuzuweisen.¹⁴⁵ Einige sind jedoch der Ansicht, dass solche Ansätze einen empathischen Umgang mit den Tätern häuslicher Gewalt erforderlich machen, was eine Verantwortungsübernahme behindern könnte.

Das Konzept der therapeutischen Beziehung ist im Kontext von häuslicher Gewalt schwierig zu untersuchen. Es verweist hier gleichzeitig auf den kooperativen Aspekt der Beziehung, auf die affektive Verbindung zwischen Klient/Teilnehmer und Therapeut/-in beziehungsweise Programmleiter/-in und auf die Fähigkeit der beiden, sich auf die Interventionsaktivitäten und -ziele zu einigen.¹⁴⁶ Die therapeutische Beziehung soll die Teilnahme am Programm erleichtern, die regelmässige Anwesenheit der Teilnehmer steigern und zu einer Reduktion der Rückfälle beitragen. Sie ist besonders zu Beginn eines Programms wichtig, wenn die Teilnehmer mit dem Stigma konfrontiert sind, das den Tätern häuslicher Gewalt anhaftet.¹⁴⁷ Eine Vertrauensbeziehung und ein Vertraulichkeitsverhältnis sind für die Schaffung eines Kontextes, in dem die Täter Hilfe in Anspruch nehmen – und erhalten – entscheidend.¹⁴⁸ Für bestimmte Fachpersonen wie Vlais (2014: 6–8) besteht jedoch das Risiko, dass der therapeutischen Beziehung Vorrang vor dem Schutz der Opfer gegeben wird. Vlais anerkennt zwar die Bedeutung eines günstigen Umfelds, einer Vertrauensbeziehung und des Gefühls der emotionalen Sicherheit zwischen Teilnehmer(n) und der Programmleitung und akzeptiert, dass letztere eine Verbündete der Täter in ihren Veränderungsbemühungen sein kann, lehnt jedoch jegliche Parallele zwischen diesen Prozessen und der Schaffung einer therapeutischen Beziehung ab.

In jedem Fall sind diese Arbeiten vor allem auf die Frage ausgerichtet, wie die Männer die verschiedenen Stufen des Veränderungs-

¹⁴¹ Insb. Miller & Rollnick (2002). Bei den Problemen, die von den Akteuren genannt werden, wird wiederholt auf einen Mangel an «Motivation» hingewiesen, der sogar zu einem Ausschlussfaktor werden kann (Barz *et al.*, 2006; Kraus, 2013).

¹⁴² Musser *et al.* (2008) und Stuart *et al.* (2007); siehe auch Walker & Bowen (2015) für ein Mentoring-Pilotprojekt in England oder Subirana-Malaret & Andrés-Pueyo (2013) für einen spanischen proaktiven Ansatz für das Halten der Teilnehmer im Programm.

¹⁴³ Nach Mackay *et al.* (2015); für Beispiele siehe z. B. De Los Galanes & Tabernero (2013), Kistenmacher & Weiss (2009) oder Taft, Murphy, Elliott & Morrel (2001).

¹⁴⁴ Murphy & Maiuro (2009a).

¹⁴⁵ Gondolf (2012: 138–47).

¹⁴⁶ Day *et al.* (2009a).

¹⁴⁷ Gemäss Stanley *et al.* (2011), insb. in Bezug auf die Arbeiten von Horvath & Bedi (2002) und von Sharf, Primavera & Diener (2010); siehe auch Hellman, Johnson & Dobson (2010).

¹⁴⁸ Campbell, Neil, Jaffe & Kelly (2010).

prozesses durchlaufen. Unsere Kenntnisse der Mechanismen, die diesen Entwicklungen zugrunde liegen, oder der Art, wie Einstellungs- und Verhaltensänderungen begünstigt werden können, werden damit nur marginal verbessert.

Die Arbeiten, die die Umstände vor dem Eintritt in ein Programm betrachten, weisen auf die Komplexität der Veränderungsprozesse hin und zeigen auf, dass die Täter zahlreiche Faktoren sowohl persönlicher (z. B. Identitätsaspekte, Stress) als auch zwischenmenschlicher (z. B. Konfrontation mit ihrer Familie und ihren Bekannten) und sozialer Art (z. B. Beschäftigung, sozio-ökonomischer Status) meistern müssen. Die Auswirkungen dieser Faktoren auf die Teilnahme an einem Programm und auf die Wiederholung häuslicher Gewalt sind nach wie vor unzureichend belegt.¹⁴⁹

4.1h Ist es zu zweit wirklich besser?

Trotz einiger ausgeprägter Widerstände hat die Paartherapie¹⁵⁰ im Bereich der häuslichen Gewalt allmählich an Bedeutung gewonnen. Die sehr früh geäußerte Kritik an dieser Therapie¹⁵¹ stammt hauptsächlich von Frauenorganisationen, die sich für die Opfer häuslicher Gewalt einsetzen, und aus Forschungsarbeiten, denen eine stenge(re) feministische Perspektive zugrunde liegt.¹⁵² Diese Einwände sind philosophischer – das systemische Interventionsparadigma verteile die Verantwortung für die Gewalttaten– und

pragmatischer Art – das Gewaltrisiko würde durch die gemeinsame Arbeit erhöht.¹⁵³

Die Ansicht, dass ein systemischer Ansatz das Opfer teilweise für die vom Täter begangenen Gewalttaten verantwortlich macht, sei eine Verkenning der gegenwärtigen systemischen Perspektiven, die auf komplexen mehrdimensionalen Kausalitätsmodellen basieren. Die Verfechterinnen und Verfechter der Paarprogramme sind der Ansicht, dass die Gewalttäter zur Verantwortung gezogen werden können und gleichzeitig an den Interaktionsarten des Paares gearbeitet werden kann.¹⁵⁴ Die häusliche Gewalt sei eine Realität in vielen Paartherapien, unabhängig davon, ob sich die Therapeutinnen und Therapeuten dessen bewusst sind oder nicht, und es sei daher entscheidend, die Interaktionsschemata von Paaren, für die eine Paartherapie klinisch indiziert(er) ist, besser zu identifizieren.¹⁵⁵

Es existieren nicht sehr viele Studien zu Paarprogrammen. Sie betreffen mehrheitlich Programme, die auf strengen Auswahlkriterien der Teilnehmenden beruhen, mit dem Ziel, Paare mit einem tiefen Gewaltniveau aufzunehmen. Im Allgemeinen werden diese Programme mit Einzelsitzungen ergänzt, an denen ein Debriefing der Paarsitzungen stattfindet und beide Partner zusätzliche individualisierte Unterstützung erhalten. Gewisse Programme umfassen auch geschlechtergetrennte Gruppensitzungen, die gleichzeitig durchgeführt werden. Die therapeutische Arbeit ist auf die Entwicklung von kognitiven Verhaltenskompetenzen, manchmal auch von Kommunikations- und

¹⁴⁹ Insb. Curwood, Degeer, Hymmen & Lehmann (2011), und vor allem die von Walker geleiteten Arbeiten in England rund um das Konzept «catalyst for change», z. B. Walker & Bowen (2015) und Walker, Bowen, Brown & Sleath (2015a, 2015b).

¹⁵⁰ Das Konzept «Paartherapie» wird hier angesichts der Vielfalt der Ansätze in einem weiten Sinn verstanden. Es scheint eher auf das Setting (Arbeit mit Paaren und nicht nur mit den Tätern) als auf einen besonderen therapeutischen Ansatz zu verweisen. In den englischen Texten wird oft der Begriff «couple counselling» verwendet.

¹⁵¹ Für ein Beispiel siehe Bograd (1984).

¹⁵² Für positive Argumente siehe Stith *et al.* (2004) und McCollum & Stith (2008); für negative Argumente siehe Babcock *et al.* (2004) und Gondolf (2012: 147–52).

¹⁵³ Diese Kritik hatte zudem direkte Auswirkungen auf die Politik und die Akkreditierungsverfahren, insbesondere in Nordamerika, wo viele Praxisstandards paartherapeutische Interventionsmodelle verbieten. In einer Übersichtsarbeit über die Praktiken stellten Maiuro & Eberle (2008) fest, dass zwei Drittel der amerikanischen Bundesstaaten, die über Akkreditierungsstandards für die Betreuung von Tätern häuslicher Gewalt verfügen, Paartherapiemodelle ausdrücklich ausschließen. Dieser Entscheid sei jedoch angesichts der Schlussfolgerungen der bis heute verfügbaren Evaluationsstudien verfrüht (McCollum, Stith, 2008); siehe auch Fussnote 24.

¹⁵⁴ McCollum & Stith (2008).

¹⁵⁵ Jacobson & Addis (1993).

Konfliktlösungskompetenzen, ausgerichtet.¹⁵⁶ Diese Programme beinhalten bis zu einem gewissen Grad mehrere Merkmale der üblichen Gruppenansätze.¹⁵⁷

Welche Daten stehen heute zur Wirksamkeit der Paarinterventionen zur Verfügung? Die erste Feststellung ist, dass die Paarprogramme anscheinend ausschliesslich bei einer Teilgruppe von Tätern häuslicher Gewalt durchgeführt worden sind. Die Teilnehmer, die nach einem strengen Evaluationsverfahren ausgewählt wurden, seien nicht repräsentativ für die von der Justiz zugewiesenen Täter. Sie seien mehrheitlich nicht sehr gewalttätig und in einer festen Paarbeziehung, die entweder noch intakt ist oder die sie unbedingt erhalten möchten. Gewisse Fachpersonen stellen daher die auf den ersten Blick ermutigenden Ergebnisse der Paarprogramme in Frage und weisen darauf hin, dass diese bei den veränderungswilligsten Tätern, die auch in einem individuellen Gruppenprogramm erfolgreich gewesen wären, erzielt wurden.¹⁵⁸

McCollum & Stith (2008) haben vier Reihen von Evaluationsstudien zu therapeutischen Ansätzen für Paare identifiziert.¹⁵⁹ Im Voraus sei jedoch darauf hingewiesen, dass fast alle diese Programme von ihren Entwicklern evaluiert wurden.

1 Die erste Forschungsgruppe befasste sich mit dem Potential zur Reduktion häuslicher Gewalt von im Suchtbereich

¹⁵⁶ Die Kritik an den Paartherapien würde sich daher hauptsächlich gegen die «klassischen» Paartherapien, die dem Kontext und den besonderen Dynamiken der häuslichen Gewalt nicht Rechnung tragen, und nicht gegen die Paararbeit an sich richten (Gondolf, 2012: 148).

¹⁵⁷ In bestimmte Fällen laufen die Gruppen- und Paararbeit parallel. Dies ist der Fall in dem von Almeida & Durkin (1999) vorgestellten kontextorientierten Modell aus Nordamerika, in dem die Paarsitzungen in Ergänzung zu den Tätergruppen stattfinden. Die Paararbeit (12 Wochen) ist spezifisch auf die Paarinteraktionen und den Missbrauch und die Gewalt der Frau und des Mannes ausgerichtet. Daneben nehmen die Männer und Frauen separat an geschlechtergetrennten Gruppen teil (6 Wochen).

¹⁵⁸ Eckhardt et al. (2013a, 2013b) und Gondolf (2012: 147–52).

¹⁵⁹ Die folgenden Abschnitte wurden mehrheitlich gemäss McCollum & Stith (2008) verfasst; für Einzelheiten siehe vollständigen Text.

durchgeführten Verhaltenstherapien für Paare. Die in verschiedenen Formaten angebotene Paararbeit soll den Paaren einerseits ermöglichen, spezifische Kompetenzen zu entwickeln und andererseits ihre Interaktionen so zu verändern, dass ein stabiles Familienumfeld geschaffen wird, was ein wichtiger Faktor für eine langfristige Substanzabstinenz ist. Üblicherweise werden diese Sitzungen gemeinsam (15–20 Paare) während einer Dauer von 5–6 Monaten durchgeführt. Die Ausschlusskriterien sind streng, insbesondere hinsichtlich der Schwere der häuslichen Gewalt (z. B. Verletzungen oder Hospitalisierung) und ihrer Art (z. B. Bedrohung mit einer Waffe). Paare werden ebenfalls ausgeschlossen, wenn beide einen Substanzmissbrauch aufweisen.¹⁶⁰ Schliesslich sehen bestimmte Protokolle ausdrücklich den Ausschluss von Paaren vor, bei denen die Frau Bedenken hinsichtlich des Prozesses und des Risikos von Gewalt oder Repressalien durch ihren Partner äussert.¹⁶¹

Die Evaluationen dieser Studienreihe stellen eine signifikante Abnahme der häuslichen Gewalt nach der Teilnahme in der Gruppe fest, die Studien, die zwei Interventionsmodalitäten vergleichen, zeigen hingegen einen grösseren Rückgang der Gewalt bei den Teilnehmern eines Paarprogramms im Vergleich zu den Teilnehmern eines individuellen Programms.¹⁶² Die Wirkung der Suchtprogramme auf die Prävalenz von häuslicher Gewalt sei auf die Tatsache zurückzuführen, dass die Partner Kompetenzen erlernen, die ihnen mehr

¹⁶⁰ McCollum & Stith (2008).

¹⁶¹ Das verhaltenstherapeutische Modell für Paare würde nicht nur eine längere Abstinenz und eine Abnahme der durch Substanzmissbrauch bedingten Probleme bewirken, sondern auch zu unbeschwerteren Paarbeziehungen, zu weniger Trennungen und Scheidungen als bei den Einzelbehandlungen und zu einem deutlichen Rückgang der Gewaltausübung führen (Fals-Stewart, O'Farrell, Birchler, Córdova, Kelley, 2005).

¹⁶² Z. B. Fals-Stewart, Kashdan, O'Farrell & Birchler (2002) und O'Farrell & Fals-Stewart (2000).

Sicherheit verleihen, insbesondere in Situationen, in denen der Partner unter Substanzeinfluss steht.¹⁶³ Diese Interpretation lege auch nahe, dass diese Programme wirksam seien für die Täter, die häusliche Gewalt mehrheitlich unter dem Einfluss von Alkohol oder Drogen ausüben.

- 2 Eine zweite Gruppe von Studien hat verschiedene Ansätze evaluiert, die ausgehend vom Paarprogramm entwickelt wurden, das Peter Neidig ursprünglich für die US-Navy erarbeitet hatte. Es handelt sich um sehr strukturierte Programme, welche die Entwicklung von Konfliktlösungskompetenzen zum Ziel haben, insbesondere der Fähigkeit, die Anzeichen von Gewalt zu erkennen und die Eskalation von Konflikten zu verhindern. Üblicherweise umfasst die Paararbeit wöchentliche Sitzungen, an denen eine Gruppe von 6–8 Paaren während eines Zeitraums von 10 Wochen teilnimmt.¹⁶⁴ Das für das Militär entwickelte Programm wurde von Brannen & Rubin (1996) bei zivilen Teilnehmern getestet. Diese verglichen die Wirksamkeit der Paarintervention mit derjenigen einer feministisch orientierten psychoedukativen Gruppenintervention, die nur für Männer bestimmt ist. Mit Ausnahme der Teilnehmer mit Alkoholproblemen, die im Paarprogramm bessere Ergebnisse erzielten, wurden zwischen den beiden Interventionsmodalitäten keine signifikanten Unterschiede nachgewiesen. Andere Forscherinnen und Forscher erarbeiteten eine zweite, abgeänderte Version des Programms von Neidig,¹⁶⁵ das Techniken zum Umgang mit Wut, Kompetenzen zu Konfliktbewältigung und Kommunikation sowie die Entwicklung der Fähigkeit zur Aushandlung von Alternativen kombiniert. Dieses Programm

wurde bei freiwillig teilnehmenden, Hilfesuchenden Paaren getestet und sollte dazu beitragen, die Anwendung von häuslicher Gewalt zu reduzieren.¹⁶⁶

- 3 Stith, McCollum & Rosen (2011) haben ein auf häusliche Gewalt ausgerichtetes Paarprogramm entwickelt, das auf einem lösungsorientierten Ansatz beruht.¹⁶⁷ Es richtet sich an Paare in einer intakten Beziehung mit leichter bis mittlerer Gewaltanwendung, bei denen beide Partner die Beziehung erhalten und die Gewalt beenden wollen. Das Hauptziel der Intervention ist das Stoppen der häuslichen Gewalt, das Nebenziel die Verbesserung der Beziehungsqualität. Die Aufnahmekriterien für das Programm betreffen auch hier die Schwere und die Art der Gewalttaten. In den ersten sechs Wochen nehmen beide Partner an geschlechtergetrennten Gruppen teil und werden regelmässig beurteilt. Danach nehmen sie – alleine oder mit anderen Paaren – an Paarsitzungen teil, an denen das Co-Therapeutenteam anwesend ist. Jede Sitzung beginnt und endet mit einem nicht gemischten Treffen. Wenn das Paar in einem individuellen Setting betreut wird, treffen der Mann und die Frau je mit einer der Personen des Co-Therapeutenteams zusammen; wenn die Betreuung in einem Gruppen-Setting erfolgt, werden die Männer und Frauen getrennt, und die beiden Gruppen treffen je mit einer der Personen des Co-Therapeutenteams zusammen. Die vorangehenden geschlechtergetrennten Sitzungen haben zum Ziel, über allfällige Gewalttaten zu diskutieren, die in der Paarbeziehung zu besprechenden Themen festzulegen und den Wunsch der Teilnehmenden zu bekräftigen, an der anschliessenden Paarsitzung teilzunehmen. Die Abschluss-sitzungen werden für ein Debriefing genutzt.

