To treat or not to treat?

Harmful sexual behavior in adolescence: Needs before risk

ter Beek, E.

Citation for published version (APA):
UITNODIGING

Voor het bijwonen van de openbare verdediging van het proefschrift

TO TREAT or NOT TO TREAT?
Harmful Sexual Behavior in Adolescence:
Needs before Risk

Door Ellis ter Beek

Woensdag 14 februari 2018 om 12.00 uur in de Agnietenkapel, Oudezijds Voorburgwal 229-231, 1012 EZ te Amsterdam

Receptie na afloop ter plaatse

Paranimfen
Agnes de Lima-Heyns
Fatima Ben Allouch

Voor vragen:
DrEllisPromoveert@gmail.com
TO TREAT
or
NOT TO TREAT?

Harmful Sexual Behavior in Adolescence:
Needs before Risk

Ellis ter Beek
Colofon

Cover art: Rosanna Hoogendoorn-Feikens, www.ateliercana.nl
Printing: Ridderprint BV, www.ridderprint.nl
ISBN: 978-94-6299-850-6

Copyright © Ellis ter Beek 2017. ALL RIGHTS RESERVED. Any unauthorized reprint or use of this material is prohibited. No part of this thesis may be reproduced, stored or transmitted in any form or by any means, without written permission of the author or, when appropriate, of the publishers of the publications.
To Treat or not to Treat?
Harmful Sexual Behavior in Adolescence: Needs before Risk

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op 14 februari 2018, te 12.00 uur

door

Ellis ter Beek
geboren te Enschede
Promotiecommissie

Promotores
Prof. Dr. G.J.I.M. Stams
Prof. Dr. J. Hendriks
Universiteit van Amsterdam

Copromotores
Dr. C.H.Z. Kuiper
Dr. R.E.A. van der Rijken
Horizon Jeugdzorg en Onderwijs
De Viersprong

Leescommissie
Prof. Dr. H.J.C. van Marle
Prof. Dr. E.J. Knorth
Prof. Dr. X.M.H. Moonen
Prof. Dr. R.J.L. Lindauer
Dr. E.S. van Vugt
Erasmus Universiteit Rotterdam
Rijksuniversiteit Groningen
Universiteit van Amsterdam
Universiteit van Amsterdam

Faculteit
Faculteit der Maatschappij en Gedragswetenschappen
Contents

Chapters | Page
---|---
1 General Introduction | 7
2 The Allocation of Sexually Transgressive Juveniles to Intensive Specialized Treatment, an Assessment of the Application of RNR Principles | 23
3 Treatment Effect on Recidivism for Juveniles who have Sexually Offended: a Multilevel Meta-Analysis | 45
4 Treatment Effect on Psychosocial Functioning of Juveniles with Harmful Sexual Behavior: a Multilevel Meta-Analysis | 69
5 Primum Non Nocere: A Comparative Study of Differences and Changes in Treatment Needs of Juveniles with Harmful Sexual Behavior Receiving Residential or Non-Residential Treatment | 97
6 General Discussion | 121
7 References | 135
8 Summary | 163

Appendices | Page
---|---
Nederlandse Samenvatting | 169
Dankwoord | 171
Curriculum Vitae | 175
Publicaties | 177
General Introduction
JUVENILE SEXUAL PROBLEM BEHAVIOR: DEFINITIONS AND TERMINOLOGY

The distinction between adequate and transgressive sexual behavior can be obscure and, therefore, subject of debate. Laws and mores have clearly differed through time, and still do culturally. In Western societies, the idea that children and juveniles are developing sexual beings is often alarming to adults, a common misconception that any sexual behavior displayed by children is inappropriate. Sexual development, however, is part of the total developmental process, and starts at birth. Pre-school children are naturally curious and exploratory, they explore their own bodies, and the bodies of others. Research indicates that 40-85% of children engage in at least some sexual behaviors or sexual play before the age of 13. When another juvenile is involved, these events are often recalled as positive or neutral if there was consent, and equality. During adolescence especially, juveniles actively experiment with sexuality and relations, discovering boundaries and developing their sense of self in relation to others (De Graaf, Mouthaan, & Van der Doef, 2014).

Problematic sexual behaviors in juveniles can range from relatively minor (e.g., excessive masturbation, verbal intimidation) to prosecutable acts (e.g., rape, child abuse; Sunday, 2016). Therefore, not all juveniles who display problematic sexual behavior automatically classify as offenders, since not all inappropriate sexual behaviors are illegal. Interactions without consent and equality, and with coercion (pressure or acts that deny free choice) are generally labeled ‘abusive’, and usually constitute an offense. From the perspective of treatment, this dissertation focuses on juveniles with harmful (coercive) sexual behavior, as well as on juveniles who have sexually offended. Depending on the subject of research, the term juvenile with transgressive, harmful or offensive sexual behavior will be used in different chapters. Harmful sexual behavior is operationalized as entailing all illegal sexual acts, as well as all other aggressive or coercive sexual behaviors. The common factor in the juveniles studied is their need for treatment to improve their level of functioning, and to prevent relapse into harmful behavior.

Prevalence

Global estimates of the prevalence of juvenile sexual victimization (i.e., unwanted, coerced sexual contact) vary widely, from 0.1 percent to 71 percent. The mean prevalence, based on self-report, was estimated to be 12.7 percent (Stoltenborgh et al., 2011). In the Netherlands, national surveys have established that one in three juveniles is victimized sexually, with five to ten percent of Dutch juveniles experiencing unwanted oral sex or intercourse (Nationaal Rapporteur, 2014).
Perpetrators of juvenile sexual abuse have been found most likely to be male adolescents or adults (Långström, Grann, & Lindblad, 2000), known to the juvenile, and member of the same household (Romans, Martin, Anderson, O’Shea, & Mullen, 1996). In the United States, Finkelhor, Ormrod, and Chaffin (2009) found juveniles to account for one third (35.6 percent) of sexual offenders who had victimized a minor. Australian 2013 crime data showed juveniles to be responsible for 18 percent of all recorded sex offenses against juveniles and adults (Warner & Bartels, 2015). In the same year in the Netherlands, this figure was 14 percent (Statline, 2017). Of all sex crime suspects in the Netherlands, 25 percent were juveniles, 98 percent was male. Most of the total amount of transgressions reported were contact offenses (61%), and most suspects (75%) were acquainted to the victim (Nationaal Rapporteur, 2014).

Due to underreporting, official criminal records generally underestimate true prevalence rates of sex crimes (White, 2011; Wittebrood, 2006). In recent years, however, the reported prevalence of juvenile sex crimes has been declining (Caldwell, 2016; Van den Berg, 2015). In the Netherlands, 1,705 juveniles were suspected of a sexual offense in 2005, in 2015 this number had dropped to 350 (Statline, 2017). This declining trend was also visible in non-sexual offending and, therefore, to be considered generic (CBS, 2017). This fact points to less crimes being committed in general, therefore broader societal explanations may apply. An explanation for the drop in reported sexual offending specifically is a possible shift towards more anonymous, non-contact types of sexual transgressions on social media and internet (e.g., viewing of (child)pornography, sexting, sextortion), which have increased in recent years, and are reported less than contact offenses (Helpwanted, 2017). Also, in the Netherlands, since 2010, an adjudication for an offense is no longer a prerequisite for juveniles to receive mandated (secure) treatment. This may have had an effect on the number of persecutions; another (civil justice) pathway has become available when intensive treatment is deemed necessary (see also, ‘treatment in the Netherlands’).

**Impact of sexual offending**

Prevention of sexual victimization is deemed important for society in general, and to be aimed at juveniles at risk of victimization and at juveniles at risk of sexual (re)offending (Nationaal Rapporteur, 2014). The psychological consequences of (juvenile) sexual victimization are substantial. Several researchers have found anxiety and trauma related disorders, chronic depression and developmental disorders, such as behavioral and attachment disorders in victims of sexual abuse (Amado, Arce, & Herraiz, 2015; Lindauer & Boer, 2012). A meta-review of the consequences of child sexual victimization (Nagtegaal, 2012), showed the additional problems of victims, as compared to non-victims, to be predominantly medical, psychological and/or sexual in nature, and of influence up until adulthood. Additionally,
victims of sexual violence were found to be at significantly higher risk for re-victimization compared to non-victims (Arrata, 2002; Nagtegaal, 2012), enhancing and/or maintaining the negative consequences of the abuse. For boys, these consequences include higher than average odds of displaying sexually offensive behavior later on in life (Salter et al., 2003; Jespersen, Lumiere, & Seto, 2009).