¹⁶³ Fals-Stewart & Kennedy (2005).

¹⁶⁴ «Peter Neidig's Domestic Conflict Containment Program» (DCCP), Neidig (1985).

¹⁶⁵ «Physical Aggression Couples Treatment» (PACT), Heyman & Neidig (1997).

¹⁶⁶ O'Leary, Heyman & Neidig (1999).

¹⁶⁷ «Domestic Violence Focused Couples Treatment» (DVFACT), Stith et al. (2011).

In einer randomisierten Studie wurden die beiden Modalitäten – Einzel oder Gruppenbetreuung – dieses Paarprogramms mit einer Kontrollgruppe, die keine Intervention erhielt, verglichen. Ausgehend vom Vergleich des Verlaufs von nur 42 Paaren, die am Programm teilgenommen hatten, gelangten die Autorinnen und Autoren zum Schluss, dass es zwischen der Einzel- und der Gruppenbetreuung keinen Unterschied gibt. Bei den Paaren, die am Gruppenprogramm teilgenommen hatten, stellten sie einen Rückgang der psychischen und physischen Gewalt und eine Verbesserung der Zufriedenheit in der Partnerschaft sowie eine Abnahme der Einstellungen, die häusliche Gewalt rechtfertigen, fest. Weder bei den Paaren, die am Einzelprogramm teilnahmen, noch bei jenen der Kontrollgruppe wurden signifikante Ergebnisse beobachtet. Bei einem Vergleich der Wiederholung häuslicher Gewalt auf der Grundlage von Daten der Frauen zur Viktimisierung war die Rückfallwahrscheinlichkeit der Männer, die am Programm teilgenommen hatten, kleiner als jene der Kontrollgruppe, und zwar unabhängig von der Betreuungsmodalität. Gemäss Aussagen der Partnerinnen hatten sechs Monate nach Beendigung des Programms nur 5,4 Prozent der Männer erneut Gewalt gegen sie ausgeübt.¹⁶⁸

- 4 Die einzige Evaluation, nämlich die von Dunford (2000) durchgeführte Forschungsarbeit, erfolgte in Form einer experimentellen Studie, mit der die Wirksamkeit von verschiedenen individuell und als Paar besuchte Täterprogramme ermittelt werden sollte. Sie wurde Ende der 1990er-Jahre auf einer Basis der US-Navy durchgeführt und beinhaltete drei Versuchsanordnungen und eine Kontrollanordnung. Insgesamt wurden 861 Paare zufällig einer der vier folgenden Gruppen zugewiesen: (a) einem kognitiv-verhaltens-therapeutischen Programm, das während

26 Wochen in einer Gruppe stattfand, gefolgt von 6 monatlichen Sitzungen; (b) einem Paarprogramm¹⁶⁹, das während 26 Wochen in der Gruppe stattfand, ebenfalls gefolgt von 6 monatlichen Sitzungen; (c) einer Gruppe mit einer engen behördlichen Überwachung (d. h. 12 monatliche Besuche einer Sozialarbeiterin oder eines Sozialarbeiters und periodische Überprüfung der Wiederholung von häuslicher Gewalt durch Kontaktieren der Ehefrau); und (d) einer Kontrollgruppe, die aus der Teilnahme des Opfers an einem Unterstützungsprogramm bestand.

Ausgehend von Daten zur Viktimisierung, welche die Opfer 6 und 12 Monate nach Beendigung des Programms bereitstellten, wurde kein signifikanter Unterschied zwischen den Gruppen festgestellt. Die beobachteten Rückfallquoten waren im Vergleich zu anderen Studien relativ tief. Die ökologische Validität dieser Ergebnisse ist jedoch beschränkt, da diese Evaluation bei Militärpflichtigen im aktiven Dienst durchgeführt wurde. Die interne Validität der Studie ist aufgrund der geringen Anzahl der Frauen, die an dieser Studie teilgenommen haben, begrenzt. Die Teilnahmequote betrug 2 Frauen auf 5 Männer, was bedeutet, dass nur wenige Paare am Paarprogramm teilgenommen haben. Während die Männer (Militärdienst) dem Programm zugewiesen wurden, nahmen die Frauen (Zivilpersonen) auf freiwilliger Basis teil.

Letztlich könnten Paarprogramme eine Interventionsmöglichkeit sein, wenn sie auf einer sorgfältigen Auswahl der Teilnehmenden beruhen. Sie könnten klinisch indiziert sein in Ergänzung oder im Anschluss an geschlechtergetrennte individuelle Programme oder seltener sogar an Stelle dieser Programme.¹⁷⁰ Es gibt immer noch zu wenige Evaluationsstudien, welche die Wirksamkeit der individuellen Gruppenprogramme und der

¹⁶⁸ Stith et al. (2004).

¹⁶⁹ Dieses psychoedukative Programm wurde von Geffner, Sorenson & Lundberg-Love (1997) entwickelt.

¹⁷⁰ Stuart et al. (2007).

Paarprogramme ermitteln. Trotz der methodischen Qualität dieser Studien sind die festgestellten Wirkungen schwach, bisweilen sogar inexistent.

Diese Ergebnisse verdienen eine kritische Würdigung, nicht nur in Bezug auf die Wirksamkeit der Paarprogramme bei einer spezifischen Tätergruppe, sondern auch im Hinblick auf die Sicherheit der Opfer, die an solchen Programmen teilnehmen.¹⁷¹

4.2 Verhaltens- und Einstellungsänderungen sowie psychologische Veränderungen

Mehrere Evaluationsstudien zu den Täterprogrammen haben den Versuch unternommen, eine Reihe von Veränderungsindikatoren in ihr Studiendesign zu integrieren. Die in diesen Studien betrachteten Veränderungen umfassen Einstellungs- und Verhaltensänderungen sowie Veränderungen von Emotionen und Problemen der psychischen Gesundheit.¹⁷² Dieses Vorgehen ist identisch wie bei Ansätzen, die sich in stärkerem Mass auf ein psychodynamisches Modell berufen.

In Europa beispielsweise werden psychometrische Daten vorwiegend in spanischen und portugiesischen Studien verwendet, wo mit validierten Instrumenten eine breite Palette von selbstberichteten Daten erhoben wurde. Die Verwendung von aktuarischen Instrumenten zur Risikoeinschätzung beschränkt sich auf Süd- und Westeuropa, während die Mehrheit der

untersuchten Studien erfassten, inwieweit die Teilnehmer sexistische Einstellungen billigen und welche Haltung sie gegenüber Frauen und geschlechtsspezifischer Gewalt vertreten. In einigen Studien wurden auch unschärfere Konzepte wie die Lebensqualität des Täters und seiner Partnerin,¹⁷³ das Wohlbefinden der Partnerin oder die Motivation des Täters – oder gar des Paares –, ohne Gewalt zu leben,¹⁷⁴ evaluiert.

In einer Zusammenfassung aller Ergebnisse gelangt das Projekt IMPACT zum Schluss, dass die Teilnahme an einem Programm nicht nur die Abnahme von häuslicher Gewalt begünstigen, sondern auch die Art der begangenen Gewalttaten verändern und das Viktimisierungsrisiko senken würde. Die Programmteilnahme erlaube, die Haltungen der Teilnehmer gegenüber den Frauen und die Ausübung von Gewalt im Allgemeinen und von geschlechtsspezifischer Gewalt im Besonderen zu verändern. Die mit der Wiederholung häuslicher Gewalt einhergehenden psychologischen Symptome würden verbessert und die Programmteilnahme zur Reduktion bestimmter Risikofaktoren beitragen, die mit der Ausübung von häuslicher Gewalt im Zusammenhang stehen (z. B. Substanz-missbrauch), und bestimmte Schutzfaktoren stärken (z. B. Lebensqualität). Gemäss einer kleineren Zahl von Studien lasse sich eine grössere Motivation der Täter zur Programmteilnahme, ein grösseres Engagement im Programm und generell eine Verbesserung der allgemeinen Zufriedenheit beobachten.

Die Ergebnisse des Projekts IMPACT sind zwar ermutigend, doch es ist zu beachten, dass die verschiedenen Studien dieser Zusammenfassung methodisch schwach sind, insbesondere aufgrund der oft fehlenden Kontrollgruppe, der kleinen Grösse der Stichproben und der Verzerrungen infolge des statistischen Umgangs mit den Programmabbrüchen.

¹⁷¹ Babcock *et al.* (2004).

¹⁷² Z. B. Arrigoni, Jiménez, Navarro & Mendoza (2013) für Ergebnisse zu kognitiven Verzerrungen (cognitive distortion) und sexistischen Einstellungen; Boira Sarto, López, Tomás-Aragonés & Gaspar (2013) zur Aufrichtigkeit der Täter; Diranzo, Murillo & Miñana (2012) für Daten zu Depression und Teilnahme am Gemeinschaftsleben; Echeburúa & Fernández-Montalvo (2009) und Echeburúa (2013) zum Selbstwertgefühl und den verschiedenen Störungen der psychischen Gesundheit; oder auch Ramírez, Giménez-Salinas Framis & De Juan Espinosa (2013) für Daten zu sexistischen Einstellungen, Verantwortungsübertragung, Impulsivität, psychische Gewalt und Ausdruck und Kontrolle von Wut; sowie Lila *et al.* (2010) zur Rechenschaftspflicht der Täter und Schuldzuweisung an das Opfer.

¹⁷³ Kraus (2013).

¹⁷⁴ Barz *et al.* (2006).

4.3 Eine verzerrte Definition des Erfolgs

Kelly & Westmarland (2015b) sind der Ansicht, dass die Definition des Kriteriums für den Erfolg der Täterprogramme zu eng gefasst wurde, und kritisieren, dass diese Definition nutzlos ist. Zudem hätten diese Programme höheren Beweisstandards zu genügen als andere (psycho-)soziale Interventionen. Zwar sind sich die Vertreterinnen und Vertreter der Täterprogramme einig, dass die Sicherheit der Opfer – Partnerinnen oder Ex-Partnerinnen – oberste Priorität hat und die Verhaltensänderungen der Täter zweitrangig sind. Es muss jedoch noch geklärt werden, wie ein solcher Erfolg zu definieren und zu messen ist.¹⁷⁵

Die Evaluation der Wirksamkeit von Täterprogrammen hängt daher davon ab, welche «Beweise» des Erfolgs die verschiedenen Methoden aus konzeptueller und pragmatischer Sicht erfassen können, das heisst welches die Indikatoren sind, die in begrenzter Zeit und mit beschränkten Budget auf angemessene Weise gemessen werden können.¹⁷⁶

4.3a Das englische Projekt MIRABAL, eine neue Ära von Evaluationsstudien?

Die Evaluation der Auswirkungen der Täterprogramme muss ausgehend von der Perspektive der Opfer vorgenommen werden.¹⁷⁷ Im Rahmen des englischen Projekts MIRABAL haben Kelly & Westmarland an der Ausarbeitung von 6 Erfolgsindikatoren gearbeitet, die quantitativ und qualitativ gemessen werden, um einen innovativen Ansatz im Bereich der Evaluationsstudien zur

Wirksamkeit von Täterprogrammen zu ermöglichen.¹⁷⁸

1. Eine respektvolle Kommunikation;
2. Ein grösserer Handlungsspielraum;
3. Sicherheit und Gewaltfreiheit für Partnerinnen und Kinder;
4. Eine sichere, positive und gemeinsame Elternschaft;
5. Eine bessere Wahrnehmung von sich selber und den anderen;
6. Sicherere, gesündere Kinder.¹⁷⁹

Angesichts des Bruchs, der dieses Projekt gegenüber der bisher traditionellen nordamerikanischen und europäischen Evaluationsforschung darstellt, ist es wichtig, an dieser Stelle diese verschiedenen Indikatoren genauer zu beschreiben und einige der Hauptergebnisse des Projekts MIRABAL zusammenzufassen. Die Ergebnisse basieren auf Bewertungen, die bei den Teilnehmern und den Opfern 12 Monate nach Beginn des Interventionsprogramms eingeholt wurden.

Beim Kriterium **respektvolle Kommunikation** zeigen die Analysen nach 12 Monaten für alle Indikatoren eine Verbesserung. Die quantitativen Indikatoren zeigen bei gewissen Punkten (z. B. verhandelt bei einem Streit) deutliche Verbesserungen und bei anderen Punkten (z. B. zeigt ein wohlbedachtes Verhalten) geringe Verbesserungen. Die qualitativen Indikatoren (z. B. hört sich ihre Meinung an; geht einer Unstimmigkeit nicht mehr aus dem Weg, sondern bringt sich ein; verhält sich deeskalierend; Aushandeln von gemeinsamen Interaktionsregeln) lassen zugänglichere Täter erkennen, die den Opfern Raum geben und überlegen, bevor sie sprechen. Die Forscherinnen und Forscher sehen darin eine Bewusstwerden der Täter der von ihnen ursprünglich ausgeübten Macht und

¹⁷⁵ Gondolf (2007) und Westmarland & Kelly (2013).

¹⁷⁶ Chung (2015); die europäischen Daphne-Programme zeigen diese Schwierigkeiten deutlich auf.

¹⁷⁷ Dies erfordert insbesondere die Anwendung einer Reihe von ethischen Leitplanken (Downes, Kelly, Westmarland, 2014; Wojnicka, Scambor, Kraus, 2016); zur Integration der Sichtweise der Kinder, siehe auch Alderson, Kelly & Westmarland (2013).

¹⁷⁸ Zur Entwicklung der Indikatoren, siehe Westmarland & Kelly (2013) und Kelly & Westmarland (2015b); für die Ergebnisse des englischen Projekts MIRABAL, siehe Kelly & Westmarland (2015a).

¹⁷⁹ Englisch: Respectful communication and improved relationship; (2) Expanded space for action; (3) Safety and freedom from violence and abuse for women and children; (4) Safe, positive, and shared parenting; (5) Enhanced awareness of self and others; et (6) Safer, healthier childhoods.

der Art, wie sie die Konversation kontrollierten und dominierten, sowie ihre bewusste Entscheidung, dies zu verändern.

Der Begriff **Handlungsspielraum** der Frauen verweist auf die Wichtigkeit, erneut ein Gefühl von ausreichender Sicherheit zu verspüren und dadurch gestärkt zu werden. Dieser Handlungsspielraum erhöht die Selbstbestimmung der Frauen, die wieder eine Stimme erhalten und die Möglichkeit, Entscheidungen zu treffen und ihr Wohlbefinden zu verbessern. Die Faktoren, die diesen Handlungsspielraum einschränken, gehen nach 12 Monaten in unterschiedlichem Mass zurück, wobei keiner vollständig verschwindet. Auf qualitativer Ebene lässt sich eine Verbesserung der Teilung der Haushalts- und Kinderbetreuungspflichten, der zwischenmenschlichen Beziehungen, aber auch eine erhöhte Bewegungsfreiheit und ein erfüllteres Gefühlsleben beobachten. Dabei ist zu beachten, dass die Frauen diese Veränderungen tendenziell ihrem eigenen Handeln zuschreiben. Sie seien nach wie vor vorsichtig und ängstlich, obwohl sie objektiv über eine grössere Freiheit verfügen. Diese Feststellung verweist auf die Notwendigkeit, nach einer andauernden Gewaltexposition neue Verhaltensweisen zu erlernen.