Juveniles with harmful sexual behavior generally (self)report specific treatment needs, indicating the presence of antisocial behavior problems, but also reporting internalizing and sexual problems (Fanniff & Kimonis, 2014; Seto & Lalumiere, 2010). Some of these treatment needs have been established as criminogenic, underscoring the importance of treatment as to prevent relapse. Different theoretical views have contributed to the understanding of the development of sexual problem behavior (e.g., biological, cognitive, behavioral, social learning, and integrative or multifactorial theories). In the following section, a concise overview of etiological theories, research findings supporting the theories, and their accompanying treatment interventions is offered.

Origins of harmful sexual behavior in juveniles

(Social) learning theories (Bandura, 1977; Pavlov, 1927; Skinner, 1974) consider -sexual-transgressions to be a learned (i.e., conditioned or modeled) behavior, which consequently might be unlearned through the administration of adequate consequences and prosocial modeling. Research has indeed shown sexually offensive juveniles to have been more exposed to porn, to have experienced more sexual abuse than non-sexually offensive peers (Burton, 2008; Fanniff & Kimonis, 2014; Leibowitz, Burton & Howard, 2012; Seto & Lalumière, 2010), and their backgrounds in general to include sexual aggression (Awad & Saunders, 1991; Grabell & Knight, 2009; Kobayashi et al., 1995; Veneziano, Veneziano, & LeGrand, 2000), underpinning a possible learned aspect of harmful sexual behavior. Providing discouraging consequences to coercive (sexual) behavior or deviant arousal, modeling and promoting / rewarding adequate behaviors, have become regular (basic) corrective therapeutic interventions.

Developmental theories (Erikson, 1968; Freud, 1965; Piaget, 1928) have explained the development of dysfunctional behavior through experiences in earlier stages of development. These experiences, however, not (just) model dysfunctional behavior, as learning theories suggested, but lead to a delay in accomplishing developmental tasks. Developmental delay is theorized to have a profound impact on an individual’s ability to complete subsequent, more complex tasks, such as starting or maintaining an intimate relationship. Thus, egocentric (sexually harmful) behavior in adolescence, or reduced empathic ability are viewed as a result of an ‘askew development’ or a lack of acquired social-emotional skills. A lack of social-emotional skills or developmental deficits are often found
in juveniles with sexually harmful behavior (Baarsma et al., 2016; Veneziano & Veneziano, 2002). Therefore, stimulating the acquirement of more adequate social-emotional skills and stimulation of (moral) development via psychoeducation and training, generally is part of treatment for juveniles with sexually harmful behavior.

Attachment theory (Ainsworth, 1989; Bowlby, 1966, 1973) has specified mechanisms by which early life trauma may affect the development of rigid interpersonal behavior patterns and creates (pervasive) intimacy deficits. According to attachment theory, internal working models (i.e., cognitive-affective representations) of early relationships shape self-image and the expectations one has of others. Self-confirmatory coercive interaction patterns may, therefore, develop when basic needs of (small) children are not met, or are violated (Hoeve et al., 2012). Attachment-based therapy focusses on restoring trust (via therapeutic relationships), improving self-image and counteracting these rigid interaction patterns. Miner and colleagues’ (2008, 2010, 2016) research on attachment styles of juveniles with sexually offensive behavior has shown anxious attachment to be -albeit indirectly- related to sexually harmful behavior of juveniles against children. Also, sexually offensive juveniles have experienced more abuse and neglect (precursors of attachment problems) and portray more (social) anxiety than their non-sexually offensive delinquent peers (Leibowitz, Burton, & Howard, 2012; Seto & Lalumière, 2010). Especially for juveniles with harmful sexual behavior who have been victims of (sexual) abuse or neglect, restoring trust and intimacy via therapeutic relationships seems imperative.

Regarding the continuation of transgressive behaviors in general, cognitive theories added the element of ‘thinking errors’ or ‘distortions’ (Yochelson & Samenow, 1976). These cognitions depict harmful behaviors as acceptable, justifiable, or harmless and thus serve as a sustaining factor. Abel and colleagues (1989) found adult child abusers to frequently make use of these cognitions. Identifying and restructuring or challenging these thoughts has become a basic tenant in cognitive (group) treatment for juveniles with sexually harmful behavior. Theories that focus on addiction added the notion that sexually harmful behavior, due to its innate physiological rewards (arousal, orgasm, and tension reduction), might become a compulsive, ritualized pre-occupation, also known as hypersexuality (Carnes, 1983). Strategies of addiction treatment (e.g., peer group therapy and the identification of ‘hotspots’ for sexual arousal) have also been found useful to aid the ‘breaking’ of an addictive or compulsive cycle, and have since been incorporated in therapy for juveniles with sexually harmful behavior.

Lastly, social-ecological theory (Bronfenbrenner, 1979) drew attention to the fact that factors explaining or maintaining antisocial (sexual) behavior in juveniles occur across ecological systems in which juveniles are embedded (e.g., school / work, peer groups,
family). Bronfenbrenner therefore stressed the importance of creating change in the (supportive) context of juveniles with problem behavior, not just a change in the juvenile himself. Contextual approaches to reducing harmful sexual behavior by juveniles have since been developed.

The now dominant cognitive-behavioral treatment (CBT) paradigm combines the knowledge as described above. In this paradigm, sexually harmful behavior is seen as a form of social behavior, shaped and maintained by the social environment. The person portraying this behavior is regarded as the result of a complex interplay of biology, emotions, cognitions and motor behavior (Ward, Polaschek, & Beech, 2006). CBT incorporates diverse treatment strategies, flexibly responding to a juvenile’s individual treatment needs and his specific social context.

Whom, What and How to treat?

In 1974, Martinson published an influential essay, stating ‘nothing works’ in offender rehabilitation and recidivism prevention. This was reacted to by several scholars, proving, via reviews and meta-analyses, that ‘not everything works for everybody’ (Andrews, Bonta, & Hoge, 1990; Andrews, et al., 1990; Gendreau, Little, & Goggin, 1996; Gendreau, Smith, & French, 2006; Lipsey & Cullen, 2007; Lipsey & Wilson, 1998; Lösel, 1995; Smith, Gendreau, & Swartz, 2009), a much-needed nuance. Based on this work, the Risk-Need-Responsivity (RNR) model for offender rehabilitation was developed (Andrews, 1995; Andrews & Bonta, 2010; Andrews et al., 1990; Gendreau, 1996), which since has become the dominant paradigm in (juvenile) offender rehabilitation. Because of its scientific rigor and great practical use, the RNR model has also been applied in adjacent fields, such as mental healthcare (Lord, 2016; Skeem, Steadman, & Manchak, 2015).

The RNR model importantly states effective rehabilitation to follow three specific guidelines: 1) the risk principle, or whom to treat (i.e., intensive, mandated treatment has to be offered to juveniles at medium to high recidivism risk only), 2) the need principle, or what to treat (i.e., criminogenic treatment needs have to be addressed), and 3) the responsivity principle, or how to treat (i.e., treatment should make use of social cognitive learning strategies, such as CBT, and be tailored to the personality, learning style, and motivation of the juvenile concerned; Andrews & Bonta, 2010). Treatments adhering to these standards have been found to be more effective in reducing recidivism than treatments that do not (Bonta & Andrews, 2007). Via meta-analysis, Hanson, Bourgon, Helmus, and Hodgson (2009) have shown the RNR principles to also apply to adults and juveniles who have sexually offended. All three principles, and the challenges applying them to juveniles with harmful sexual behavior, will be discussed below.
Risk

Research into recidivism risk factors for harmful sexual behavior in juveniles is hampered by low, and declining, rates of sexual recidivism by juveniles. Studies performed between the years 2000 and 2015 reported a weighted mean sexual recidivism rate of 2.75 percent (Caldwell, 2016). Most juveniles with sexual harmful behavior do not reoffend sexually (Cale, Smallbone, Rayment-McHugh, & Downling, 2016), and desistance is norm rather than exception (Lussier, Van den Berg, Bijleveld, & Hendriks, 2012); sexual transgressions by juveniles remain predominantly adolescence limited. To date, several specific risk factors for sexual reoffending by juveniles have, however, been established. Worling and Långström (2003), reviewed previous research and found sexual deviation (e.g., interest in prepubescent children or sexual violence), prior criminal sanctions for sexual assault(s), having made two or more victims, victimization of strangers, lack of intimate peer relationships/social isolation, and incomplete offense-specific treatment sufficiently empirically supported. Two other factors were deemed promising: poor relationship with parents and attitudes supportive of sexual offending. In 2011, Carpentier and Proulx examined 351 male adolescents who had sexually offended and found paternal abandonment, childhood sexual victimization, association with younger children, and having victimized a stranger to be associated with a higher risk for sexual recidivism. Christiansen and Vincent (2013), using a large sample of 39,249 juvenile offenders, of which 695 were juveniles who had offended sexually, added prior non-sexual offending, hands-off offending, offending against a child, younger age at time of initial offense, poor school performance and non-school attendance, and replicated prior sexual offending as risk factors for sexual reoffending. The risk factors found predominantly represent classifications of the juveniles’ offenses (i.e., number of victims, type of victim, type of offensive behavior, number of previous arrests) or of their past (i.e., victimization, abandonment, non-treatment completion).

Non-sexual recidivism among juveniles with sexually harmful behavior is actually more common than sexual recidivism; in contemporary studies, 30 percent of juveniles with harmful sexual behavior relapse into non-sexual transgressions (Caldwell, 2016). The most influential established risk factors for generic delinquent behavior include an antisocial behavior pattern, antisocial attitudes and values, antisocial associates, and substance abuse (Andrews, Bonta, & Wormith, 2006). Juveniles with harmful sexual behavior displaying these characteristics are, therefore, considered at higher risk for (repeat) non-sexual transgressions. They, however, also tend to reoffend sexually more often (Chu & Thomas, 2010; Drew, 2013; Hissel et al., 2006; Parks & Bard, 2006). Risk factors for generic delinquent behavior are, therefore, usually included in sexual recidivism risk assessment and allocation practices.
Introduction

Through systematic review of recidivism risk assessment tools for juveniles with harmful sexual behavior, Hempel, Buck, Cima, and Van Marle (2011) found none of the instruments reviewed (J-SOAP-II, J-SORRAT-II, ERASOR, JRAS, SAVRY, and PCL:YV) undisputed. Low sexual reoffending rates make recidivism hard to predict, most contemporary actuarial instruments seem to overrate the actual risk of reoffending in these juveniles. Hempel and colleagues (2011) therefore concluded that long term assessment of risk, based on characteristics in (developmental) transition, is not to be advised.

Need

The RNR model specifically focuses on the prevention of recidivism by targeting dynamic risk factors, also referred to as criminogenic needs. Many of the above mentioned risk factors for sexual reoffending, however, are considered static (not to be influenced by therapy) and, therefore, provide a relatively small foundation for treatment purposes. In general, juveniles with harmful sexual behavior have been found to differ from otherwise delinquent or antisocial juveniles by presenting more extensive histories of early sexual exposure / abuse and physical and emotional abuse or neglect, more atypical sexual interests, poorer social relationships, higher levels of anxiety, and lower self-esteem (Seto & Lalumiere, 2010). Fanniff and Kimonis (2014) additionally found a lower level of callous unemotional traits in juveniles with harmful sexual behavior. Thus, in general, these juveniles display higher levels of internalizing problems, than do juveniles with other types of transgressive behavior.

Typology research, furthermore, indicated juveniles with harmful sexual behavior to form a heterogeneous group regarding their treatment needs. Peer abusing juveniles and juveniles with a ‘mixed offending pattern’, which includes non-sexual problem behavior, have relatively more in common with non-sexual juvenile offenders than do juveniles with a child victim (≥ 5 years younger and below the age of 12). The ‘mixed’ and ‘peer victim’ groups show higher levels of antisocial conduct problems (Drew, 2013; Leroux, Pullman, Montayne, & Seto, 2016). Intrapsychic problems seem dominant in juveniles with child victims (Hendriks, 2006). Notably, a relatively large group of juveniles with harmful sexual behavior does not report any intra- or interpersonal difficulties; they score within the normal range on psychosocial measures, possibly pointing to situational or developmental phase-bound explanations of harmful sexual behavior (Caldwell, 2010; Ryan, Leversee & Lane, 2010; Van Outsem, 2009; Worling, 2013). Contextual treatments may prove a better fit for these juveniles.

In sum, treatment needs relatively common in juveniles with harmful sexual behavior (e.g., low self-esteem, social anxiety), may not have been empirically established as criminogenic due to the low prevalence rate of sexual reoffending. Hence, treatment for juveniles with
harmful sexual behavior is usually not strictly limited to dynamic risk factors or criminogenic treatment needs, as proposed by the RNR model. Protective factors and non-criminogenic psychosocial treatment needs especially seem somewhat underexposed in the RNR model.

Responsivity

In the RNR model, social learning strategies are advocated under the responsivity principle. As described, modern (established) treatments for juveniles with harmful sexual behavior make use of cognitive behavioral based treatment, incorporating social learning strategies, thereby generally adhering to the RNR model. Additionally, responsive treatment needs to adapt itself to the personality, learning style, and motivation of the individual juvenile in treatment. The heterogeneity of juveniles with harmful sexual behavior offers a challenge to the individual therapist. Protocolled treatments for juveniles with harmful sexual behavior need to be able to cope with very different juveniles with different treatment needs. David Prescott’s motto ‘personalize the manual, do not manualize the person’ (personal communication) is well known and valued among therapists working with juveniles with harmful sexual behavior. Responsive treatment needs to be tailored to individual learning styles and motivations for change, which importantly may not overlap their assessed criminogenic treatment needs. CBT for juveniles with harmful sexual behavior needs to be able to address a very diverse set of possible underlying problems or social-psychological mechanisms and, therefore, make use of various treatment techniques. This may influence treatment integrity, which is thought of as an important prerequisite for treatment success (Lipsey, 2009).

In sum, (intensive, mandated) treatment should be aimed at those juveniles at high risk for sexual recidivism, but these juveniles are hard to identify prospectively, that is, at allocation. Treatment should be aimed at specific dynamic criminogenic needs, which may offer a restricted view on (the diverse) treatment needs of juveniles with harmful sexual behavior. In addition, delivering treatment responsively may affect treatment integrity.

Treatment in the Netherlands

In the Netherlands, juveniles primarily in need of treatment for conduct problems are generally treated separately from juveniles primarily showing harmful sexual behavior. A broad spectrum of specialized treatments is available. A behavioral training order (Jonker, De Haas, Reijmers, Rekers, & Van Berlo, 2011) and specialized outpatient treatment (Hendriks, 2011) are the least intrusive and most frequently used types of specialized treatment offered throughout the country (Hendriks, 2013). These treatments are generally indicated for low to medium risk offenders, able to (safely) stay at home during treatment.
Juveniles in need of mandated treatment may be allocated to this type of treatment via one of two ‘pathways’, 1) via civil law and 2) via juvenile justice court. Dutch law on youth care (Hirsch Ballin, 2007) states that the invasion of fundamental children’s rights, such as restricting freedom of movement, must be kept to a minimum and, therefore, is only applicable (by civil court order) if so advised by an authorized psychologist, who evaluates whether placement in a Secure Youth Care (SYC) facility is in accordance with the juvenile’s treatment needs. For adjudicated juveniles (12 to 18 years of age at the time of the offense), mandated treatment in a Forensic Youth Care (FYC) facility is available. In accordance with the European Rights of Children (Council of Europe, 1996), only juveniles who pose the highest possible risk to society are to receive treatment in a maximum security residential facility (which is offered by an FYC). This policy generally adheres to RNR principles, basing allocation to intensive mandated treatment on levels of risk and treatment need.

A third, community-based, treatment option, making use of both allocation pathways, is multisystemic therapy for juveniles with problem sexual behavior (MST-PSB; Dwyer & Letourneau, 2011). A description of the three treatment modalities is offered in following.