Die quantitativen Indikatoren verweisen auf eine deutliche und signifikante Abnahme der Gewalt im Rahmen des Kriteriums **Sicherheit und Gewaltfreiheit**. Dies trifft besonders auf die Indikatoren der körperlichen oder sexuellen Gewalt zu und weniger auf diejenigen, die missbräuchliche Verhaltensweisen reflektieren. Die qualitativen Indikatoren (z. B. fühlt sich sicher) gehen in die gleiche Richtung und legen keine Verlagerung hin zu anderen Formen von Kontrolle nahe. Veränderungen zeigen sich auch beim Verständnis und der Definition von Gewalt durch die Täter. Die Täter würden in ihren Berichterstattungen nach 12 Monaten ehrlicher erscheinen, seien fähiger, über ihr Verhalten zu reden und würden vermehrt Verantwortung übernehmen.

Die quantitativen Indikatoren im Zusammenhang mit einer **sicheren, positiven**

und gemeinsamen Elternschaft zeigen alle zumindest eine minimale Verbesserung. Die qualitativen Indikatoren sind schwieriger zu bewerten, da die Hälfte der Teilnehmer während des Programms keinen Kontakt zu ihren Kindern hatte.

Die Programtteilnehmer würden nach 12 Monaten auch eine bessere **Wahrnehmung von sich selber und den anderen** zeigen, da sie die Folgen ihres gewalttätigen Verhaltens für die Partnerin und ihre Kinder verstehen (akzeptieren) würden. Auch wenn Verbesserungen erkennbar seien, werde für einige der Weg bis zu einer konkreten Verantwortungsübernahme noch weit sein. Gewisse Täter hätten sich aber vor anderen Zuhörern als jenen der Gruppe zu ihrer Verantwortung bekannt, doch dies seien wenige.¹⁸⁰

Die Indikatoren, die sich auf die **gesünderen Kinder** beziehen, lassen Verbesserungen erkennen, insbesondere hätten die Kinder «weniger Angst um die Mutter» und «weniger Angst vor dem Täter». Ihre Validität ist jedoch beschränkt, da die Verbesserungen die Wahrnehmung der Situation der Mutter widerspiegeln und einige Väter seit Programmbeginn keinen Kontakt zu ihren Kindern hatten.¹⁸¹ Gemäss den Forscherinnen und Forschern bleibt noch viel zu tun, damit sich die Väter bewusst werden, welche Auswirkungen häusliche Gewalt auf die Kinder ungeachtet ihres Alters haben, und verstehen, dass sie ein angsterfülltes und spannungsgeladenes Klima schaffen.

Schliesslich veranschaulichen die von Kelly & Westmarland im Rahmen dieses innovativen Projekts zusammengetragenen quantitativen und qualitativen Daten sehr gut, dass die Kriterien «Programm beendet» und «keine neuen Gewalthandlungen» ungenügend sind. Dies trifft nicht nur auf das Wissen über die Schemata häuslicher Gewalt und deren Folgen

¹⁸⁰ Ein Kriterium, das Parallelen zum Transtheoretischen Modell aufweist, siehe S. 30 ff.

¹⁸¹ Dieses Kriterium ist besonders wichtig, wenn die Zuweisung zu einem Programm vom Sozialdienst oder von der Kinderschutzbehörde angeordnet wird.

zu, sondern auch auf die geleistete Arbeit im Rahmen eines Täterprogramms und auf die Erwartungen – und Hoffnungen – der Opfer.¹⁸²

¹⁸² Kelly & Westmarland (2015b).

5 Schlussfolgerungen

Diese Programme sind ein «unvollendetes Werk» und erfordern eine ständige Evaluation, damit ihre Möglichkeiten zur Förderung der Sicherheit von Frauen und Kindern ermittelt werden können.¹⁸³

Auf internationaler Ebene besteht nach wie vor Uneinigkeit über die Wirksamkeit der Täterprogramme. Diese Uneinigkeit ist teilweise auf das Unterlassen zurückzuführen, eine breite Palette potentieller Ergebnisse zu berücksichtigen, konzentrierte sich doch die Mehrheit der Studien auf ein übermässig enges Verständnis des Begriffs «Erfolg» (...). Der Erfolg muss neu definiert werden und sich nicht nur an der Strafjustiz, sondern auch am Gesundheits- und Sozialwesen orientieren.¹⁸⁴

Die Arbeit mit Tätern häuslicher Gewalt ist nicht sehr populär, weder in der Gemeinschaft, noch bei anderen Fachpersonen.¹⁸⁵ Trotz unternommener Anstrengungen bleiben einige überzeugt, dass die Lösung für häusliche Gewalt in den Gesetzen und nicht in den Programmen «zur Änderung» der Täter häuslicher Gewalt zu finden ist.

Von den verhafteten Tätern wird jedoch nur eine Minderheit strafrechtlich verfolgt, und nur ein paar wenige werden sanktioniert.¹⁸⁶ Die Androhung einer Freiheitsstrafe kann sich zwar sowohl auf kollektiver als auch auf individueller Ebene als starkes Symbol erweisen, aber langfristig kann sie keine Lösung sein. Wiederholungstaten häuslicher Gewalt werden nicht immer angezeigt, und viele Formen problematischen Verhaltens

stellen trotz ihrer schädlichen Wirkung auf die Gesundheit keinen Straftatbestand dar. Zudem ist die Inhaftierung eine kostspielige Lösung, die weder eine Verhaltensänderung noch eine Beendigung der Gewalt über das Ende der Haftstrafe hinaus bewirkt.¹⁸⁷ Die Täterprogramme sind daher eine notwendige, aber – zumindest heute – unzureichende Realität.

Die Evaluationsstudien gelangen zu gemischten Ergebnissen hinsichtlich der Wirksamkeit der Täterprogramme, Wiederholungstaten häuslicher Gewalt signifikant zu reduzieren. Mehrere der in der Schweiz, in Europa und anderswo in der Welt verfügbaren Studien weisen methodische Verzerrungen auf, welche die Validität ihrer Schlussfolgerungen einschränken. Wie bereits andere vor uns¹⁸⁸ gelangen wir unter Berücksichtigung der zurzeit verfügbaren Literatur zum Schluss, dass die Täterprogramme die Wiederholung häuslicher Gewalt in geringem Mass zu reduzieren vermögen. Dabei kann heute aufgrund der Variabilität der Programme und der Teilnehmer nicht festgestellt werden, *was* funktioniert, *wie* es funktioniert und *für welche Täterkategorien* es funktioniert.

Zwar scheint die körperliche Gewalt während der Teilnahme an einem Interventionsprogramm häufig abzunehmen, doch das gewalttätige Verhalten verschwindet nicht ganz. Die Programme begünstigen das Bewusstsein für das gewalttätige Verhalten und seine Auswirkungen auf Dritte. Die Partnerinnen fühlen sich sicherer und berichten oft über positive Veränderungen, sofern ihre Erwartungen realistisch sind. Dieser Befund ist der Grund dafür, warum die Dienstleistungen für die Opfer mit den Verantwortlichen der Täterprogramme

¹⁸³ Nach Chung (2015 : 182): «The programs are a 'work in progress' requiring continuing evaluation to identify their possibilities for promoting women's and children's safety.»

¹⁸⁴ Nach Westmarland & Kelly (2013: 1092): « [T]here remains ongoing disagreement internationally about [the] effectiveness [of perpetrator programs]. Part of this disagreement stems from the failure to consider a broad range of potential outcomes, with most research focusing on an overly narrow understanding of what 'success' means (...). [S]uccess needs to be redefined and connected not just to criminal justice, but also to health and social care agendas. »

¹⁸⁵ Gemäss den Fachgesprächen, die im Rahmen des Projekts MIRABAL geführt wurden, und gemäss unseren eigenen Studien.

¹⁸⁶ Hester & Westmarland (2006) und Hester (2006)

¹⁸⁷ Kelly & Westmarland (2015b).

¹⁸⁸ Insb. Babcock *et al.* (2004), Chung (2015), Eckhardt *et al.* (2013a, 2013b), Edleson (2012), Gondolf (2004, 2012) und Kelly & Westmarland (2015a).

koordiniert werden müssen oder zumindest regelmässige Kontakte stattfinden sollten.¹⁸⁹

Gestützt auf die derzeit verfügbaren Daten ist eine besondere Privilegierung eines Interventionsmodells zulasten eines anderen nicht gerechtfertigt, da sich heute die meisten Programme durch eine Form von Eklektizismus auszeichnen.¹⁹⁰ Aus dem gleichen Grund ist die Forschung noch nicht in der Lage, weder die spezifischen Mechanismen festzulegen, die den positiven Veränderungen bei den Tätern zugrunde liegen, noch die Strategien zu identifizieren, mit denen sich diese Mechanismen am besten begleiten lassen.

Die Meinungen gehen auch betreffend die Notwendigkeit auseinander, die Programme zu individualisieren oder sie an bestimmte Merkmale der Täter häuslicher Gewalt anzupassen. Bei dieser Frage scheinen die Verfechterinnen und Verfechter feministisch orientierter psychoedukativer Ansätze, wie Gondolf, und jene von psychodynamischen Ansätzen, wie Dutton, fast systematisch gegenläufiger Meinung zu sein. Ungeachtet der Positionierung scheint es wichtig, gründliche Überlegungen zur Relevanz der verschiedenen Individualisierungsparameter anzustellen. Zu diesem Punkt ist zu beachten, dass mehrere jüngere Studien alternative oder ergänzende Ansätze zu den sogenannten klassischen Programmen getestet haben, die insbesondere dem Transtheoretischen Modell und den Techniken der Motivierenden Gesprächsführung entnommen wurden. Gewisse Ergebnisse sind zwar vielversprechend, aber die verfügbaren Daten reichen noch nicht zur Festlegung aus, in welchem Masse solche Ansätze breiter eingesetzt werden könnten, insbesondere im Rahmen von gerichtlichen Anordnungen.

¹⁸⁹ Während das Modell in England klar abgegrenzt zu sein scheint, ist dies anderswo in Europa weniger der Fall. Den Opfern wird zwar am Rande der Täterprogramme Unterstützung angeboten, aber deren Modalitäten sind sehr unterschiedlich und reichen von telefonischer Unterstützung bis zu Diskussionsgruppen (Gloor, Meier, 2002; Kraus, 2013; Küken, Wypych, 2011). Manchmal sei man daher von einer echten Integration der Interventionen weit entfernt.

¹⁹⁰ Siehe oben, Titel 1.2.

Die präventive Wirkung der Täterprogramme scheint untrennbar mit der Unterstützung verbunden zu sein, die vom Justizapparat – und in einem grösseren Kontext vielleicht auch von der Gemeinschaft – bei der Koordination, der Zuweisung und der Orientierung der Täter häuslicher Gewalt und bei deren Sanktionierung im Fall von Nichtbeachtung der Regeln oder Nichterscheinen geleistet wird. Die Programme, die als Teil einer koordinierten Reaktion mehrerer Stellen entwickelt wurden, würden die besten Ergebnisse erzielen. Die Sicherheit und der Schutz der Partnerinnen würden teilweise ebenfalls von der Qualität der koordinierten Reaktion abhängen. Trotz dieser Verbesserungsansätze bleiben die hohen Abbruchquoten in den Täterprogrammen das Haupthindernis für eine wirksame Bekämpfung häuslicher Gewalt.

Welche Schlüsse sollen somit im Evaluationsbereich gezogen werden, ausser dass es weitere Qualitätsevaluationen zu den verschiedenen Programmtypen braucht? Denn die Evaluationsforschung im Bereich der Täterprogramme hat mit der Zunahme – und in einem gewissen Sinn mit der Diversifikation – der Interventionsmodelle nicht Schritt gehalten. Dieser Aufruf zu erhöhter methodischer Qualität gilt sowohl für das Studiendesign und den Kontext, in dem die Studien durchgeführt werden, als auch für die Indikatoren, die für die Bewertung des «Erfolgs» der verschiedenen Programme herangezogen werden. Diese Indikatoren müssen geeignet sein, komplexe dynamische Faktoren zu erfassen, die mit häuslicher Gewalt in Verbindung stehen und über das blosses Verschwinden der körperlichen Gewalt hinausgehen.

Mehrere Fachpersonen fordern eine vermehrte Durchführung von randomisierten experimentellen Studien, wobei dieser Wunsch nicht einvernehmlich ist. Eine Privilegierung des Grundsatzes der Randomisierung zulasten des Ermessensspielraums der Strafverfolgungs- und der Strafvollzugsbehörden wirft effektiv sowohl ethische als auch praktische Fragen auf. Andere Fachpersonen rufen daher dazu auf,

vermehrt strengere Evaluationen in einem weiteren Kontext, namentlich über Partnerschaften zwischen Forschung und Praxis, durchzuführen.

Unabhängig von ihrem Kontext müssen diese Evaluationen in der Lage sein, die Komplexität der massgeblichen Veränderungsmechanismen aufzuzeigen. Die jüngeren Evaluationen sind zwar ermutigend, es liegt aber noch ein weiter Weg vor uns.

6 Anhang

Dieser Anhang umfasst eine Zusammenfassung einiger der Texte, die für diese narrative Zusammenfassung verwendet wurden. Am Anfang werden Einzelstudien vorgestellt, die über eine gute methodische Qualität verfügen oder die angesichts der angestrebten Ziele von besonderer Bedeutung sind. Anschliessend folgen Informationen zu den Schlussfolgerungen einer Reihe von systematischen Übersichtsarbeiten und den wichtigsten Metaanalysen, die sich spezifisch mit der Wiederholung von häuslicher Gewalt nach einer Täterintervention befassen. Schliesslich wird auf einige Aufsätze und Arbeitsdokumente hingewiesen, die im Rahmen des europäischen Projekts IMPACT realisiert wurden. Innerhalb der einzelnen Gruppen sind die Studien nach Jahr, in aufsteigender Reihenfolge geordnet.

6.1 Einzelstudien¹⁹¹

Edleson & Syers (1990)

Randomisierte Studie zur relativen Wirksamkeit, ohne Kontrollgruppe, n = 283

Das Forscherteam hat 3 verschiedene Interventionsmodalitäten verglichen: (a) ein strukturiertes psychoedukatives Gruppenprogramm; (b) ein Selbsthilfeprogramm («self-help») für Gruppen; und (c) eine Intervention, welche diese beiden Modalitäten kombiniert. Es wurden immer auch 2 Interventionsdauern verglichen, nämlich 12 oder 32 Wochen. Freiwillige ($n = 181$) und von der Justiz zugewiesene Teilnehmer ($n = 102$) wurden nach dem Zufallsprinzip einer der 6 Experimentalgruppen zugeordnet.

Die anhand von Viktimisierungsdaten der Partnerinnen oder, wenn die (Ex-)Partnerinnen nicht erreichbar waren, anhand von selbstberichteten Daten vorgenommene

Feststellung der Wiederholung von körperlicher häuslicher Gewalt und von Drohungen 6 Monate nach Beendigung der Intervention, haben keine signifikanten Unterschiede weder zwischen der kurzen und der langen Programmdauer noch zwischen den 3 Interventionsmodalitäten ergeben. Das Forscherteam legte jedoch nahe, dass die kurzen und stark strukturierten Programme wirksamer seien. Diese Schlussfolgerung ist in der Literatur mehrmals anzutreffen.

The Ontario Experiment: Palmer, Brown & Barrera (1992)

Randomisierte Studie, mit Kontrollgruppe, n = 59

In dieser kanadischen Studie wurden 59 männliche Täter häuslicher Gewalt zufällig (Block-Randomisierung) einer der Versuchsbedingungen – Gruppenprogramm – oder einer Kontrollbedingung – keine Intervention – zugewiesen. Das wenig strukturierte Gruppenprogramm verfolgte einen psychoedukativen Ansatz, enthielt aber auch Elemente der kognitiven Verhaltenstherapie. Es dauerte 10 Wochen und war auf die Verantwortungsübernahme der Täter und die Entwicklung von Wut- und Konfliktmanagementkompetenzen ausgerichtet. 6 Monate nach Beendigung des Programms war die häusliche Gewalt gestützt auf Polizeiakten in der Kontrollgruppe signifikant höher (31 %) als in der Experimentalgruppe (10 %), was einer mittleren Effektgrösse entspricht ($d = 0.54$). Das Forscherteam wollte die Wiederholung der häuslichen Gewalt auch gestützt auf Viktimisierungsdaten messen, aber die Rücklaufquote war zu tief: Nur 22 % der Opfer haben 12 Monate nach Beendigung des Programms an der Datenerhebung teilgenommen.

Diese Studie ist zwar methodisch solid, wurde jedoch mit einer kleinen Stichprobe durchgeführt, was ihr viel Kritik einbrachte. Sie hat aber trotz der wenigen berücksichtigten Fälle das Verdienst, auf einer Population zu basieren, die besonders repräsentativ für alle

¹⁹¹ Für eine detaillierte Übersicht über einige der wichtigsten Studien, Metaanalysen und systematischen Übersichtsarbeiten, siehe z. B. Cluss & Bodea (2011).

Täter ist, die normalerweise von der Justiz zugewiesen werden.

Brannen & Rubin (1996)

Randomisierte Studie zur relativen Wirksamkeit, ohne Kontrollgruppe, n = 49 Paare

Die Forscherinnen und Forscher verglichen die Wirksamkeit einer Paarintervention mit jener einer feministisch orientierten psychoedukativen Gruppenintervention, die ausschliesslich für Männer bestimmt ist. 49 Paare wurden zufällig einer dieser Bedingungen zugewiesen. Ausgehend von selbstberichteten Daten gelangte das Forscherteam zum Schluss, dass die Paarintervention die körperliche Gewalt unmittelbar nach dem Programm wirksamer zu reduzieren vermochte als die Gruppenintervention. Doch 6 Monate nach Beendigung des Programms konnte kein signifikanter Unterschied mehr nachgewiesen werden, ausser bei den Teilnehmern mit einem Alkoholproblem, die im Paarprogramm bessere Ergebnisse zu erzielen scheinen.

Saunders (1996)

Randomisierte Studie zur relativen Wirksamkeit, ohne Kontrollgruppe, n = 218

Im Rahmen eines Programms wurden 218 Männer zufällig folgenden Interventionen zugeteilt: (a) einem feministisch orientierten kognitiv-verhaltens-therapeutischen Gruppenprogramm von 20 Wochen Dauer; oder (b) einem psychodynamischen Gruppenprogramm, das ebenfalls 20 Wochen dauerte. Drei Viertel der Teilnehmer waren von der Justiz zugewiesen worden (Strafvollzugs- oder Bewährungs-dienste). Der Forscher ging nicht davon aus, dass die Rückfallquoten zwischen den Teilnehmern der beiden Programme verschieden ausfallen würden, vielmehr rechnete er mit unterschiedlichen Ergebnissen je nach Persönlichkeitstyp der Männer.

Die Analyse der Viktimisierungsdaten hat 22 oder mehr Monate nach Beendigung des Programms für die Indikatoren erlittene Gewalt, Angst, allgemeine Veränderung des Verhaltens des Täters und Gleichstellung in der Paarbeziehung, keinen signifikanten Unterschied zwischen den beiden Programmen gezeigt. Insgesamt und gemäss den Viktimisierungsdaten wurden 30,9 % der Männer im kognitiv-verhaltenstherapeutischen Programm und 28,8 % im psychodynamischen Programm im Verlauf von 3 bis 54 Monate rückfällig. Im Einklang mit den Hypothesen des Forschers wiesen die Männer mit einer Suchtthematik im psychodynamischen Programm eine geringere Rückfälligkeit auf, während die Männer mit einer antisozialen Persönlichkeit im strukturierten kognitiv-therapeutischen Programm weniger häufig rückfällig wurden.

In einer methodisch gut aufgebauten Langzeitstudie, in der auch die Vollständigkeit der durchgeführten Interventionen und die Auswirkungen von Programmabbrüchen analysiert wurden, gelangte Saunders zum Schluss, dass der Ansatz «one size fits all» für Täter häuslicher Gewalt möglicherweise nicht der geeignetste Ansatz ist.

The San Diego Naval Experiment: Dunford (2000)

Randomisierte Studie, mit Kontrollgruppe, n = 861 verheiratete Paare

Insgesamt wurden 861 Paare aus der US-Navy zufällig einer der 4 folgenden Bedingungen zugeteilt: (a) einem kognitiv-verhaltens-therapeutischen Gruppenprogramm von 26 Wochen Dauer, gefolgt von 6 monatlichen Sitzungen; (b) einem Paarbehandlungs-programm, das während 26 Wochen im Multigruppen-Format, ebenfalls gefolgt von 6 monatlichen Sitzungen, stattfand; (c) einer engen behördlichen Überwachung; und (d) einer Kontrollgruppe, die aus der

Teilnahme des Opfers an einem Unterstützungsprogramm bestand.¹⁹²

Ausgehend von Viktimisierungsdaten, welche die Opfer 6 und 12 Monaten nach Beendigung der verschiedenen Interventionen bereitstellten, wurde kein signifikanter Unterschied zwischen den Gruppen festgestellt. Die ausgehend von Polizeidaten berechneten Rückfallquoten waren im Vergleich zu anderen Studien relativ niedrig: 12 Monate nach Beendigung des Programms waren 3–6 % der Täter über alle Gruppen hinweg erneut von der Polizei festgenommen worden. Die von den Opfern bereitgestellten Daten zeichnen ein anderes Bild: Über alle Studienbedingungen hinweg fühlten sich 12 Monate nach Beendigung des Programms 33–42 % der Opfer in Gefahr, 27–35 % waren gestossen oder geschlagen und 16–21 % körperlich verletzt worden. Zwischen den verschiedenen Interventionsmodalitäten wurde jedoch ungeachtet der verwendeten Daten kein signifikanter Unterschied festgestellt.

Trotz ihrer methodischen Strenge ist die ökologische Validität dieser Evaluation beschränkt, da sie bei einer Studienpopulation bestehend aus Männern im aktiven Militärdienst durchgeführt wurde. Dies bedeutet, dass die Männer keine oder nur eine unbedeutende kriminelle Vorgeschichte haben, einer Beschäftigung nachgehen, kein Suchtproblem haben und in einer strukturierten Gemeinschaft leben. Zudem besteht möglicherweise aufgrund der Null-Toleranz für häusliche Gewalt in der Armee eine starke formelle Kontrolle.

The Broward Experiment: Feder & Forde (2000) und Feder & Dugan (2002, 2004)

Randomisierte Studie, mit Kontrollgruppe, n = 404

In dieser Studie wurde die Wirksamkeit eines angeordneten Hilfsprogramms bei Tätern in Florida, die wegen häuslicher Gewalt verurteilt worden waren («misdemeanor»), untersucht.

404 Männer wurden für 5 Monate zufällig einer der folgenden Interventionen zugeteilt: (a) einer Versuchsbedingung – 12 Monate Bewährung, in Kombination mit einem angeordneten Hilfsprogramm gemäss Duluth-Modell von 26 Wochen; oder (b) einer Kontrollbedingung – nur 12 Monate Bewährung. Etwa ein Drittel der Männer nahmen nicht an der Intervention teil, der sie zugewiesen worden waren.

Daten zur Rückfälligkeit wurden bei den Tätern, den Opfern und in den amtlichen Dossiers erhoben (d. h. ausgeübte Gewalttaten, Verhaftungen, Verletzung der Bewährungsauflagen). Die Forscherinnen und Forscher haben die Daten unter Berücksichtigung der zugeteilten Intervention («intent to treat») und der tatsächlich erhaltenen Behandlung («dosage») analysiert: Sie stellten bei den Einstellungen, den Glaubenssätzen und den Verhaltensweisen hinsichtlich häuslicher Gewalt keinen signifikanten Unterschied zwischen den Gruppen fest. Die Viktimisierungsdaten ergaben auch keine Unterschiede: Bei beiden Gruppen wurden neue Gewalttaten gemeldet. Auch bei der Wiederverhaftung und der Verletzung der Bewährungsauflagen zeigte sich kein Unterschied: In jeder Gruppe waren während der Bewährungsfrist 24 % ein- oder mehrmals erneut verhaftet worden. Den Forscherinnen und Forscher zufolge war es den Männern, die sich um eine regelmässige Teilnahme am Programm füllten, auch gleichgültig, wenn sie rückfällig wurden. Mit anderen Worten, die Teilnehmer, die das Programm beendeten, würden einer Untergruppe von Tätern angehören, bei denen die Rückfall-wahrscheinlichkeit ungeachtet der Betreuung sowieso kleiner war.

Da die Teilnehmer, die der Kontrollbedingung zugeteilt worden waren, nicht berechtigt waren, an irgend einem Täterprogramm teilzunehmen, seien gemäss dem Forscherteam viele Personen (z. B. Sozialhelfer, Staatsanwälte und Bewährungsbeamte) der Ansicht gewesen, dass diese Studie die Opfer gefährden würde, und hätten ihre Ablehnung offen gezeigt, was

¹⁹² Siehe auch Titel 4.1h.

das Experiment beeinträchtigt hätte. Die Tatsache, dass nur ein Viertel an den Gesprächen teilgenommen habe, sei eine Konsequenz dieser ungünstigen Stimmung. Dies ändert jedoch nichts daran, dass die Rücklaufquote zu niedrig ist, um aussagekräftige Vergleiche zwischen der Experimental- und der Kontrollgruppe vorzunehmen.

O'Leary, Heyman & Neidig (1999)

Randomisierte Studie zur relativen Wirksamkeit, ohne Kontrollgruppe, n = 70 Paare

Diese Studie hatte zum Ziel, eine Paarbetreuung (Multigruppen-Paartherapie) mit einem feministisch orientierten Gruppenprogramm zu vergleichen. Beide Interventionen beruhten auf einem kognitiv-verhaltenstherapeutischen Ansatz. Im feministisch orientierten Programm wurde der Mann als verantwortlich für die ausgeübte Gewalt angesehen; die Multigruppen-Paartherapie ging davon aus, dass sowohl der Mann als auch die Frau für eine Verringerung ihrer Beziehungskonflikte verantwortlich waren. Teilgenommen haben Paare auf freiwilliger Basis. Die teilnehmenden Paare waren auf der Suche nach fachlicher Hilfe, ihre Beziehung war intakt, aber in allen Beziehungen hatten in der Vergangenheit mindestens 2 vom Mann ausgeübte Gewalttaten stattgefunden, bei denen die Frau keine Verletzungen davon getragen hatte und sich nicht medizinisch behandeln lassen musste.

40 Paare wurden der Paartherapie zugewiesen und 30 dem Gruppenprogramm; beide Programme dauerten 14 Wochen. Die teilnehmenden Paare mussten sich mit dem Grundsatz der Randomisierung (Randomisierung mit Warteliste) einverstanden erklären, und die Frauen, die separat befragt wurden, mussten bestätigen, dass sie sich beim Gedanken, zusammen mit ihrem Partner in einer Therapiegruppe teilzunehmen, wohl fühlten.

Es wurden viele selbstberichtete Daten zusammengetragen: vor und nach der Intervention sowie 12 Monate nach Beendigung des Programms. Sie umfassten die Häufigkeit psychischer und körperlicher Gewalt, Anzeichen von Dominanz und Isolation, Angstgefühle, Verantwortungsübertragung, Depression, dyadische Anpassung, Gewalt-ausübung und Angst nach den Sitzungen und Zufriedenheit der Teilnehmerinnen und Teilnehmer. Über alle Interventionen hinweg sind laut dem Forscherteam die schwerwiegenden körperlichen Gewalttaten um 51 %, die mittelschweren körperlichen Gewalttaten um 55 % und die psychische Gewalt um 47 % zurückgegangen. 12 Monate nach Beendigung des Programms war nur 1 von 4 Männern nicht erneut gewalttätig geworden, während 2 von 3 Männern keine schwerwiegende Gewalt ausgeübt hatten. Verbesserungen wurden bei der ehelichen Anpassung, den Depressionserkrankungen bei den Frauen und der Verantwortungs-übernahme der Täter festgestellt, auch hier ungeachtet der Intervention.

In Bezug auf die Sicherheit führten Paarkonflikte, die im Anschluss an die Programmsitzung stattfanden, bei beiden Interventionsarten in 2 % der Fälle zu körperlicher Gewalt. Ebenfalls ohne Unterschied zwischen den Gruppen gaben die Männer und die Frauen an, mit dem Programm zufrieden zu sein.

The Brooklyn Experiment: Taylor, Davis & Maxwell (2001, 2004)

Randomisierte Studie, mit Kontrollgruppe, n = 376

Insgesamt wurden 376 Männer, die für häusliche Gewalt verurteilt worden waren, zufällig folgenden Interventionen zugeteilt: (a) einem Gruppenprogramm nach Duluth-Modell von insgesamt 40 Stunden, das in einer Version von entweder 8 oder von 26 Wochen durchgeführt wurde; oder (b) einer angeordneten gemeinnützigen Arbeit (GA) von 40 Stunden, die innerhalb von 2 Wochen

geleistet werden musste. Die Staatsanwaltschaft, das Gericht und der Verurteilte mussten sich vorgängig damit einverstanden erklären, dass der Verurteilte ausgelost werden könnte. Dieses Setting kann die Stichprobe verzerrt haben, da ausschliesslich Täter aufgenommen wurden, die für die Teilnahme an einem Programm sehr motiviert waren.

Die Daten zur Rückfälligkeit wurden 6 und 12 Monate nach dem Entscheid aus vielfältigen Quellen zusammengetragen (d. h. Klagen, Verhaftungsberichte, Viktimisierungsdaten). Nach 6 Monaten zeigten die amtlichen Daten in der Experimentalgruppe eine statistisch niedrigere Rückfallquote als in der Kontrollgruppe (10 % gegenüber 21 %). Nach 12 Monaten war das Ergebnis gleich: 16 % gegenüber 26 %. Hingegen war keiner der Vergleiche, die auf den Viktimisierungsdaten beruhten, statistisch signifikant. Nach 6 Monaten gaben 22 % der Opfer der Männer der Experimentalgruppe an, erneut Gewalt erlitten zu haben, gegenüber 21 % bei der Kontrollgruppe; nach 12 Monaten betrug diese Zahl 15 % bei der Experimentalgruppe im Vergleich zu 22 % bei der Kontrollgruppe.

Anschliessend führten die Forscherinnen und Forscher eine Reihe von logistischen Regressionen durch, indem sie das Alter des Täters, seinen Zivilstand, seine ethnische Zugehörigkeit, seinen sozial-beruflichen Status und seine kriminelle Vorgeschichte prüften. Bei einer Analyse der Polizeidaten 12 Monate nach Programmende (8-wöchige Version) wurde eine Abnahme der Rückfälligkeit um 20,9 % im Vergleich zur Kontrollgruppe (GA) festgestellt; gestützt auf die Viktimisierungsdaten betrug die Abnahme 54,4 %. Bei der langen Programmversion (26 Wochen) wurde eine Abnahme der Rückfälligkeit um 43,4 % im Vergleich zur Kontrollgruppe (GA) festgestellt; gestützt auf die Viktimisierungsdaten betrug die Reduktion 79,3 %. Doch nur eine Reduktion war statistisch signifikant: Die Rückfallverminderung beim 26-wöchigen Programm, die anhand von amtlichen Daten

12 Monate nach Beendigung der Programms berechnet wurde. Gemäss dem Forscherteam sei dies auf die niedrige Teilnahmequote der Opfer an den Kontrollgesprächen 6 und 12 Monate nach Programmende zurückzuführen; bei 35 % der Opfer sei eine Kontaktaufnahme nicht möglich gewesen.

Trotz eines strikten experimentellen Designs konnten die Verzerrungen in den Evaluationsstudien aufgrund der Nichterreichung der Opfer nicht ausgeglichen werden. Obwohl das Autorenteam auf eine deutliche Rückfallverminderung hinwies, war nur einer der statistischen Vergleiche signifikant.¹⁹³ Mit dem langen Programm wurden zwar bessere Ergebnisse beobachtet, es kann jedoch nicht festgestellt werden, ob dies am Programm selber oder einfach an der Tatsache lag, dass diese Täter über einen längeren Zeitraum eng überwacht und deshalb weniger rückfällig wurden. Zudem hätten die zugelassenen Teilnehmer anscheinend kaum 3 % der 11 000 in diesem Zeitraum verurteilten Täter ausgemacht, was eine Anwendung dieses Modells auf breiterer Ebene stark einschränkt.