MST-PSB is aimed at adjudicated and non-adjudicated juveniles aged 10 to 18 years, who exhibit sexual problem behavior, have complex (family) issues, and pose a risk to society (Boonstra & Van der Rijken, 2010). Since MST-PSB is a community-based, at home treatment, the family system actively takes responsibility for safety issues. MST-PSB is based on regular MST, and targets youth and family problems within and between the multiple systems in which the family members are embedded. The overarching goal of MST is to empower parents and adolescents with the skills and resources needed to cope with their familial and extra-familial problems. Using treatment strategies derived from strategic family therapy, structural family therapy, behavioral parent training, and cognitive-behavioral therapy, MST addresses intrapersonal (e.g., cognitive), familial, and extra familial (e.g., peer, school, neighborhood) factors that are known to be linked with a juvenile’s serious antisocial behavior, including sexual offending. If, and when, specific individual interventions are used to modify the juvenile’s social perspective-taking skills, belief system, or attitudes believed to contribute to offending and the sexual assault cycle, the parent is reinforced as the change agent, and is directly involved in the intervention. The exact nature of interventions applied varies for each family, depending on their strengths and weaknesses (Borduin & Schaeffer, 2001). Contraindications for the Dutch MST-PSB program are an IQ < 70, acute psychiatric problems, and severe forms of autism (Boonstra & Van der Rijken, 2010).

Secure Youth Care (SYC) offers treatment to non-adjudicated juveniles (aged 12 to 18 years) with Severe Sexual Problems (SSP). During their stay, cognitive behavioral based group treatment is offered by youth care professionals, and monitored by a behavioral scientist. Social learning strategies are explicitly used to stimulate change, and the specific
interventions applied are tailored to the juvenile’s treatment needs. At admittance, (criminogenic) treatment needs are assessed by means of the Juvenile Sex Offender Assessment Pack-Dutch version (J-SOAP-D; Bullens, Van Horn, & Van Eck, 2012), and used for treatment planning purposes. Group therapy ‘Out of the Circle’ (Koster & Tel, 2010), a central part of the treatment offered, focusses on the sexual misconduct. It assesses the cycle of sexual aggression (triggers, thoughts, behaviors, and consequences), and creates an individualized risk management plan. Individual therapy (as to address possible victimization experiences or trauma) may supplement the group treatment offered. The family system of the juvenile is actively involved in the treatment of the juvenile, furlough at home and family guidance is used to transfer skills and knowledge. Contraindication for placement is an IQ < 70; these juveniles receive treatment in a secure institution for the mentally disabled.

Forensic Youth Care is available for adjudicated juveniles (12 to 18 years of age at the time of the offense) with Severe Sexual Problems (FYC-SSP). Their measure (a ‘PIJ-maatregel’; court ordered placement in a forensic institution for juveniles) is usually imposed for two years, and may remain imposed for a maximum period of six years. One year of aftercare (outpatient care and supervision) is also mandatory. At commencement of treatment, a behavioral scientist performs a comprehensive offense analysis, together with the juvenile, to determine treatment needs. For risk assessment purposes, tools such as the J-SOAP-D (Bullens et al., 2012) and the Structured Assessment of Violence Risk in Youth (SAVRY; Lodewijks, Doreleijers, de Ruiter & de Wit-Grouls, 2006) are regularly administered. A stay in FYC-SSP implies treatment through YOUTURN, a comprehensive, phased, cognitive behavioral based method (Stuurgroep YOUTURN, 2009) that focuses on all developmental tasks for adolescents (Erikson, 1968). YOUTURN is supplemented with group therapy ‘Out of the Circle’ (Koster & Tel, 2010) and individual therapy. The juvenile’s family system is involved in treatment, family guidance is used to transfer skills and knowledge. Contraindication for placement is an IQ < 70, these juveniles receive treatment in a forensic very intensive care (VIC) treatment group for the mentally disabled.

Notably, the three (mandatory) treatment options available (MST-PSB, SYC-SSP, and FYC-SSP) are all theoretically aimed at a small group of very problematic juveniles that are deemed at risk for sexual recidivism. Clear indication criteria that adequately distinguish between these treatment types, however, are lacking.
STUDY AIM

This dissertation aims to contribute to the knowledge on treatment of juveniles with sexually transgressive behavior. It provides insight into current allocation practices, reviews contemporary research on the effects of treatment, and further specifies what individual treatment needs are most common and most successfully targeted in treatment.

Samples

This dissertation uses two Dutch samples:

1. A sample of 86 juveniles who received MST-PSB, SYC-SSP, or FYC-SSP treatment for harmful sexual behavior between January 2010 (when all three types of treatment were operational) until April 2012, and for whom comprehensive background information was coded and a risk assessment instrument; J-SOAP-D (Bullens et al., 2012) was (retrospectively) scored.

2. A sample of 36 juveniles who received MST-PSB, or SYC-SSP treatment for harmful sexual behavior between February 2012 (when measurements started) up until October 2015, and completed the Adolescent Sexual Abuser Project Assessment Measures-Dutch Revised version (ASAP-D; Van Outsem, Beckett, Bullens, Vermeiren, Van Horn, & Doreleijers, 2006) at admission and at completion of treatment.

Furthermore, this dissertation (quantitatively) reviews primary studies on juveniles with sexually transgressive behavior reporting on the effect of treatment on recidivism, including 13 independent samples, with 77 effect sizes and 1,726 participants, and primary studies on juveniles with sexually transgressive behavior reporting on the effect of treatment on psychosocial functioning, including 31 independent samples, with 362 effect sizes, and 1,342 participants.
Chapter 1

DISSERTATION OUTLINE

Central aim of this dissertation is to determine what ‘type’ of juvenile with harmful sexual behavior should be treated by what ‘type’ of intensive, mandated treatment, aimed at which individual treatment needs. Are there differences in approach or effects that could guide allocation of these juveniles to better matched care? How can treatments improve their results by incorporating contemporary research findings into their practices? The following sub-questions were formulated, which are answered in the corresponding chapters:

Chapter 2. Were juveniles with harmful sexual behavior in intensive, mandated treatment in the Netherlands, allocated according to RNR principles? What are (main) characteristics of the juveniles in intensive, mandated treatment?

Chapter 3. Is treatment of juveniles with harmful sexual behavior effective in reducing recidivism? Do participant, treatment or study characteristics moderate this effect?

Chapter 4. Is treatment of juveniles with harmful sexual behavior effective in improving psychosocial functioning? Do participant, treatment or study characteristics moderate this effect?

Chapter 5. What are the effects on psychosocial functioning obtained by intensive, mandated treatment for juveniles with harmful sexual behavior in the Netherlands and how may treatments improve their results?

This dissertation closes with a general discussion on the main findings, practical implications, limitations, and recommendations for future research (Chapter 6).
The Allocation of Sexually Transgressive Juveniles to Intensive Specialized Treatment, an Assessment of the Application of RNR Principles

Ter Beek, E.
Van der Rijken, R.E.A.
Kuiper, C.H.Z.
Hendriks, J.
Stams, G.J.J.M.

doi:10.1177/0306624X16674684
ABSTRACT

The aim of this study was to compare results of and develop guidelines for mandatory allocation of sexually transgressive juveniles to Multisystemic Therapy - Problem Sexual Behavior (MST-PSB), Secure Youth Care (SYC), or Forensic Youth Care (FYC), based on the Risk-Need-Responsivity (RNR) model. Results of current allocation showed a population with relatively less treatment needs to receive community based MST-PSB, compared to populations receiving residential SYC and FYC. Furthermore, estimated recidivism risk levels did not always support the need for risk reduction by the imposition of limitation of freedom of movement and maximum supervision, provided by all three treatment modalities. Based on the assessed sexual recidivism risk, 38% of the juveniles in FYC, 7% in SYC, and 24% in MST-PSB received treatment that was too intensive, which is considered detrimental to motivation and development. Future allocation practices could benefit from assessing treatment needs and recidivism risk, by use of an actuarial tool.
INTRODUCTION

A significant proportion of sexual offenses are considered to be committed by juveniles. Studies estimate that between 20% of rapes and 30% to 50% of child molestations are committed by adolescent males (Barbaree & Marshall, 2006), for whom consequently allocation to treatment is an option. Research into allocation practices, however, is scarce. Such research may prove useful, especially among juveniles who have shown sexually harmful behavior; a notoriously heterogeneous group. Several studies into typologies, specific mental illnesses, traits, executive functioning, and personality profiles of adolescents who have sexually offended have been conducted (e.g., Adjorlolo & Egbenya, 2016; Butler & Seto, 2002; Drew, 2013; Glowacz & Born, 2012; Hart-Kerhoff, Doreleijers, Jansen, Van Wijk, & Bullens, 2009; Hendriks, 2006; Hissel, Bijleveld, Hendriks, Jansen, & Collot d’Escurey-Koenigs, 2006; Kjellgren, Wassberg, Carlberg, Långström, & Svedin, 2006; Lawing, Frick, & Cruise, 2010; Margari et al., 2015; Purcel, 2010). Caldwell (2010), after conducting a meta-analytic review, stated that developmental issues (i.e., incomplete maturation, impulsiveness or exploratory behavior) might be important factors in the etiology of juvenile sex offending. Seto and Lalumiere (2010), also through a meta-analytic study, found sexually transgressive juveniles to have risk factors in common with juvenile non-sex offenders (i.e., age of onset, antisocial thought and behavior patterns), but also found them to differ (i.e., more sexual victimization and deviancy, more social and emotional dysfunctions, and more learning problems), suggesting other risk factors may (also) play a role in the development of sexual transgressive behavior. Lussier, Van den Berg, Bijleveld, and Hendriks (2012) defined a small subgroup (10% of their total sample, \( N = 498 \)) of sexually transgressive adolescents who sexually offended at a relatively high rate, and of whom 60% persisted in sexual reoffending even beyond adolescence; the high rate, slow desisters. None of the aforementioned traits or typologies, however, were found exclusively in this group, making it hard to identify them prospectively, that is, at allocation.

For allocation practices, the etiology of transgressive behavior is very important. For when deemed primarily antisocial or adolescence limited in nature, non-specialist treatment may proof a good fit. When origins are deemed more sexually deviant or pervasive in nature, a specialized type of treatment, aiming prominently at, for example, specific cognitive distortions, may proof the best fit. We aim to contribute to the literature on juvenile sex offender typology by offering a description and comparison of recidivism risk levels and treatment needs of juveniles who have sexually offended in the Netherlands.
Treating Juveniles who have Sexually Offended in the Netherlands

In Europe, important differences exist between juvenile justice systems. Countries all have developed their own approach in offering juveniles the help they need and/or punishing undesirable behavior (Walgrave & Mehlbye, 1998). For example; in the Netherlands the lower limit for adjudication is twelve years of age, whereas in the neighboring countries of Great Britain and Belgium this limit respectively is ten versus eighteen years of age (Brouwers, 2007). In most European countries a continuum between youth protection services and a juvenile justice system is available. Both systems usually are equipped to respond to transgressive behavior and are able to mandate supervision or treatment. In comparison to other European countries, in the Netherlands a relatively large number of juveniles are detained (and receive treatment there). Only in Germany, and in England and Wales juveniles are imprisoned more often (Kalidien & De Heer-De Lange, 2015). In contrast, in Sweden and Denmark, detaining juveniles is very rare. Core of the Dutch policy on justice involved interventions is to intervene only when the development of a juvenile is threatened. Compared to other European countries, justice involvement is less prevention focused (Bol, 2002).

In Figure 1 a representation of the Dutch process of (mandatory) allocation into intensive treatment for sex offending is presented. Non-specialized treatment options (i.e., treatment aimed at transgressive behavior in general) or non-intensive treatments available (i.e., psychoeducation on sexual transgressions (Jonker, De Haas, Reijmers, Rekers & Van Berlo, 2011), and specialized outpatient treatment (Hendriks, 2011)) are not included in this diagram. In both allocation pathways (‘civil law’ on the left and ‘criminal justice’ on the right) an evaluation is performed in order to establish the etiology of the problem behavior, and determine risks and treatment needs. Different professionals and methods, however, are used. In the ‘civil law pathway’, firstly, a youth protection board caseworker assesses the reported situation. When the decision is made to resort to mandated residential care, an independent psychologist has to agree. Specialized contextual treatment can be administered without an independent assessment of treatment needs. In the ‘criminal justice pathway’ the prosecutor’s office firstly establishes whether or not prosecution is feasible. If a decision to prosecute is reached, an independent forensic psychological and psychiatric evaluation, making use of structured risk assessments, is performed in order to advise the juvenile judge.

If specialized intensive treatment is necessary, a five to seven months community-based treatment; multisystemic therapy for juveniles with problem sexual behavior (MST-PSB; Borduin, Henggeler, Blaske, & Stein, 1990) is available via both pathways. MST-PSB is aimed at adjudicated and non-adjudicated juveniles aged 10 to 18 years, who exhibit sexually deviant behavior, have complex (family) issues, and pose a risk to society (Borduin, et. al.,
Since MST-PSB is a community-based type of treatment, the family system involved needs to actively take responsibility for safety issues during the at home treatment. The exact nature of interventions applied varies for each family, depending on their strengths and weaknesses, although the interventions mainly involve cognitive behavioral and family therapeutic techniques (Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 2009). Contraindications for the MST-PSB program are an IQ < 70, acute psychiatric problems, and severe forms of autism (Boonstra & Van der Rijken, 2010).

Figure 1. Schematic process of allocation to intensive specialized treatment
Another intensive specialized treatment available for non-adjudicated juveniles (aged 12 to 18 years), is mandated specialized treatment in a Secure Youth Care (SYC) facility. Dutch law on youth care (Hirsch Ballin, 2007) states that the invasion of fundamental children's rights, such as restricting freedom of movement, must be kept to a minimum, and therefore is only applicable (by civil court order) if so advised by an authorized psychologist. SYC should adhere to minimalizing imposed restrictions of freedom and the shortest possible stay, with juveniles receiving further treatment in less restrictive settings (Veldhuijzen van Zanten-Hyllner, 2011). The contraindication for placement is an IQ < 70; these juveniles receive treatment in institutions for the mentally disabled. One specialized SYC treatment group (10 beds) is available in the Netherlands.

A second possibility for adjudicated juveniles who have sexually offended (12 to 18 years of age at the time of the offense) is treatment in Forensic Youth Care (FYC), of which there are 16 beds available in the Netherlands. In accordance with the European Rights of Children (Council of Europe, 1996), only juveniles who pose the highest possible risk to society are to receive treatment in a maximum security residential facility. Their treatment measure (a ‘PIJ-maatregel’; placement in a forensic institution for juveniles) is usually imposed for a minimum of two years, and may remain imposed for a maximum period of six years. One year of aftercare (care and supervision outside the institution) is also mandatory. A contraindication for placement is an IQ < 70; these juveniles receive treatment in a forensic treatment group for the mentally disabled.

<table>
<thead>
<tr>
<th>Capacity</th>
<th>Legal Framework</th>
<th>Target Group</th>
<th>Risk/Need Assessment</th>
<th>Specific Treatment modules</th>
<th>Contraindications</th>
</tr>
</thead>
<tbody>
<tr>
<td>MST-PSB</td>
<td>24 clients per year</td>
<td>Voluntary &amp; Mandated, Adjudicated &amp; Non adjudicated</td>
<td>Sexually deviant behavior, complex issues, and a risk to society</td>
<td>10-18 years of age</td>
<td>Risk Management Plan, CBT and Family Therapeutic techniques</td>
</tr>
<tr>
<td>SYC</td>
<td>10 beds Mandated, Non adjudicated</td>
<td>Sexually deviant behavior, complex issues, and a risk to society</td>
<td>12-18 years of age</td>
<td>Risk Management Plan J-SOAP-D, Group CBT, Out of the Circle, Individual treatment</td>
<td>IQ &lt; 70</td>
</tr>
<tr>
<td>FYC</td>
<td>16 beds Mandated, Adjudicated</td>
<td>Sexually deviant behavior, complex issues, and a risk to society</td>
<td>12-18 years of age</td>
<td>Risk Management Plan J-SOAP-D SAVRY, YOUTURN, Out of the Circle, Individual treatment</td>
<td>IQ &lt; 70</td>
</tr>
</tbody>
</table>

**Figure 2.** Main characteristics of intensive specialized treatments offered
Notably, the three intensive treatment options available are all theoretically aimed at a relatively small group of very problematic and ‘at risk’ sexually transgressive juveniles. Clear differentiating indication criteria between the treatment types are lacking.