Diese Studie ist jedoch auch aus methodischer Sicht interessant, da sie eine Auseinandersetzung mit mehreren praktischen Schwierigkeiten möglich macht, die sich bei der Umsetzung eines «sauberen» Experiments in diesem Bereich stellen. In dieser Studie waren die Gerichte berechtigt, die einem Teilnehmer zufällig «zugeteilte» Sanktion zu ändern, wenn sie der Ansicht waren, dass diese mit einer Gefahr einhergehen könnte. So wurden 14 % der Teilnehmer, die zufällig der Kontrollgruppe (GA) zugeteilt worden waren, in das Gruppenprogramm umgeteilt; der umgekehrte Fall ist hingegen nie eingetreten. Das Forscherteam befasst sich schliesslich auch mit der Frage, wie mit Teilnehmern,

¹⁹³ Feder & Wilson (2005: 250-1) haben ausserdem durch den Einschluss dieser Studie in ihre Metaanalyse äusserst interessante Ergebnisse erzielt, die auf einige Unterschiede zwischen den Auswirkungen des kurzen und des langen Programms hinweisen; sie legen nahe, dass andere Elemente als nur das Programm Auswirkungen auf die Rückfälligkeit hätten.

welche die ihnen zugewiesene Intervention nicht beendet haben, umzugehen ist.

MacLeod, Pi, Smith & Rose-Goodwin (2010)

Programmevaluation, n = 1425

Diese Studie hat 53 zertifizierte Täterprogramme verglichen und dabei das Augenmerk auf die Auswirkungen der Tätermerkmale wie auch der Merkmale des Systems gerichtet, in das die verschiedenen Programme eingebettet sind. Diese psychoedukativen Programme, die auf die Entwicklung besonderer Kompetenzen ausgerichtet waren, dauerten alle 52 Wochen. Insgesamt wurden dem Forscherteam 1457 Dossiers übermittelt, 1425 davon enthielten vollständige Einzeldaten; 685 Männer hatten den Eintrittsfragebogen ausgefüllt, und die vor und nach der Intervention erhobenen Daten lagen für 233 Männer vor.

Die Forscherinnen und Forscher zogen einige Schlussfolgerungen aus ihrer Arbeit, so die Tatsache, dass Täter häuslicher Gewalt, die von der Justiz einem Programm zugewiesen werden, nicht repräsentativ für alle von häuslicher Gewalt betroffenen Personen seien. Die untersuchten Teilnehmer waren insgesamt weniger gebildet, ärmer, mehrheitlich hispanischer Herkunft und wiesen häufiger eine kriminelle Vorgeschichte auf. Nach Angaben der Teilnehmer wohnt ein Drittel immer noch mit dem Opfer und ebenfalls ein Drittel lebt mit Kindern zusammen. Das Forscherteam stellte einen gewissen Eklektizismus der Interventionen fest, die sich teilweise auf die kognitive Verhaltenstherapie, teilweise auf das Duluth-Modell und teilweise auf andere Ansätze beriefen. Die Vollständigkeit der Programme variierte je nach Rechtsordnung, doch dieser Umstand schien eher auf die Merkmale der verschiedenen Programmen zugeteilten Tätern als auf die Systemmerkmale zurückzuführen zu sein.

Unter den Teilnehmern, die an 52 Sitzungen teilgenommen hatten, stellte das Forscherteam gewisse positive Veränderungen im Bereich

ihrer Einstellungen und Glaubenssätze fest (z. B. Verantwortungsübernahme, Wutkontrolle, Empathie). Die besten statistischen Prädiktoren für eine erneute Verhaftung des Täters nach Eintritt in das Programm waren die persönlichen Merkmale des Täters und nicht jene des Programms oder der Rechtsordnung. Besser ausgebildete und ältere Männer mit einer «kürzeren» kriminellen Vorgeschichte und ohne Suchtproblematik wiesen die tiefste Rückfallwahrscheinlichkeit auf. Schliesslich machte es keinen Unterschied, ob das Gericht oder der Bewährungsdienst für die Überwachung zuständig war.

6.2 Systematische Übersichtsarbeiten

Smedslund, Dalsbø, Steiro, Winsvold & Clench-Aas (2011)

6 Studien, kognitive Verhaltenstherapie in Gruppen, experimentelles Design (Randomisierung und Kontrollgruppe)

Die 6 Studien, die von den Forscherinnen und Forschern ausgehend von sehr strengen Kriterien ausgewählt wurden, waren alle in den USA durchgeführt worden und umfassten insgesamt 2343 Teilnehmer.

Die erste Studie war eine Metaanalyse von 4 Studien, von denen jede eine Versuchs- (kognitive Verhaltenstherapie) und eine Kontrollbedingung (keine Behandlung) verglich. Insgesamt umfasste sie 1771 Teilnehmer. Das relative Rückfallrisiko betrug 0.86 zugunsten der Experimentalgruppe, Vertrauensintervall (CI) 95 % [0.54, 1.38]. Dies entspricht einer kleinen Effektgrösse, und angesichts des Vertrauensintervalls kann keine klare Aussage zur Wirksamkeit des Programms gemacht werden. Die fünfte Studie verglich ein kognitiv-verhaltenstherapeutisches Programm mit einem psychodynamischen Programm. Die Stichprobe umfasste 218 Männer. Das relative Rückfallrisiko betrug 1.07, 95 % CI [0.68, 1.68] zugunsten des psychodynamischen Programms, aber die Ergebnisse waren nicht eindeutig. In der

sechsten Studie wurde eine kombinierte kognitiv-verhaltenstherapeutische Behandlung, die auf häusliche Gewalt und Suchtprobleme ausgerichtet war, mit einer Selbsthilfegruppe, «Twelve Steps»,¹⁹⁴ verglichen. Die bei 58 Teilnehmern durchgeführten Analysen zeigten eine Effektgrösse von 0.30, 95 % CI [-0.22, 0.81] zugunsten des alkoholspezifischen Teils der kombinierten Behandlung.

Die Forscherinnen und Forscher kommen zum Schluss, dass nicht genügend randomisierte Studien vorliegen, um fundierte Aussagen zu einer eventuellen Wirksamkeit der kognitiven Verhaltenstherapie zur Reduktion der Wiederholung von häuslicher Gewalt zu machen.

Akoensi, Koehler, Lösel & Humphreys (2013)

11 Studien, wovon 7 einen Rückfallindikator gemessen haben, nur 1 mit einer Kontrollgruppe

Auf der Grundlage von vordefinierten Kriterien haben die Forscherinnen und Forscher 12 europäische Studien ausgewählt, welche die Wirksamkeit eines Täterprogramms evaluiert hatten. Diese aus verschiedenen Ländern – Zypern (1), Finnland (1), Deutschland (1), Spanien (4), Schweden (1) und dem Vereinigten Königreich (4) – stammenden Studien basierten auf unterschiedlich grossen Stichproben von 9–322 Teilnehmern. Alle evaluierten Programme beriefen sich auf einen gemischten Ansatz und kombinierten psychoedukative und kognitiv-verhaltenstherapeutische Ansätze mit feministischen Techniken. Kein Programm sah ausdrücklich eine individualisierte Betreuung vor, die sich nach den Merkmalen des Täters richtete. Alle Interventionen fanden in Gruppen statt und dauerten zwischen 3 und 30 Wochen, im Durchschnitt 20 Wochen.

Das Forscherteam konnte keine komplexen Analysen durchführen: Nur 7 der 12 Studien hatten einen Indikator für gewalttätiges

Verhalten gemessen, nur 1 Studie hatte eine Kontrollgruppe verwendet und alle wiesen methodische Verzerrungen auf. Die Abbruchquoten waren relativ hoch, logischerweise jedoch tiefer als bei den Studien, die mit «gefangenen» Studienpopulationen, wie Gefängnisinsassen, durchgeführt wurden. Nur wenige Studien hatten verschiedene Datenquellen für die Evaluation der Wiederholung häuslicher Gewalt verwendet. Die Beobachtungsdauern waren unterschiedlich: Die Hälfte der Studien hatte lediglich nach Beendigung des Programms Daten erhoben, und keine der übrigen beinhaltete eine Datenerhebung später als 12 Monate nach Betreuungsende.

Zwar legen mehrere Studien eine positive Wirkung der evaluierten Interventionen nahe, aber angesichts der Qualität des experimentellen Designs lässt sich die Wirksamkeit der Täterprogramme auf die Wiederholung häuslicher Gewalt kaum aussagekräftig beurteilen.

Eckhardt, Murphy, Whitaker, Sprunger, Dykstra & Woodard (2013a, 2013b)

30 Studien, psychoedukative oder kognitiv-verhaltenstherapeutische Programme oder alternative Ansätze, (randomisiertes) experimentelles oder quasi-experimentelles Design mit mindestens 2 Bedingungen

Hier hat das Forscherteam eine sehr detaillierte systematische Übersichtsarbeit zu sogenannten klassischen, wie feministisch orientierten psychoedukativen (Duluth-Modell) oder kognitiv-verhaltenstherapeutischen Programmen, und zu sogenannten alternativen Programmen durchgeführt. Es wurden auch Übersichtstabellen erstellt. Die Forschenden berücksichtigten nur randomisierte Studien oder quasi-experimentelle Studien, die eine aktive Experimentalgruppe (Programm) mit irgendeiner Kontrollgruppe verglichen hatten. Ausgewählt wurden 20 Studien, die klassische

¹⁹⁴ Ein ursprünglich in den USA entwickeltes Programm für die Anonymen Alkoholiker.

Programme evaluiert hatten, und 10 Studien, die alternative Ansätze bewertet hatten.¹⁹⁵

Das Forscherteam kommt zum Schluss, dass die Evaluationsstudien zu den Täterprogrammen hinsichtlich der Rückfallverminderung zu mehrdeutigen Ergebnissen gelangen und dass viele Studien methodische Verzerrungen aufweisen. Es weist darauf hin, dass die Evaluationen die Gruppenwirkung betreffen. Da die Entwicklungen der verschiedenen Teilnehmer unterschiedlich sind, besteht kein Zweifel daran, dass es Erfolgsgeschichten gibt. Zum Abschluss verweist das Autorenteam auf die ermutigenden Ergebnisse gewisser alternativer Ansätze in jüngeren und häufig methodisch strikteren Studien.

NICE (2014)

Die 34 für diese Arbeit ausgewählten Evaluationsstudien befassten sich zum einen mit der Wiederholung häuslicher Gewalt («outcome evaluation»), die mithilfe von amtlichen oder selbstberichteten Daten oder Viktimisierungsdaten ermittelt wurde, und zum anderen mit der Umsetzung dieser Programme und der jeweiligen Prozesse («process evaluation»).

Gestützt auf diese Zusammenfassung hat das Forscherteam eine Liste von 17 Empfehlungen für die Betreuung von Tätern, Opfern und Kindern erstellt, die von der Planung der Dienstleistungen und der Beteiligung der Institutionen und lokalen Organisationen, über die Schaffung eines günstigen Umfelds für die Offenlegung häuslicher Gewalt bis zu Relevanz der individualisierten Interventionen und der Ausbildung der Programmmitarbeitenden reicht.

Projet IMPACT, Daphne III (Hester, Lilley, O'Prey, Budde, 2014)

Im Rahmen des Projekts IMPACT¹⁹⁶ haben die Forscherinnen und Forscher eine narrative

Zusammenfassung vorgelegt. Unter Anwendung von weniger strikten Kriterien haben sie 65 Evaluationsstudien aus 13 Ländern aus ganz Europa ausgewählt: aus dem Norden (Dänemark, Finnland, Island und Schweden, 8), dem Süden (Spanien und Portugal, 23), dem Westen (Irland, Vereinigtes Königreich, 20), dem Osten (Kroatien, 1) und aus Zentraleuropa (Deutschland, Österreich, Niederlande und der Schweiz, 13).

Die Entscheidung, die Kriterien weit zu fassen, war teilweise auch durch das Interesse des Forscherteams bedingt, nicht nur die auf die Wiederholung häuslicher Gewalt ausgerichteten Evaluationen zu untersuchen, sondern auch die Studien, die sich mit der Implementierung und der Umsetzung der Programme und deren Prozesse befassten. Eingeschlossen wurden experimentelle, quasi-experimentelle und nicht-experimentelle Studien.

Die Ergebnisse sind sehr detailliert und beinhalten viele Tabellen. Hier sei nur auf einige Ergebnisse bezüglich der Wirkung der Programme hingewiesen. Ein Drittel der ausgewählten Studien hatten Beobachtungsdaten erhoben; 5 Studien nach 6 Monaten oder weniger, 9 zwischen 6 und 12 Monaten und 11 nach 12 Monaten. Die Hälfte der Studien (57 %) ergaben positive Veränderungen, 13 zeigten statistisch signifikante Ergebnisse und 37 vielversprechende, aber nicht signifikante Ergebnisse. Diese Schlussfolgerungen basierten jedoch auf sehr unterschiedlichen Indikatoren, sowohl für die Verhaltens- und Einstellungsveränderungen sowie für die psychologischen Veränderungen als auch für die Ermittlung der Motivation und der Zufriedenheit.

6.3 Metaanalysen

Davis & Taylor (1999)

¹⁹⁵ Die Autorinnen und Autoren haben in dieser Arbeit auch die Interventionsprogramme für die Opfer untersucht.

¹⁹⁶ Siehe auch die Texte am Ende des Anhangs.

Metaanalyse, 5 quasi-experimentelle (gepaarte, nicht äquivalente Gruppen) und 2 experimentelle Studien

Beim Versuch zu bestimmen, ob die Teilnahme an einem Täterprogramm rückfallvermindernd wirkt, stellten die Forscherinnen und Forscher fest, dass die Zahl der Studien mit einem Design, das die Beantwortung dieser Frage erlaubt, beschränkt ist. Sie weisen jedoch darauf hin, dass die Studien, welche die Auswahlkriterien erfüllen, mehr oder weniger alle zum Schluss kommen, dass die Täterprogramme funktionieren und die beobachteten Wirkungen substantiell sind. Es wurde nicht eine Metaanalyse im eigentlichen Sinn durchgeführt, sondern der Durchschnitt der Effektgrößen der 5 quasi-experimentellen Studien berechnet. Das Forscherteam schätzt, dass die Teilnahme an einem Täterprogramm einer Effektgröße von 0.41 entspricht, was klein ist. Aus den vorliegenden Ergebnissen könne jedoch nicht auf die grössere Wirksamkeit des einen oder anderen Programms geschlossen werden. Die Frage, ob bestimmte Täter bessere Ergebnisse erzielen als andere oder ob sie mit bestimmten Programmtypen bessere Ergebnisse erzielen, lässt sich zum jetzigen Zeitpunkt nicht beantworten.

In einem Vergleich der experimentellen mit den quasi-experimentellen Studien widerlegt das Forscherteam die Hypothese, dass die Wirksamkeit des Programms mit zunehmender methodischer Qualität der Evaluation verschwinden würde. Es widerlegt auch die Hypothese, dass die Programme nicht rückfallvermindernd wirken würden, räumt aber ein, dass es noch nicht viele Studien gibt und dass sowohl die Zahl der experimentellen und quasi-experimentellen als auch ihre methodische Qualität erhöht werden müssen.

Babcock, Green & Robie (2004)

Metaanalyse, 5 experimentelle Studien und 17 quasi-experimentelle Studien, Vergleich der Behandlungsgruppe mit (a) Abbrüchen; (b) Nicht-Erscheinen oder Ausschluss, oder (c) gepaarte Gruppe, ohne Intervention

Die Forscherinnen und Forscher haben eine Metaanalyse von 22 Studien durchgeführt, welche die Wirksamkeit von Täterprogrammen zur Reduktion der Wiederholung häuslicher Gewalt evaluiert hatten, die auf feministisch orientierten psychoedukativen (Duluth-Modell) oder kognitiv-verhaltens-therapeutischen Interventionen oder anderen Ansätzen basierten. In den Analysen wurde dem Designtyp und den Behandlungsmodalitäten Rechnung getragen. Es wurde kein signifikanter Unterschied zwischen den beiden Ansätzen festgestellt. Die Evaluationsstudien der feministisch orientierten psychoedukativen Programme ergaben eine kleine Effektgröße $d = 0.35$ (95 % CI: 0.15 bis 0.55), im Vergleich zu $d = 0.29$ (95 % CI: 0.01 bis 0.60) für die kognitiv-verhaltenstherapeutischen Interventionen. Die Wirkung der Behandlung war insgesamt gering. Mit anderen Worten, die Teilnahme an den Interventionsprogrammen würde nur geringfügig zu einer Rückfallverminderung beitragen.

Feder & Wilson (2005) und Feder, Wilson & Austin (2008)

Metaanalyse, 4 experimentelle Studien, 6 quasi-experimentelle Studien, nur mit einer gepaarten Gruppe oder mit statistischen Kontrollen

In dieser Metaanalyse befassten sich die Forscherinnen und Forscher mit den von der Justiz zugewiesenen Teilnehmern. Sie haben die Angaben zu den Rückfällen in den amtlichen Daten und die der Opfer separat analysiert. Berücksichtigt wurden die Studien von Programme, die von der Justiz entweder in Form einer Sanktion oder eines Diversionsprogramms (vor dem Prozess) angeordnet worden waren und die mindestens 6 Monate nach Beendigung der Intervention Daten erfasst hatten.