**Evidence-based Allocation: Literature Review**

A well-known and empirically-based model for guiding allocation to mandatory treatment is the Risk-Need-Responsivity (RNR) model (Andrews, Bonta, & Hoge, 1990; Andrews & Bonta, 2010; Andrews, Bonta & Wormith, 2006; Andrews, Bonta & Wormith, 2011; Bonta & Andrews, 2007). The RNR model is built on general personality and cognitive social learning theory. Importantly, it states that treatment allocation should be based on three principles. The first is the risk principle, which matches the level of service to the client’s risk to re-offend (who should be treated). The second is the need principle, which assesses criminogenic needs as to target these in treatment (what should be treated). The third is the responsivity principle, which maximizes a client’s ability to learn from an intervention by providing cognitive behavioral treatment and by tailoring the intervention to the learning style, motivation, abilities, and strengths of the client (how the client should be treated).

The inaccurate matching of treatment intensity with treatment need and risk level has been shown to lead to a waste of resources (Bonta & Andrews 2007; Cecile & Born, 2009; Hendriks & Bijleveld, 2008). Low risk offenders receiving intensive treatment, especially in residential settings, have even shown an increase in criminal behavior (Bonta, Wallace-Capretta & Rooney, 2000; Lowenkamp, Latessa & Holsinger, 2006). Adhering to the RNR model is associated with a reduction of recidivism, while non-adherence is associated with increased recidivism or null effects on reoffending (Andrews & Dowden, 2007). As a consequence, the characteristics of treatment should generally match treatment needs and risk levels. As treatment needs and risk levels increase, the intensity of the treatment should increase accordingly. Hanson, Bourgon, Helmus, and Hodgson (2009) have, via meta-analysis, shown the RNR principles to apply to adults and juveniles who have sexually offended. Skeem, Steadman, and Manchak (2015), based on their review of RNR literature, stated that in general there is a large body of empirical support for the RNR model to be applied to mentally disordered clients in need of treatment.

**Evidence Based Allocation in the Netherlands: Hypotheses**

When adhering to the first RNR principle (Risk), in the allocation of sexually transgressive juveniles to intensive treatment in the Netherlands, only ‘high risk’ juveniles are expected to receive FYC, for this constitutes mandatory, long-term, intensive and comprehensive treatment at a maximum security level. At the commencement of FYC treatment, a behavioral specialist and the juvenile perform a comprehensive offense analysis to
determine (criminogenic) treatment needs. For risk assessment, tools such as the Juvenile Sex Offender Assessment Protocol – Dutch (J-SOAP-D; Bullens, Van Horn, & Van Eck, 2012) and the Structured Assessment of Violence Risk in Youth (SAVRY; Lodewijks, Doreleijers, de Ruiter, & de Wit-Grouls, 2006) are regularly administered. Promoting the safety of society is a primary directive of FYC; fostering a juvenile's well-being and health are important secondary goals. Juveniles are notably restricted to a reduced degree and for a shorter period of time in community-based MST-PSB and residential SYC. That said, the freedom restrictions in MST-PSB and SYC are still considerable; both offer 24-hour monitoring, a focus on risk management, and comprehensive protocolled treatment by skilled professionals targeting all areas of development. In MST-PSB, at admission, a risk management plan is made based on the referrer's, the parents', and the juvenile's description of the problem behavior and, if present, using the judicial files. Clinical, non-structured, assessment demonstrates which factors constitute a risk. Parents monitor the risk management plan and are addressed as responsible for the safety of their own and other children. They are stimulated to ask and receive support from their informed and activated social network (e.g., extended family, neighborhood), community (e.g., school, local police), and the professional MST-PSB counsellors. In SYC, the J-SOAP-D (Bullens et al., 2012) is scored at admission for risk assessment and treatment planning purposes. In the first period of residential SYC treatment, no leave outside the institution is granted (until a safety plan is in place). Therefore, when adhering to the Risk principle, only medium to high risk cases are expected to be allocated to MST-PSB and SYC.

Notably, for the assessment of recidivism risk in juveniles who have sexually offended, two instruments are regarded ‘state of the art’ (Christiansen & Vincent, 2013; Lodewijks & Domburg, 2012). These are the Estimate of Risk of Adolescent Sexual Recidivism (ERASOR; Worling & Curwen, 2001) and the Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II) by Prentky and Righthand (2003). The ERASOR is not used in the Netherlands. The J-SOAP Dutch version (J-SOAP-D) is a translated and adapted Dutch version of the J-SOAP-II (Bullens et al., 2012). Sexually transgressive adolescents, however, are a relatively small group, and sexual recidivism is rare among them (Caldwell, 2010; Lobanov-Rostovsky, 2015; Mulder, Vermunt, Brand, Bullens, & Van Marle, 2012). Therefore, it is difficult to determine risk factors through research, as larger numbers of reoffenders are needed for robust statistical analysis of predictors. Hempel, Buck, Cima, and Van Marle (2011), in a review of risk assessment instruments for juveniles who have sexually offended, concluded that no one instrument to date has shown unequivocal positive results in predicting future offending. Specialized tools, such as J-SOAP-II, best predict sexual recidivism. Thus, long-term judgments are never to be based solely upon any single risk instrument and the term ‘high risk’ must be used with caution, for it is a potentially unwarranted and therefore damaging label.
Regarding the Need principle, Hanson, Bourgon, Helmus, and Hodgson (2009) defined a ‘criminogenic need’ as a dynamic risk factor that is theoretically and empirically associated with criminal behavior and can be influenced by means of treatment. In an overview by Worling and Långström (2003), the following empirically supported criminogenic factors for sexual reoffending in adolescents were defined: sexual deviation (e.g., interest in prepubescent children or sexual violence), prior criminal sanctions for sexual assault(s), two or more victims, stranger victimization, lack of intimate peer relationships/social isolation, and incomplete offense-specific treatment. Two further factors were deemed “promising”: poor relationship with parents and attitudes supportive of sexual offending. Christiansen and Vincent (2013), using a large sample of 39,249 juvenile offenders, of which 695 were juveniles who had offended sexually, recently added prior non-sexual offending, prior sexual offending, hands-off offending, offending against a child, and younger age at time of initial offense as risk factors for sexual reoffending. Some of these factors are static (not to be influenced by therapy).

MST-PSB targets deficits in the adolescent’s cognitive processes (denial, empathy, distortions), family relationships (cohesion, empowering parents, monitoring skills, supervision), peer relationships, and school performance. In SYC, cognitive behavioral based treatment is offered by youth care professionals, and monitored by a behavioral specialist. Criminogenic treatment needs are assessed and central to therapy. Social learning strategies are explicitly used to stimulate change. Group therapy “Out of the Circle” (Koster & Tel, 2010) is additionally administered, which focusses on the sexual misconduct and on creating an individualized risk management plan with the juvenile. The plan is presented by the juvenile to his parents/social network, providing insight into the transgressive behavior and its origins. Individual therapy may supplement the group treatment offered. A stay in FYC implies treatment through YOUTURN, a comprehensive, phased, cognitive behavioral based method (Stuurgroep YOUTURN, 2009) that focuses on all developmental tasks for adolescents (Erikson, 1968). YOUTURN is supplemented with group therapy “Out of the Circle” (Koster & Tel, 2010) and individual therapy. In sum, the dynamic factors; sexual deviation, social isolation, relationship with parents, and attitudes supportive of sexual offending, are described as central elements in FYC, SYC, and MST-PSB treatment, with MST-PSB taking a more holistic approach to all factors mentioned, and focusing primarily on treatment of the family (parental) system. Due to the out of home placement of the juvenile, FYC and SYC focus less on parental training; yet sexual deviation, social isolation, and attitudes of the juvenile are specifically central to therapy.