Das Forscherteam hat eine marginale Wirkung der Programme auf die Rückfälligkeit festgestellt und darauf hingewiesen, dass die Unterschiede in den mit einer allgemeinen Studienpopulation durchgeführten Studien kleiner seien und dass sie verschwinden

würden, wenn die Rückfälligkeit anhand von Viktimisierungsdaten gemessen würde. Das Team kam zum Schluss, dass nicht mit Sicherheit gesagt werden könne, dass die damals verwendeten Programme wirksam seien, und hat vorgeschlagen, dass die Justiz andere Interventionsmodelle einsetzen solle.

In Bezug auf die amtlichen Daten ergaben die Berechnungen des Forscherteams für die experimentellen Studien eine bescheidene Rückfallverminderung, $d = 0.26$, 95 % CI [0.03, 0.50], was einem Rückgang von rund 13 bis 20 % des Rückfallrisikos entspricht. Die Analysen mit den Viktimisierungsdaten waren nicht signifikant. Unter den quasi-experimentellen Studien gab es zum einen Studien, welche die freiwilligen und die von der Justiz zugewiesenen Teilnehmer verglichen, und zum anderen solche, welche die Teilnehmer, die ihre Betreuung beendeten, mit Personen verglichen, die nicht ausgewählt wurden, nicht erschienen sind oder abgebrochen haben. Die Differenz zwischen den freiwilligen und den zugewiesenen Teilnehmern war nicht signifikant, $d = -0.14$, 95 % CI [-0.44, 0.31]. Hingegen war der Unterschied zwischen den Personen, welche die Betreuung beendeten, und den anderen signifikant, $d = 0.97$, 95 % CI [0.12, 1.82]. Aufgrund des Designs können diese Unterschiede jedoch nicht nur der Behandlung zugeschrieben werden. Wahrscheinlich spielten auch andere Variablen eine Rolle, namentlich bereits bestehende Unterschiede zwischen den Gruppen.

Ausgehend von den Viktimisierungsdaten konnte weder für die experimentellen Studien noch für die quasi-experimentellen Studien eine Wirkung der Behandlung nachgewiesen werden. Die mittlere Effektgrösse für die experimentellen Studien betrug nahezu null und war statistisch nicht signifikant, $d = 0.01$, 95 % CI [-0.11, 0.13]; bei den quasi-experimentellen Studien legte sie gar eine negative Wirkung der Behandlung nahe, aber dieses Ergebnis war ebenfalls nicht signifikant, $d = -0.11$, 95 % CI [-0.12, 0.11].

Das Forscherteam bezeichnet die Differenz zwischen den amtlichen Daten und den Viktimisierungsdaten als «Schwelleneffekt»: Die Teilnahme würde zwar die Wahrscheinlichkeit von schweren Gewalttaten senken, die anderen Formen von Gewalt würden jedoch nicht verschwinden. Dieser Hypothese kommt eine gewisse Gültigkeit zu, da mehrere dieser Studien die ausgeübte Gewalt mithilfe der «Conflict Tactics Scales» (CTS-Methode) gemessen haben. Dabei handelt es sich um ein Instrument, mit dem das Vorliegen mehrerer körperlicher Gewalttaten von geringerer Schwere gemessen wird.

Im Vergleich zu Babcock *et al.* (2004) haben Feder *et al.* (2008) für die Auswahl ihrer Studien konservativere Kriterien angewandt, insbesondere haben sie die quasi-experimentellen Studien nach der Art ihrer Kontrollgruppe unterschieden und diese separat analysiert. Werden nur die experimentellen Studien betrachtet, so sind die Schlussfolgerungen von Babcock *et al.* (2004) und Feder *et al.* (2008) ähnlich. Auf der Grundlage von amtlichen Daten schliessen Babcock *et al.* (2004) auf eine Effektgrösse von $d = 0.12$, 95 % CI [0.02, 0.22], gegenüber $d = 0.26$ (95 % CI 0.03-0.50) bei Feder *et al.* (2008). In beiden Metaanalysen ist die Wirkung schwach, aber signifikant. Ausgehend von den Viktimisierungsdaten beträgt die Effektgrösse gemäss Babcock *et al.* (2004) $d = 0.09$, 95 % CI [-0.02, 0.21], gegenüber $d = 0.01$, 95 % CI [-0.11, 0.13] bei Feder *et al.* (2008); das entspricht in beiden Metaanalysen einer nicht signifikanten Wirkung.

Miller, Drake & Nafziger (2013)

Metaanalysen, 9 experimentelle oder quasi-experimentelle Studien mit statistischen Kontrollen

Die Forscherinnen und Forscher haben 34 amerikanische und kanadische Studien bestimmt, welche die Wirksamkeit von Täterprogrammen auf das Rückfallverhalten von Tätern häuslicher Gewalt untersucht haben, die von der Strafjustiz einer Betreuung

zugewiesen worden waren. Mit Ausnahme von 4 Studien, die sich mit der Wirksamkeit von Paarinterventionen befassten, betrafen alle Studien geschlechtergetrennte Programme. Die Forschenden berücksichtigten ausschliesslich Studien, die eine Form von Kontrollbedingung und Informationen beinhalteten, die eine Berechnung der Effektgrössen für alle zugeteilten Teilnehmer («intent to treat») erlaubten, sowie solche, in denen die Abbruchquote weniger als 10 % betrug. Sie haben zudem Studien bevorzugt, deren Angaben zu den Rückfallquoten auf amtlichen Daten beruhten.

Nach einer Analyse wählten sie 9 Studien aus, die sie als genügend strikt einstufen, was einem Total von 11 Effektgrössen entspricht. Die mittlere Effektgrösse für alle Studien war nicht signifikant, das heisst, die angeordnete Hilfe hatte keine Wirkung auf die Wiederholung häuslicher Gewalt. Nach einem individuellen Vergleich der Ansätze nach dem Duluth-Modell (5) mit anderen eklektischeren Modellen (4) gelangte das Forscherteam zum Schluss, dass erstere keine Wirkung auf die Rückfälligkeit hatten, während letztere alle rückfallvermindernd waren, wenn auch nicht signifikant. Eine gemeinsame Analyse der 4 alternativen Modelle zeigte mit einer mittleren Rückfallverminderung von 33 % eine signifikante Wirkung auf die Reduktion von häuslicher Gewalt. Dabei ist jedoch zu beachten, dass es sich um Interventionen handelt, die nicht unbedingt allen Tätern angeboten werden können. Genauer gesagt handelte es sich um eine kognitiv-verhaltenstherapeutische Gruppe, eine Intervention zur Verbesserung der Paarbeziehung, eine Suchtbehandlung und ein Multigruppen-Paartherapie-Programm.

Arias, Arce & Vilariño (2013)

Metaanalyse, 6 experimentelle Studien, 13 quasi-experimentelle Studien, mit einer Beobachtungsdauer von mindestens 6 Monaten

Das Forscherteam berücksichtigte für die Berechnung der Rückfallquoten amtliche Daten (Polizei) und Viktimisierungsdaten und

gab an, dass die mit letzteren Daten erzielten Ergebnisse immer höher waren. Bei den quasi-experimentellen Studien betrugen die ausgehend von den amtlichen Daten berechneten Rückfallquoten je nach Studie zwischen 4,6 % und 31,8 % bei einer Beobachtungsdauer von 7 bis 10 Monaten, und zwischen 33 % und 34 % bei einer Berechnung gestützt auf Viktimisierungsdaten. Bei den experimentellen Studien betrugen die ausgehend von den amtlichen Daten berechneten Rückfallquoten nach 12 Monaten je nach Studie zwischen 10 % und 69,6 %, und zwischen 0 % und 35 % bei einer Berechnung gestützt auf Viktimisierungsdaten.

Das Forscherteam hat in seinen Analysen auch die verschiedenen statistischen Moderatorvariablen berücksichtigt (z. B. Beobachtungsdauer, Interventionsart). Insgesamt kommt die Metaanalyse zum Schluss, dass die Behandlung von Tätern häuslicher Gewalt nicht wirksam ist, auch wenn gewisse Programme eine Rückfallverminderung nahe legen. Auf insgesamt 46 Effektgrössen, was einer Stichprobe von 18 941 Teilnehmern entspricht, betrug das mittlere gewichtete δ 0.41, 90 % CI [-0.12, 0.94], was einer positiven, nicht signifikanten Wirkung der Behandlung entspricht. Die Forscherinnen und Forscher haben auch eine negative Wirkung der Behandlung und eine Erhöhung der Rückfallwahrscheinlichkeit um 6 % beobachtet. Die Wirkung der Behandlung sei jedoch nicht gleich null; eine Null-Wirkung sei gleich wahrscheinlich wie ein Wirkungsgrad von 38 %. Die Wirksamkeit der Interventionen hänge weder von der Beobachtungsdauer noch von der Interventionsart ab. Daher erachtet das Forscherteam die derzeit vorliegenden wissenschaftlichen Daten nicht als schlüssig.

6.4 Aufsätze und Arbeitsdokumente

Zusätzlich zu den Leitlinien (IMPACT Toolkit), die eine Harmonisierung der Prozesse zur Evaluation der Täterprogramme bezwecken, wurden im Rahmen des europäischen Projekts

IMPACT mehrere sehr interessante Aufsätze und Arbeitsdokumente verfasst.¹⁹⁷

IMPACT Toolkit

Der im Rahmen des Projekts IMPACT¹⁹⁸ entwickelte «Werkzeugkasten» hatte folgende Ziele: (1) Bereitstellung von Instrumenten, mit denen die Anbieter und Vermittler von Täterprogrammen deren Auswirkungen sowohl auf die Täter als auch auf die Opfer und Kinder analysieren können; (2) Bereitstellung einer Struktur für Programme, die einer Selbst- oder Fremdevaluation unterzogen werden sollen; und (3) Förderung der Erhebung von vergleichbaren Evaluationsdaten für die Forscherinnen und Forscher.¹⁹⁹

Im Aufsatz «Implementing the IMPACT Toolkit», der Erste einer Reihe zur Beobachtung der Verbreitung und Entwicklung des Toolkits, vermittelt Jones (2015) verschiedene praktische Ratschläge, wie mit den vom Projektteam bereitgestellten Papier- und Internetinstrumenten gearbeitet werden kann. Sie demonstriert, wie mit dem Fragebogen Projekte evaluiert werden können, und geht ausführlich auf die vor und nach der Intervention und während der Beobachtungsdauer zu stellenden Fragen ein.

Working towards Safety: Supporting women alongside DVIP's perpetrator programme

Rajagopalan, Price, Langston & Potter (2015) erörtern die Beziehungen, welche die Dienste für Täter häuslicher Gewalt in England zu den Partnerinnen und Ex-Partnerinnen ihrer Klienten pflegen. Sie erläutern, wie diese proaktive Unterstützung der Opfer zur deren

Empowerment beitragen kann, und gleichzeitig die Verantwortungsübernahme der Täter unterstützt. Diese Dienstleistungen seien besonders geeignet für Opfer, welche die Paarbeziehung – zumindest vorerst – nicht beenden möchten, sowie für Opfer, die Kinder mit dem Täter haben und einen regelmässigen Kontakt zwischen dem Täter und seinen Kindern sicherstellen möchten oder müssen. Bei diesen Dienstleistungen für die Opfer geht es um mehr als Informationsvermittlung, Risikoevaluation und Validierung der Aussagen der Täter über ihr gewalttätiges Verhalten. «Es handelt sich um ein frauenzentriertes Interventionsangebot, das die Gewalt- und Missbrauchsoffer darin bestärkt, sicherere Entscheidungen für sich selber zu treffen und damit zu beginnen, ihre Lebenserfahrungen und das Missbrauchsverhalten durch ihn zu entflechten.»²⁰⁰

Partnerships with Victims' Services in Work with Perpetrators

Logar (2015) setzt sich mit einem sogenannten holistischen Ansatz der Täterarbeit auseinander. Sie befasst sich insbesondere mit den Schwierigkeiten, in bestimmten Kontexten Partnerschaften zwischen – feministischen – Frauenorganisationen, die spezialisierte Dienstleistungen für Gewaltopfer und deren Kinder anbieten, und den Männergruppen und den Diensten, die Täterprogramme durchführen, aufzubauen. Sie geht auf die sozialen Geschlechterungleichheiten, welche die Täterarbeit auszeichnen, aber auch auf die Partnerschaften zwischen verschiedenen Institutionen und Verbänden ein. Anhand des österreichischen Modells werden konkrete Anhaltspunkte aufgezeigt.²⁰¹

¹⁹⁷ Die hier ausführlich beschriebenen Dokumente sowie alle Publikationen im Zusammenhang mit den europäischen Projekten WWP und IMPACT sind unter folgendem Link abrufbar: <http://www.work-with-perpetrators.eu>.

¹⁹⁸ IMPACT Toolkit guidelines (2014); Toolkit from the Daphne III project 'IMPACT: Evaluation of European perpetrator programmes' (2014).

¹⁹⁹ Im Rahmen der Vorbereitung des Toolkits hatte das IMPACT-Projektteam auch eine kritische Analyse der verschiedenen Indikatoren publiziert, die für die Täterprogramme in Europa verwendet werden, siehe Geldschläger, Ginés, Nax & Ponce (2014).

²⁰⁰ Nach Rajagopalan et al. (2015: 2): «It is about a woman-focussed intervention service supporting a victim of violence and abuse as she is empowered to make safer choices for herself, and begins the process of disentangling her life experiences from his abuse of her.»

²⁰¹ Die Frage von institutionsübergreifenden Partnerschaften wird in Logar & Vargová (2015) ebenfalls ausführlich behandelt, insbesondere unter dem Aspekt der Ausbildung der Programmmitarbeitenden.

Responding to diverse ethnic communities in domestic violence perpetrator programmes

Debbonaire (2015) stellt interessante Überlegungen zum Konzept der ethnischen Zugehörigkeit an, die im Bereich der Täterarbeit einzigartig sind. Sie fragt nach der Relevanz dieses Konzepts und weist darauf hin, dass die Entwicklung von Kenntnissen zu den kulturellen Überzeugungen und Haltungen von ethnisch unterschiedlichen Gruppen in der Arbeit vor Ort von Nutzen sind. Sie schlägt einen interkulturellen Ansatz vor, der bei einer Reihe von Personen, die verschiedenen ethnischen Gemeinschaften angehören, angewandt wird. Weiter weist sie darauf hin, dass den spezifischen Gegebenheiten, wie Sprache oder Kulturschranken, Rechnung zu tragen sei, um zu verstehen, wie die Kultur, die selber multidimensional ist, auf vielfältige Weise zahlreiche Männlichkeitskulturen einschliesst.²⁰²

²⁰² Zur Männlichkeit siehe auch Wojnicka (2015).

7 Quellenverzeichnis

Aeberhard N, Frei P, *Alkohol und Häusliche Gewalt: Handbuch für die Beratungspraxis*, Bern, Schweiz: Blaues Kreuz Verlag, 2015.

Akoensi TD, Koehler JA, Lösel F, Humphreys DK, Domestic violence perpetrator programs in europe, part II: A systematic review of the state of evidence, *International Journal of Offender Therapy and Comparative Criminology*, 57: 1206-25, 2013.

Alderson S, Kelly L, Westmarland N, *Domestic violence perpetrator programmes and children and young people*, London and Durham, England: London Metropolitan University and Durham University, 2013.

Alexander PC, Morris E, Stages of change in batterers and their response to treatment, *Violence and Victims*, 23: 476-92, 2008.

Almeida RV, Durkin T, The cultural context model: Therapy for couples with domestic violence, *Journal of Marital and Family Therapy*, 25 : 313-24, 1999.

Andrews DA, Bonta J, *The psychology of criminal conduct*, New Providence, NJ: Matthew Bender & Co., 5th ed., 2010 [1994].

Arce R, Fariña F, Diseño e implementación del programa galicia de reeducación de maltratadores: Una respuesta psicosocial a una necesidad social y penitenciaria, *Intervención Psicosocial*, 19: 153-66, 2010.

Arce R, Fariña F, Intervención con penados en libertad por violencia de género: El "programa galicia de reeducación de maltratadores de género". in F Fariña, R Arce, G Bucla-Casal, eds., *Violencia de género: Tratado psicológico y legal*, Madrid, Spain: Biblioteca Nueva, 235-49, 2009.

Arias E, Arce R, Vilariño M, Batterer intervention programmes: A meta-analytic review of effectiveness, *Psychosocial Intervention*, 22: 153-60, 2013.

Arrigoni F, Jiménez JL, Navarro JI, Mendoza P, Aplicación de un programa terapéutico en hombres violentos contra la pareja, *Anuario de Psicología Jurídica*, 23: 3-9, 2013.

Babcock JC, Green CE, Robie C, Does batterers' treatment work? A meta-analytic review of domestic violence treatment, *Clinical Psychology Review*, 23: 1023-53, 2004.

Bächli-Biétry J, *Evaluationsbericht des modellversuchs, lernprogramme als neue interventionsform in der strafjustiz* [Online], 2006, <https://www.bj.admin.ch/bj/de/home/sicherheit/smv/modellversuche/evaluationsberichte.html>.

Baobaid M, Kovacs N, MacDiarmid L, Tremblay E, A culturally integrative model of domestic violence response for immigrant and newcomer families of collectivist backgrounds, in H Johnson, BS Fisher, V Jaquier, eds., *Critical issues on violence against women: International perspectives and promising strategies*, London: Routledge, 154-65, 2015.

Barz M, Helfferich C, Kavemann B, Aksu-Yagci S, Breunig A, Sommer B, *Häusliche Gewalt beenden: Verhaltensänderung von Tätern als Ansatzpunkt. Eine Evaluationsstudie zum Vorgehen und Wirkung von Täterprogrammen im Kontext von Interventionsprojekten gegen häusliche Gewalt in Baden-Württemberg*, Stuttgart: Landesstiftung Baden-Württemberg gGmbH, 2006.

Begun AL, Shelley G, Strodthoff T, Short L, Adopting a stages of change approach for individuals who are violent with their intimate partners, *Journal of Aggression, Maltreatment & Trauma*, 5: 105-27, 2001.

Bennett L, Hsieh C-M, Stoops C, Underclass men in batterer intervention programs: Disorders and disparities, *Families in Society: The Journal of Contemporary Social Services*, 91: 394-400, 2010.

Bograd M, Family systems approaches to wife battering: A feminist critique, *American Journal of Orthopsychiatry*, 54: 558-68, 1984.

Boira Sarto S, Jodrá Esteban P, Psicopatología, características de la violencia y abandonos en programas para hombres violentos con la pareja: Resultados en un dispositivo de intervención, *Psicothema*, 22: 593-9, 2010.

Boira Sarto S, López Y, Tomás-Aragónés L, Gaspar AR, Efficacy of different treatment modalities in men convicted of intimate partner violence, *Anales de Psicología / Annals of Psychology*, 29: 19-28, 2013.

Braaf R, *Elephant in the room: Responding to alcohol misuse and domestic violence*, Sydney: University of New South Wales, Australian Domestic and Family Violence Clearinghouse, 2012.

Brannen SJ, Rubin A, Comparing the effectiveness of gender-specific and couples groups in a court-mandated spouse abuse treatment program, *Research on Social Work Practice*, 6: 405-24, 1996.

Campbell M, Neil JA, Jaffe PG, Kelly T, Engaging abusive men in seeking community intervention: A critical research & practice priority, *Journal of Family Violence*, 25: 413-22, 2010.

Capaldi DM, Kim HK, Typological approaches to violence in couples: A critique and alternative conceptual approach, *Clinical Psychology Review*, 27: 253-65, 2007.

Carbajosa P, Boira Sarto S, Estado actual y retos futuros de los programas para hombres condenados por violencia de género en España, *Psychosocial Intervention*, 22: 145-52, 2013.

Casey S, Day A, Howells K, The application of the transtheoretical model to offender populations: Some critical issues, *Legal and Criminological Psychology*, 10: 157-71, 2005.

Cavanaugh MM, Gelles RJ, The utility of male domestic violence offender typologies: New directions for research, policy, and practice, *Journal of Interpersonal Violence*, 20: 155-66, 2005.

Chung D, Behavior change programs for intimate partner violence abusers: A means to promote the safety of women and children?, in H Johnson, BS Fisher, V Jaquier, eds., *Critical issues on violence against women: International perspectives and promising strategies*, London: Routledge, 171-82, 2015.

Cluss PA, Bodea A, *The effectiveness of batterer intervention programs*, Pittsburgh, PA: University of Pittsburg, 2011.

Coalition for Evidence-Based Policy Working Group, *Which study designs can produce rigorous evidence of program effectiveness? A brief overview* [working paper] [Online], 2006, http://coalition4evidence.org/wp-content/uploads/2012/12/RCTs_first_then_match_c-g_studies-FINAL.pdf.

Complicated Matters UK, London: AVA, Against Violence & Abuse, 2013.

Corvo K, Dutton D, Chen W, Toward evidence-based practice with domestic violence perpetrators, *Journal of Aggression, Maltreatment & Trauma*, 16: 111-30, 2008.

Cullen FT, The twelve people who saved rehabilitation: How the science of criminology made a difference, *Criminology*, 43: 1-42, 2005.

Curwood SE, DeGeer I, Hymmen P, Lehmann P, Using strength-based approaches to explore pretreatment change in men who abuse their partners, *Journal of Interpersonal Violence*, 26: 2698-715, 2011.

Davis RC, Taylor BG, Does batterer treatment reduce violence? A synthesis of the literature, *Women and Criminal Justice*, 10: 69 - 93, 1999.

Day A, Chung D, O'Leary P, Carson E, Programs for men who perpetrate domestic violence: An examination of the issues underlying the effectiveness of intervention programs, *Journal of Family Violence*, 24: 203-12, 2009a.

Day A, O'Leary P, Chung D, Justo D, eds., *Domestic violence: Working with men. Research, practice experiences and integrated responses*, Annandale, Australia: The Federation Press, 2009b.

de los Galanes MJ, Tabernero C, El impacto del entrenamiento cognitivo-conductual. Un estudio exploratorio con agresores de género, *Anuario de Psicología Jurídica*, 23: 11-9, 2013.

Debbonaire T, *Responding to diverse ethnic communities in domestic violence perpetrator programmes*. WWP-EN Expert essay [Online], 2015, <http://www.work-with-perpetrators.eu>.

Diranzo RC, Murillo ML, Miñana AC, Cambios psicosociales en un programa de intervención con hombres penados por violencia contra la mujer, *Revista de Psicología*, 21: 159-86, 2012.

Dobash RE, Dobash RP, Cavanagh K, Lewis R, eds., *Changing violent men*, Thousand Oaks, CA: Sage, 2000.

Downes J, Kelly L, Westmarland N, Ethics in violence and abuse research: A positive empowerment approach [Online], *Sociological Research Online*, 19, 2014, <http://www.socresonline.org.uk/19/1/2.html>.

Dunford FW, The San Diego Navy experiment: An assessment of interventions for men who assault their wives, *Journal of Consulting and Clinical Psychology*, 68: 468-76, 2000.

Dutton DG, Corvo K, The Duluth model: A data-impervious paradigm and a failed strategy, *Aggression and Violent Behavior*, 12: 658-67, 2007.

Dutton DG, Corvo K, Transforming a flawed policy: A call to revive psychology and science in domestic violence research and practice, *Aggression and Violent Behavior*, 11: 457-83, 2006.

Echauri JA, Fernández-Montalvo J, Martínez M, Azkarate JM, Effectiveness of a treatment programme for immigrants who committed

gender-based violence against their partners, *Psicothema*, 25: 49-54, 2013.

Echeburúa E, Adherencia al tratamiento en hombres maltratadores contra la pareja en un entorno comunitario: Realidad actual y retos de futuro, *Psychosocial Intervention*, 22: 87-93, 2013.

Echeburúa E, Fernández-Montalvo J, Evaluación de un programa de tratamiento en prisión de hombres condenados por violencia grave contra la pareja, *International Journal of Clinical and Health Psychology*, 9: 50-20, 2009.

Echeburúa E, Sarasua B, Zubizarreta IZ, de Corral P, Evaluación de la eficacia de un tratamiento cognitivo-conductual para hombres violentos contra la pareja en un marco comunitario: Una experiencia de 10 años (1997-2007), *International Journal of Clinical and Health Psychology*, 9: 199-217, 2009.

Eckhardt C, Holtzworth-Munroe A, Norlander B, Sibley A, Cahill M, Readiness to change, partner violence subtypes, and treatment outcomes among men in treatment for partner assault, *Violence and Victims*, 23: 446-75, 2008.

Eckhardt CI, Murphy C, Black D, Suhr L, Intervention programs for perpetrators of intimate partner violence: Conclusions from a clinical research perspective, *Public Health Reports*, 121: 369-81, 2006.

Eckhardt CI, Murphy CM, Whitaker DJ, Sprunger J, Dykstra R, Woodard K, The effectiveness of intervention programs for perpetrators and victims of intimate partner violence, *Partner Abuse*, 4: 196-231, 2013a.

Eckhardt CI, Murphy CM, Whitaker DJ, Sprunger J, Dykstra R, Woodard K, Addendum (detailed tables): Effectiveness of intervention programs for perpetrators and victims of intimate partner violence, *Partner Abuse*, 4: 1-26, 2013b.

Eckhardt CI, *Stages and processes of change and associated treatment outcomes in partner*

assaultive men, Washington, DC: National Institute of Justice, 2004.

Edleson JL, *Groupwork with men who batter: What the research literature indicates* [Online], Harrisburg, PA: VAWnet, a project of the National Resource Center on Domestic Violence, 2012, <http://www.vawnet.org>.

Edleson JL, Syers M, The relative effectiveness of group treatments for men who batter, *Social Work Research*, 26: 10-7, 1990.

Egger T, *Beratungsarbeit und Anti-Gewalt-Programme für Täter und Täterinnen häuslicher Gewalt in der Schweiz. Eine Bestandesaufnahme der Institutionen und ihrer Arbeit*, Bern: Eidg. Büro für die Gleichstellung von Frau und Mann (EBG), 2008.

Elinoff Acker S, *Unclenching our fists: Abusive men on the journey to nonviolence*, Nashville, TN: Vanderbilt University Press, 2013.

Fals-Stewart W, Kashdan TB, O'Farrell TJ, Birchler GR, Behavioral couples therapy for drug-abusing patients: Effects on partner violence, *Journal of Substance Abuse Treatment*, 22: 87-96, 2002.

Fals-Stewart W, Kennedy C, Addressing intimate partner violence in substance-abuse treatment, *Journal of Substance Abuse Treatment*, 29: 5-17, 2005.

Fals-Stewart W, O'Farrell TJ, Birchler GR, Córdova J, Kelley ML, Behavioral couples therapy for alcoholism and drug abuse: Where we've been, where we are, and where we're going, *Journal of Cognitive Psychotherapy*, 19: 229-46, 2005.

Farrington DP, Methodological quality standards for evaluation research, *The Annals of the American Academy of Political and Social Science*, 587: 49-68, 2003.

Farrington DP, Welsh BC, Randomized experiments in criminology: What have we learned in the last two decades?, *Journal of Experimental Criminology*, 1: 9-38, 2005.

Feder L, Dugan L, A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment, *Justice Quarterly*, 19: 343-75, 2002.

Feder L, Dugan L, *A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment*, Washington, DC: U.S. Department of Justice, National Institute of Justice, 2004.

Feder L, Forde DR, *A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment*, Washington, DC: U.S. Department of Justice, National Institute of Justice, 2000.

Feder L, Wilson DB, A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior?, *Journal of Experimental Criminology*, 1: 239-62, 2005.

Feder L, Wilson DB, Austin S, Court-mandated interventions for individuals convicted of domestic violence, *Campbell Systematic Reviews*, 12: 1-46, 2008.

Flood M, Pease B, Factors influencing attitudes to violence against women, *Trauma, Violence, and Abuse*, 10: 125-42, 2009.

Galvani S, *The role of alcohol in violence against women*, Saarbrücken, Germany: Lambert Academic, 2010.

Geffner R, Sorenson SB, Lundberg-Love PK, *Violence and sexual abuse at home: Current issues in spousal battering and child maltreatment*, New York: Haworth Maltreatment & Trauma Press, 1997.

Geldschläger H, Ginés O, Nax D, Ponce Á, *Outcome measurement in European perpetrator programmes: A survey*. WP1 from the Daphne III project 'IMPACT: Evaluation of European perpetrator programmes' [Online], 2014, <http://www.work-with-perpetrators.eu>.

Gilchrist G, Blazquez A, Segura L, Geldschläger H, Valls E, Colom J, Torrens M, Factors associated with physical or sexual

intimate partner violence perpetration by men attending substance misuse treatment in Catalunya: A mixed methods study, *Criminal Behaviour and Mental Health*, 25: 239-57, 2015.

Gilchrist G, Radcliffe P, McMurran M, Gilchrist L, The need for evidence-based responses to address intimate partner violence perpetration among male substance misusers, *Criminal Behaviour and Mental Health*, 25: 233-8, 2015.

Gillioz L, De Puy J, Ducret V, *Domination et violence envers la femme dans le couple*, Lausanne: Payot, 1997.

Gloor D, Meier H, *Evaluation des Pilotprojekts « Soziales Trainingsprogramm für Gewaltausübende Männer »*, Zürich : Basler Interventionsprojekt gegen Gewalt in Ehe und Partnerschaft « Halt-Gewalt » und der Interventionsstelle gegen häusliche Gewalt des Kantons Basel-Landschaft, 2002.

Gloor D, Meier H, *Evaluation des zweiten Pilotjahrs 2002 « Soziales Trainingsprogramm für Gewaltausübende Männer »*, Zürich : Basler Interventionsprojekt gegen Gewalt in Ehe und Partnerschaft « Halt-Gewalt » und der Interventionsstelle gegen häusliche Gewalt des Kantons Basel-Landschaft, 2003.

Goldkamp JS, *Role of drug and alcohol abuse in domestic violence and its treatment: Dade County's domestic violence court experiment*, Washington, DC: U.S. Department of Justice, National Institute of Justice, 1996.

Gondolf EW, Characteristics of court-mandated batterers in four cities, *Violence Against Women*, 5: 1277-93, 1999.

Gondolf EW, Evaluating batterer counseling programs: A difficult task showing some effects and implications, *Aggression and Violent Behavior*, 9: 605-31, 2004.

Gondolf EW, Limitations of experimental evaluation of batterer programs, *Trauma, Violence, & Abuse*, 2: 79-88, 2001.

Gondolf EW, *The future of batterer programs: Reassessing evidence-based practice*, Lebanon, NH: Northeastern University Press, 2012.

Gondolf EW, The weak evidence for batterer program alternatives, *Aggression and Violent Behavior*, 16: 347-53, 2011.

Gondolf EW, Theoretical and research support for the Duluth model: A reply to Dutton and Corvo, *Aggression and Violent Behavior*, 12: 644-57, 2007.

Gondolf EW, Williams OJ, Culturally focused batterer counseling for African-American men, *Trauma, Violence, & Abuse*, 2: 283-95, 2001.

Hagemann-White C, Kavemann B, Beckmann S, Brandfass U, Grieger K, Leopold B, Rabe H, Katenbrink J, *Täterarbeit im Kontext von Interventionsprojekten gegen häusliche Gewalt. Gemeinsam gegen häusliche Gewalt. Kooperation, Intervention, Begleitforschung, band iii*, Berlin: BMFSFJ, 2004.

Hamberger LK, Hastings J, Court-mandated treatment of men who assault their partner: Issues, controversies, and outcomes, in NZ Hilton, ed., *Legal responses to wife assault*, Newbury Park, CA: Sage, 188-229, 1993.

Hamilton L, Koehler JA, Lösel FA, Domestic violence perpetrator programs in Europe, part I: A survey of current practice, *International Journal of Offender Therapy and Comparative Criminology*, 57: 1189-205, 2013.

Hellman CM, Johnson CV, Dobson T, Taking action to stop violence: A study on readiness to change among male batterers, *Journal of Family Violence*, 25: 431-8, 2010.

Hester M, Lilley S-J, *Domestic and sexual violence perpetrator programmes, article 16 of the Istanbul Convention: A collection of papers on the Council of Europe Convention on preventing and combating violence against women and domestic violence*, Strasbourg: Council of Europe, 2014.

Hester M, Lilley S-J, O'Prey L, Budde J, *Overview and analysis of research studies evaluating European perpetrator programmes*. WP2 from the Daphne III project 'IMPACT: Evaluation of European perpetrator programmes' [Online], 2014, <http://www.work-with-perpetrators.eu>.