The Responsivity principle comprises two elements: (i) general and (ii) specific responsivity. General responsivity calls for the use of cognitive social learning methods to influence behavior. Cognitive social learning strategies are found to be the most effective regardless
of the type of offender. Core correctional practices, such as prosocial modeling, the appropriate use of reinforcement and disapproval, and problem solving (Dowden & Andrews, 2004), comprise the specific skills represented in a cognitive social learning approach. Specific responsivity is a ‘fine tuning’ of the cognitive behavioral intervention. It takes into account the strengths, learning style, personality, motivation, and biosocial (e.g., gender, race) characteristics of the individual. MST-PSB, SYC, and FYC all use cognitive social learning based methods to influence behavior. They also instruct their professionals on how to take into account the individual strengths, learning style, and personality of the juvenile being treated (i.e., specific responsivity). In the current study, however, differences in the responsivity of the three treatment forms were not assessed, so no statements on differences in responsivity can be made.

In summary, the three most intensive treatment modalities available in the Netherlands for young sexually transgressive juveniles appear to show an overlap in RNR criteria. Firstly, all three should only be administered to juveniles with a high and specific treatment need. Based on legislation and therapy characteristics, however, some differences in target group are expected. For medium to high risk sexually transgressive juveniles, MST-PSB or SYC treatment seems an adequate choice. Within this group, only those in unsafe living situations may need to be placed in SYC. Finally, only high risk (re)offenders may need to be treated in a maximum security setting (an FYC), as to also protect society.

**STUDY AIM**

The current study aims to strengthen the process of allocation to specialized treatment, by a retrospective investigation of the recidivism risk and treatment needs of the current target groups of (intensive) contextual and residential treatments. Currently, the differences and similarities between the various intensive treatment populations are unknown, as is the degree to which mandatory treatment allocation follows RNR principles. This study aims to further illuminate these areas and provide guidelines for future allocation practices.

**METHOD**

**Participants**

Case file information was coded for 86 adolescent boys (MST-PSB: N=25; SYC: N=29; FYC: N=32). This group constituted all juveniles (no girls, IQ ≥ 70) admitted to high intensity specialized treatment in the Netherlands between January 2010 and April 2012. This study thus investigated the total Dutch population of sexually transgressive juveniles allocated
to intensive specialized treatment. The majority of the boys were of Dutch origin (77%) and their mean age was 14.8 years (SD = 2.1) at the start of treatment. Table 1 provides an overview of background characteristics.

### Table 1. Overview of background characteristics

<table>
<thead>
<tr>
<th></th>
<th>MST-PSB N = 25</th>
<th>SYC N = 29</th>
<th>FYC N = 32</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total IQ</td>
<td>93.6 (15.6)</td>
<td>86.6 (10.7)</td>
<td>90.0 (13.4)</td>
<td>1.321 (0.274)</td>
</tr>
<tr>
<td>Age at index beh.</td>
<td>12.5 (2.5)</td>
<td>13.3 (1.5)</td>
<td>15.1 (1.4)</td>
<td>15.277 (&lt;0.001**)</td>
</tr>
<tr>
<td>Number of victims</td>
<td>2.0 (1.1)</td>
<td>3.5 (3.2)</td>
<td>3.5 (3.2)</td>
<td>2.216 (0.117)</td>
</tr>
<tr>
<td><strong>% (N)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex only</td>
<td>84.0% (21)</td>
<td>55.2% (16)</td>
<td>56.3% (18)</td>
<td>6.152 (0.046)</td>
</tr>
<tr>
<td>First offenders</td>
<td>44.0% (11)</td>
<td>17.2% (5)</td>
<td>12.5% (4)</td>
<td>8.690 (0.013)</td>
</tr>
<tr>
<td>Victims 12- and 5y↓</td>
<td>32.0% (8)</td>
<td>29.6% (8)</td>
<td>59.4% (19)</td>
<td>18.657 (0.005**)</td>
</tr>
<tr>
<td>Solo offenders</td>
<td>87.5% (21)</td>
<td>92.6% (25)</td>
<td>100% (32)</td>
<td>3.920 (0.141)</td>
</tr>
<tr>
<td>Single parent home</td>
<td>64.0% (16)</td>
<td>55.2% (16)</td>
<td>31.3% (10)</td>
<td>14.044 (0.081)</td>
</tr>
<tr>
<td>Prior out of home pl.</td>
<td>28.0% (7)</td>
<td>72.4% (21)</td>
<td>56.3% (18)</td>
<td>10.802 (0.005**)</td>
</tr>
<tr>
<td>Prior neglect</td>
<td>52.0% (13)</td>
<td>86.2% (25)</td>
<td>68.7% (22)</td>
<td>7.473 (0.024)</td>
</tr>
<tr>
<td>Prior abuse</td>
<td>40.0% (10)</td>
<td>51.7% (15)</td>
<td>43.3% (13)</td>
<td>0.813 (0.666)</td>
</tr>
<tr>
<td>Prior sexual abuse</td>
<td>16.0% (4)</td>
<td>25.0% (7)</td>
<td>31.3% (10)</td>
<td>3.020 (0.554)</td>
</tr>
<tr>
<td>Domestic violence</td>
<td>45.8% (11)</td>
<td>60.7% (17)</td>
<td>50% (16)</td>
<td>1.263 (0.531)</td>
</tr>
<tr>
<td>Pervasive disorder</td>
<td>12.0% (3)</td>
<td>24.1% (7)</td>
<td>53.1% (17)</td>
<td>12.091 (0.002**)</td>
</tr>
<tr>
<td>Disruptive disorder</td>
<td>76.0% (19)</td>
<td>62.1% (18)</td>
<td>43.8% (14)</td>
<td>6.187 (0.045)</td>
</tr>
<tr>
<td>Symptoms of (a) Personality disorder</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>43.8% (14)</td>
<td>28.219 (&lt;0.001**)</td>
</tr>
<tr>
<td>Paraphilia</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>25.0% (8)</td>
<td>14.885 (0.001**)</td>
</tr>
</tbody>
</table>

*a, b: significant post-hoc difference.

* = significant at a 0.05 level, ** = significant at a 0.01 level, *** = significant at a 0.001 level.

### Treatment forms

MST-PSB is provided by de Viersprong treatment center in the southern region of the Netherlands. Mean treatment duration of the MST-PSB group examined in this study was 7 months (SD =1.1, range 3-8 months). SYC treatment for sexually transgressive youth is provided by Horizon, center for youth care and education. Mean stay in the SYC group examined in this study, including aftercare, was 22 months (SD = 10.3, range 1-53 months).
There are two specialized FYC treatment groups for sexually transgressive behavior in the Netherlands, both at the Den Hey-Acker state institution for juveniles. Mean stay of the FYC group examined in this study, including aftercare, was 47 months (SD = 14.8, range 7-69 months).

**Instruments**

Due to the availability of a validated Dutch version, the J-SOAP Dutch version (J-SOAP-D) (Bullens et al., 2012) was selected as the instrument to be used in the assessment of the need and risk levels in this study. The J-SOAP-D was designed to be used with boys aged 12 to 18 years who had been adjudicated for sexual offenses, as well as non-adjudicated youths with a history of sexually coercive behavior. Prentky and colleagues (2010) also established the use of the J-SOAP-II as valid for pre-adolescents with a high risk of reoffending. Its items were generated from an extensive review of risk factor literature (Prentky & Righthand, 2003). The J-SOAP-D consists of two static criminogenic need scales and two dynamic criminogenic need scales which are used to estimate the level of recidivism risk. The scale items are scored on a three-point scale according to the instruction manual, after which an average score per scale is calculated (ranging from zero to one). Scale 1 measures sexual drive and preoccupation (i.e., previous sex offenses, number of victims, duration of sex offense history, and sexual victimization history). Scale 2 measures impulsive and antisocial behavior (i.e., caregiver consistency, school behavior problems, previous arrests, and history of physical assault/exposure to domestic violence). Scale 3 measures intervention items (i.e., accepting responsibility, internal motivation for change, empathy, shame and guilt, and cognitive distortions). Finally, scale 4 measures community stability and adjustment (i.e., management of sexual urges and desires, stability of living situation, stability in school, and a positive support system). A complete overview of the scales, items, and scoring criteria is available in the J-SOAP-D manual (Bullens et al., 2012). Recidivism risk is classified in two ways, namely sexual recidivism risk and general recidivism risk, which are coded as low, medium, or high.