Hester M, Making it through the criminal justice system: Attrition and domestic violence, *Social Policy and Society*, 5: 79-90, 2006.

Hester M, Westmarland N, Domestic violence perpetrators, *Criminal Justice Matters*, 66: 34-5, 2006.

Heyman RE, Neidig P, Physical aggression couples treatment, in WK Halford, HJ Markman, eds., *Clinical handbook of marriage and couples interventions*, New York: Wiley, 589-617, 1997.

Holtzworth-Munroe A, Meehan JC, Typologies of men who are maritally violent, *Journal of Interpersonal Violence*, 19: 1369-89, 2004.

Horvath AO, Bedi RP, The alliance, in JC Norcross, ed., *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*, New York: Oxford, 37-69, 2002.

IMPACT Toolkit guidelines, Daphne III project 'IMPACT: Evaluation of European perpetrator programmes' [Online], 2014, <http://www.work-with-perpetrators.eu>.

IMPACT Toolkit, Daphne III project 'IMPACT: Evaluation of European perpetrator programmes' [Online], 2014, <http://www.work-with-perpetrators.eu>.

Jacobson NS, Addis ME, Research on couples and couple therapy: What do we know? Where are we going?, *Journal of Consulting and Clinical Psychology*, 61: 85-93, 1993.

Jaquier V, Stein H, Van Daele F, *Accompagner le changement: Le Service pour auteur·e·s de violence conjugale de neuchâtel. Recherche sur mandat de l'Office de la politique familiale et de l'égalité (OPFE) du*

canton de Neuchâtel, Neuchâtel: Université de Neuchâtel, Centre romand de recherche en criminologie, 2016.

Jaspard M, Brown E, Condon S, Firdion J-M, Fougeyrollas-Schwebel D, Houel A, Lhomond B, Maillochon F, Saurel-Cubizolles M-J, Schiltz M-A, *Les violences envers les femmes en France: Une enquête nationale*, Paris: La Documentation Française, 2002.

Jewell LM, Wormith JS, Variables associated with attrition from domestic violence treatment programs targeting male batterers, *Criminal Justice and Behavior*, 37: 1086-113, 2010.

Johnson H, Dawson M, *Violence against women in Canada: Research and policy perspectives*, Oxford: Oxford University, 2010.

Johnson MP, *A typology of domestic violence: Intimate terrorism, violent resistance, and situational couple violence*, Lebanon, NH: Northeastern University, 2008.

Jones AS, D'Agostino RB, Gondolf EW, Heckert A, Assessing the effect of batterer program completion on reassault using propensity scores, *Journal of Interpersonal Violence*, 19: 1002-20, 2004.

Jones C, *Implementing the IMPACT Toolkit*. WWP-EN Expert essay [Online], 2015, <http://www.work-with-perpetrators.eu>.

Kelly L, Dubois L, *Combating violence against women: Minimum standards for support services*, Strasbourg, France: Council of Europe, 2008.

Kelly L, Westmarland L, *Domestic violence perpetrator programmes: Steps toward change [Project MIRABAL]*, London and Durham, England: London Metropolitan University and Durham University, 2015a.

Kelly L, Westmarland N, New approaches to assessing effectiveness and outcomes of domestic violence perpetrator programs, in H Johnson, BS Fisher, V Jaquier, eds., *Critical issues on violence against women:*

International perspectives and promising strategies, London: Routledge, 183-94, 2015b.

Killias M, Simonin M, De Puy J, *Violence experienced by women in Switzerland over their lifespan: Results from the International Violence Against Women Survey*, Bern: Stämpfli, 2005.

Kistenmacher BR, Weiss RL, Motivational interviewing as a mechanism for change in men who batter: A randomized controlled trial, in CM Murphy, RD Maiuro, eds., *Motivational interviewing and stages of change in intimate partner violence*, New York: Springer, 25-42, 2009.

Klein AR, Crowe A, Findings from an outcome examination of Rhode Island's specialized domestic violence probation supervision program: Do specialized supervision programs of batterers reduce reabuse?, *Violence Against Women*, 14: 226-46, 2008.

Kraus H, *Trainingsprogramm zur Beendigung von gewalttätigem Verhalten in Paarbeziehungen: Interne Evaluation (1999-2010)*, Wien: Männerberatung Wien, 2013.

Küken H, Wypych B, *Vorstellung der Evaluationsergebnisse der Interventionsstelle gegen häusliche Gewalt Südpfalz*, 2011.

Labriola M, Rempel M, Davis RC, *Testing the effectiveness of batterer programs and judicial monitoring: Results from a randomized trial at the Bronx Misdemeanor Domestic Violence Court*, Washington, DC: National Institute of Justice, 2005.

Lehmann P, Simmons CA, eds., *Strengths-based batterer intervention: A new paradigm in ending family violence*, New York: Springer, 2009.

Levesque DA, Gelles RJ, *Does treatment reduce recidivism in men who batter? A metaanalytic evaluation of treatment outcome* [Communication], Program Evaluation and Family Violence Research: An International Conference, Durham, NH, 1998.

Lila M, Catalá A, Conchell R, García A, Lorenzo MV, Pedrón V, Terreros E, Una experiencia de investigación, formación e intervención con hombres penados por violencia contra la mujer en la Universidad de Valencia: Programa contexto, *Intervención Psicosocial*, 19: 167-79, 2010.

Lipton DS, Martinson R, Wilks J, *The effectiveness of correctional treatment: A survey of treatment evaluation studies*, New York: Praeger, 1975.

Logar R, *Partnerships with victims' services in work with perpetrators*. WWP-EN Expert essay [Online], 2015, <http://www.work-with-perpetrators.eu>.

Logar R, Rösemann U, Zürcher U, eds., *Gewalttätige Männer ändern (sich). Rahmenbedingungen und Handbuch für ein soziales Trainingsprogramm*, Bern/Stuttgart/Wien: Haupt, 2002.

Logar R, Vargová BM, *Effective multi-agency cooperation for preventing and combating domestic violence. Training of trainers manual*, Strasbourg: Council of Europe, 2015.

Mackay E, Gibson A, Lam H, Beecham D, *Perpetrator interventions in Australia, part 1: Literature review*, Alexandria, Australia: ANROWS, Australia's National Research Organisation for Women's Safety, 2015.

MacLeod D, Pi R, Smith D, Rose-Goodwin L, *Batterer intervention systems in California: An evaluation*, Washington, DC: U.S. Department of Justice, National Institute of Justice, 2010.

Maiuro RD, Eberle JA, State standards for domestic violence perpetrator treatment: Current status, trends, and recommendations, *Violence and Victims*, 23: 133-55, 2008.

McCollum EE, Stith SM, Couples treatment for interpersonal violence: A review of outcome research literature and current clinical practices, *Violence & Victims*, 23: 187-201, 2008.

McMurran M, Motivational interviewing with offenders: A systematic review, *Legal and Criminological Psychology*, 14: 83-100, 2009.

Miller MW, Drake E, Nafziger M, *What works to reduce recidivism by domestic violence offenders*, Olympia, WA: Washington State Institute for Policy, 2013.

Miller WR, Rollnick S, *Motivational interviewing: Preparing people to change*, New York: Guilford Press, 2002.

Moreillon L, Druey J, *Programme imposés pour auteur-e-s de violence dans le couple : Étude d'applicabilité dans le système judiciaire vaudois*, Lausanne: Sur mandat de la Commission cantonale de lutte contre la violence domestique (CCLVD) et du Bureau cantonal de l'égalité entre les femmes et les hommes (BEFH) du Canton de Vaud, 2012.

Mösch Payot P, *Gutachten: Anordnung von Pflichtberatung und Lernprogrammen im Rahmen von strafrechtlichen Sanktionen, insb. als Weisungen*, Bern: Im Auftrag des Eidg. Büro für die Gleichstellung von Frau und Mann (EBG), 2012.

Moyer R, *To BIP, or not to BIP?* [Online], 2004, <http://www.biscmi.org/aquila/positive-effects-of-bips/>.

Mullender A, Burton S, Dealing with perpetrators, in J Taylor-Browne, ed., *What works in reducing domestic violence*, London: Whiting and Birch, 59-94, 2001.

Murphy CM, Baxter VA, Motivating batterers to change in the treatment context, *Journal of Interpersonal Violence*, 12 : 607-19, 1997.

Murphy CM, Maiuro RD, eds., *Motivational interviewing and stages of change in intimate partner violence*, New York: Springer, 2009b.

Murphy CM, Maiuro RD, Understanding and facilitating the change process in perpetrators and victims of intimate partner violence: An introduction and commentary, in CM Murphy, RD Maiuro, eds., *Motivational interviewing and*

stages of change in intimate partner violence, New York: Springer, 3-22, 2009a.

Musser PH, Semiatin JN, Taft CT, Murphy CM, Motivational interviewing as a pregroup intervention for partner-violent men, *Violence and Victims*, 23: 539-57, 2008.

Neidig P, Friedman DH, Collins BS, Domestic conflict containment: A spouse abuse treatment program, *Social Casework: The Journal of Contemporary Social Work*, 66: 195-204, 1985.

NICE, *Domestic violence and abuse: How social care, health services and those they work with can identify, prevent and reduce domestic violence and abuse*, London: National Institute for Health and Care Excellence, 2014.

Oelemann B, Lempert J, ... *dann habe ich zugeschlagen*, Hamburg, Deutschland: Konkret Literatur Verlag, 1995.

O'Farrell TJ, Fals-Stewart W, Behavioral couples therapy for alcoholism and drug abuse, *Journal of Substance Abuse Treatment*, 18: 51-4, 2000.

O'Leary KD, Heyman RE, Neidig PH, Treatment of wife abuse: A comparison of gender-specific and conjoint approaches, *Behavior Therapy*, 30: 475-505, 1999.

Palmer SE, Brown RA, Barrera ME, Group treatment for abusive husbands: Long-term evaluation, *American Journal of Orthopsychiatry*, 62: 276-83, 1992.

Phillips R, Kelly L, Westmarland N, *Domestic violence perpetrator programmes: An historical overview*, London and Durham, England: London Metropolitan University and Durham University, 2013.

Prochaska J, DiClemente C, *The transtheoretical approach: Crossing traditional boundaries of therapy*, Homewood, IL: Krieger, 1984.

Prochaska JO, DiClemente CC, Stages of change in the modification of problem behaviors, in M Hersen, RM Eisler, PM Miller, eds., *Progress in behavior modification*, Newbury Park, CA: Sage, 184-218, 1992.

Radatz DL, Wright EM, Integrating the principles of effective intervention into batterer intervention programming: The case for moving toward more evidence-based programming, *Trauma, Violence, & Abuse*, 2015.

Rajagopalan V, Price P, Langston J, Potter F, *Working towards safety: Supporting women alongside DVIP's perpetrator programme*. WWP-EN Expert essay [Online], 2015, www.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/Expert%20Essays/Working_towards_Safety_DVIP_Rajagopalan.pdf.

Ramírez MP, Giménez-Salinas Framís A, de Juan Espinosa M, Evaluación de la eficacia del programa de tratamiento con agresores de pareja (PRIA) en la comunidad, *Psychosocial Intervention*, 22: 105-14, 2013.

RESPECT, The RESPECT accreditation standard [Online], 2012, <http://respect.uk.net/wp-content/themes/respect/assets/files/accreditation-standard.pdf>.

Robert P, *L'évaluation des politiques de sécurité et de prévention en europe*, Guyancourt, France: Groupe européen de recherche sur les normativités, 2009.

Saunders DG, Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interaction of abuser traits and treatment models, *Violence and Victims*, 11: 393-414, 1996.

Saunders DG, Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interaction of abuser traits and treatment models, *Violence and Victims*, 11: 393-414, 1996.

Shadish WR, Cook TD, Campbell DT, *Experimental and quasi-experimental designs*

for generalized causal inference, Boston, MA, US: Houghton, Mifflin and Company, 2002.

Sharf J, Primavera LH, Diener MJ, Dropout and therapeutic alliance: A meta-analysis of adult individual psychotherapy, *Psychotherapy: Theory, Research, Practice, Training*, 47: 637-45, 2010.

Shepard MF, Pence EL, eds., *Coordinating community responses to domestic violence: Lessons from Duluth and beyond*, Thousand Oaks, CA Sage, 1999.

Sherman LW, Farrington DP, Welsh BC, MacKenzie DL, *Evidence-based crime prevention*, New York: Routledge, 2006,

Smedslund G, Dalsbø TK, Steiro A, Winsvold A, Clench-Aas J, Cognitive behavioural therapy for men who physically abuse their female partner, *Cochrane Database of Systematic Reviews*, 2011.

Stanley N, Borthwick R, Graham-Kevan N, Chamberlain R, *Strength to change: Report of the evaluation of a new initiative for perpetrators of domestic violence*, Preston, UK: University of Central Lancashire, 2011.

Stella Toolkit UK, London: AVA, Against Violence & Abuse, 2007.

Stith SM, McCollum EE, Rosen KH, *Couples therapy for domestic violence: Finding safe solutions*, Washington, DC: American Psychological Association, 2011.

Stith SM, Rosen KH, McCollum EE, Thomsen CJ, Treating intimate partner violence within intact couple relationships: Outcomes of multi-couple versus individual couple therapy, *Journal of Marital Family Therapy*, 30: 305-18, 2004.

Stuart GL, Temple JR, Moore TM, Improving batterer intervention programs through theory-based research, *JAMA*, 298: 560-2, 2007.

Subirana-Malaret M, Andrés-Pueyo A, Retención proactiva y adherencia terapéutica en programas formativos para hombres

maltratadores de la pareja, *Psychosocial Intervention*, 22: 95-104, 2013.

Sullivan CM, Rumpitz MH, Campbell R, Eby KK, Davidson WS, Retaining participants in longitudinal community research: A comprehensive protocol, *The Journal of Applied Behavioral Science*, 32: 262-76, 1996.

Taft CT, Murphy CM, Elliott JD, Morrel TM, Attendance-enhancing procedures in group counseling for domestic abusers, *Journal of Counseling Psychology*, 48: 51-60, 2001.

Taylor BG, Davis RC, Maxwell CD, The effects of a group batterer treatment program: A randomized experiment in Brooklyn, *Justice Quarterly*, 18: 171-201, 2001.

Tjaden P, Thoennes N, *Extent, nature, and consequences of intimate partner violence. Findings from the National Violence Against Women Survey*, Atlanta, GA: Centers for Disease Control and Prevention, 2000.

Van Voorhis P, Cullen FT, Applegate B, Evaluating interventions with violent offenders, *Federal Probation*, 50: 17-27, 1995.

Vlais R, *Domestic violence perpetrator programs: Education, therapy, support, accountability 'or' struggle?* [Online], Melbourne, Australia: No To Violence, Male Family Violence Prevention Association, 2014, <http://ntv.org.au/>.

Walker K, Bowen E, Brown S, Sleath E, Desistance from intimate partner violence: A conceptual model and framework for practitioners for managing the process of change, *Journal of Interpersonal Violence*, 30: 2726-50, 2015a.

Walker K, Bowen E, Brown S, Sleath E, Subjective accounts of the turning points that facilitate desistance from intimate partner violence, *International Journal of Offender Therapy and Comparative Criminology*, 2015b.

Walker K, Bowen E, Mentoring serial and high-risk perpetrators of intimate partner violence in the community: Engagement and initiating

change, *Criminal Behaviour and Mental Health*, 25: 299-313, 2015.

Weisz A, *Spouse Assault Replication Program: Studies on effects of arrest on domestic violence*, 2001.

Welsh BC, *Evidence-based crime prevention: The scientific basis, trends, and results*, Ottawa, Canada: National Crime Prevention Centre, Department of Public Safety and Emergency Preparedness Canada, 2007.

Westmarland N, Kelly L, Why extending measurements of 'success' in domestic violence perpetrator programmes matters for social work, *British Journal of Social Work*, 43: 1092-110, 2013.

White RJ, Gondolf EW, Implications of personality profiles for batterer treatment, *Journal of Interpersonal Violence*, 15: 467-88, 2000.

Wojnicka K, *Masculinities, men and violence*. WWP-EN Expert essay [Online], 2015, <http://www.work-with-perpetrators.eu/>.

Wojnicka K, Scambor C, Kraus H, New pathways in the evaluation of programmes for men who perpetrate violence against their female partners, *Evaluation and Program Planning*, 57: 39-47, 2016.