Research on the reliability of the J-SOAP-II shows a good to excellent interrater reliability, ranging from 0.75 to 0.91, with an average of 0.83 (Prentky & Righthand, 2003). Validity outcomes (Elkovitch, Viljoen, Scalora & Ullman, 2008; Hecker, Scoular, Righthand, & Nangle, 2002; Martinez, Flores & Rosenfeld, 2007; Powers-Sawyer & Miner, 2009; Prentky et al., 2010; Rajlic & Gretton, 2010; Viljoen et al., 2008) however, vary. An area under the curve (AUC) value of 0.50 indicates that an estimate is considered equal to chance (i.e., flipping a coin is considered as accurate as using the instrument). A value between 0.70 and 0.75 is considered moderate; a value greater than 0.75 to 0.80 is considered good (Fawcett, 2006). The J-SOAP AUC values measured for general recidivism (all types of re-offending) range...
from 0.53 (Elkovitch et al., 2008) to 0.76 (Martinez et al., 2007). J-SOAP AUC values measured for sexual recidivism range from 0.44 (Elkovitch et al., 2008) to 0.83 (Prentky et al., 2010). Therefore, the validity of the J-SOAP-D is considered to be restricted.

**Procedure**

The study complied with the APA ethical principles regarding research with human participants. Permission for the retrospective case file research performed was granted by the Dutch Ministry of Justice and the participating organizations. All boys receiving treatment via one of the included treatment forms from January 2010 (when all three types of treatment were operational) until April 2012 (when case files were scored) were included. Case files contained referral information (including psychodiagnostic evaluations), treatment reports, and judicial documents on criminal history. A broad range of items was coded by a single assessor. Special attention was given to diagnostic history, treatment history, and offense history, as well as to protective and risk factors. All items were operationalized in a codebook and the majority concerned factual information; therefore, little interpretation of data was required. Nevertheless, seven files were scored by a second assessor in order to establish inter-coder reliability through the use of intra-class correlation (ICC). Following the guidelines by Shrout (1989), single ICCs were calculated for the two-way random effects model, with absolute agreement as a criterion. The single ICC provides information about the reliability of the risk assessment performed by a single evaluator. Critical ICC values are defined as follows: ICC $\geq 0.81$ = substantial, $0.60 \leq ICC \leq 0.80$ = moderate, $0.41 \leq ICC \leq 0.60$ = fair, ICC $\leq 0.40$ = slight (Shrout, 1989). In the scoring of case files, single measures showed an ICC of 0.99 (i.e., substantial reliability).

As a J-SOAP-D was not available for all juveniles, J-SOAP-D questionnaires were completed retrospectively by a trained assessor, based on the case file information available at admission. Twenty cases (23% of the total sample) were scored by a second trained assessor as to establish interrater reliability. Calculation of single measures for the J-SOAP set yielded an ICC of 0.71 (i.e. moderate reliability). One case file was excluded from the study due to insufficient information. This case did not substantially differ from the other cases.

**Statistical analyses**

All statistical analyses were performed using the IBM Statistical Package for the Social Sciences version 22.

In order to analyze group differences on the four treatment need scales, a one-way ANOVA procedure and Bonferroni post-hoc analyses with a two-tailed probability level of $< 0.05$ was used. We tested the differences in mean score (the higher this score, the more problematic
juveniles -on average- are perceived to be). No standardization of J-SOAP-D treatment need scores is available, so no statements about a ‘low, medium, or high treatment need’ measured is possible. Dunnet’s one-tailed post-hoc analysis was used for the “Community stability” scale, because MST-PSB clients were expected to show relatively lower (less problematic) scores than SYC or FYC clients on this scale.

To compare the estimated sexual and general recidivism risk of the three treatment groups, Chi-square tests with a two-tailed probability level of < 0.05 were performed. Percentages of low, medium, and high at recidivism risk cases in each of the treatment programs were calculated and compared, expecting an even distribution between MST-PSB and SYC and relatively more high at risk juveniles in FYC.

As to assess the clinical relevance of differences found, Cohen's d was calculated, using the calculator of Wilson (2013) and formulas of Lipsey and Wilson (2001). Effect sizes in terms of Cohen's d are considered small (0.10 to 0.30), medium (0.30 to 0.50), or large (≥ 0.50) (Cohen, 1992). The power of the analysis was also assessed, and showed the sample size to be too small to detect medium size effects, at p = .05 and power = .80, requiring a total of 133 participants more or less equally distributed across categories. Thus, using the current population, even when they represent the total population in a period of two years, small -but possibly clinically relevant- effect sizes could not be detected. Therefore, effect sizes that just failed to reach significance (trends; p. < 0.1) were reported.

RESULTS

Although Table 2 (pre-treatment comparison of treatment needs) shows that comparison of Sexual drive scores failed to reach significance, it did show a trend towards a difference between the groups. Post-hoc analyses revealed a trend towards MST-PSB clients scoring as less problematic than FYC clients (p = 0.056; Cohen’s d = 0.66). On the Impulsive/antisocial scale, SYC-clients scored highest (most problematic). MST-PSB clients scored least problematic on these items. Post-hoc analyses confirmed that MST-PSB juveniles scored significantly less antisocial than did SYC juveniles (p = 0.015; Cohen’s d = 0.78).

On Intervention items, FYC clients scored most problematic, whereas MST-PSB clients scored least problematic. Post-hoc analyses confirmed that MST-PSB juveniles scored significantly higher on problem insight and motivation than did FYC juveniles (p = 0.039; Cohen’s d = 0.67). Regarding Community stability, MST-PSB clients scored the lowest, which indicates a higher stability. Post-hoc analyses showed that MST-PSB clients scored significantly lower than SYC and FYC clients (p = 0.021 and p = 0.028; Cohen’s d = 0.69 and 0.56, respectively).
Table 2. Comparison of criminogenic treatment needs/J-SOAP scale scores

<table>
<thead>
<tr>
<th></th>
<th>MST-PSB Mean (SD)</th>
<th>SYC Mean (SD)</th>
<th>FYC Mean (SD)</th>
<th>F (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexual drive</td>
<td>0.44± (0.18)</td>
<td>0.54± (0.20)</td>
<td>0.57± (0.19)</td>
<td>2.99 (0.056)</td>
</tr>
<tr>
<td>Impulsiveness/antisocial</td>
<td>0.32± (0.27)</td>
<td>0.51± (0.23)</td>
<td>0.44± (0.24)</td>
<td>4.23 (0.018)</td>
</tr>
<tr>
<td>Intervention items</td>
<td>0.70± (0.29)</td>
<td>0.78± (0.20)</td>
<td>0.84± (0.13)</td>
<td>3.23 (0.045)</td>
</tr>
<tr>
<td>Community stability</td>
<td>0.50± (0.20)</td>
<td>0.64± (0.19)</td>
<td>0.63± (0.24)</td>
<td>3.31 (0.042)</td>
</tr>
</tbody>
</table>

¹ Range = 0 – 1; higher scores indicate increasingly problematic juveniles. ²: significant post-hoc difference. * = p < 0.05. † = trend, p < 0.10

As Table 3 (pre-treatment comparison of estimated recidivism risk) shows, the assessed sexual recidivism risk did not differentiate between the three treatment groups. There was a trend indicating that juveniles with low sexual recidivism risk received MST-PSB more often than FYC (p = .087, Cohen’s d = 0.32).

General recidivism risk showed a significant differentiation (p = 0.045). Low at general recidivism risk juveniles were treated more often by MST-PSB than by SYC (Cohen’s d = 0.68).

Table 3. Comparison of estimated recidivism risk/J-SOAP scores

<table>
<thead>
<tr>
<th></th>
<th>MST-PSB % (N)</th>
<th>SYC % (N)</th>
<th>FYC % (N)</th>
<th>χ² (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexual recidivism risk</td>
<td></td>
<td></td>
<td></td>
<td>6.37 (0.173)†</td>
</tr>
<tr>
<td>Low</td>
<td>24.0% (6)</td>
<td>7.1% (2)</td>
<td>9.4% (3)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>40.0% (10)</td>
<td>42.9% (12)</td>
<td>28.1% (9)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>36.0% (9)</td>
<td>50.0% (14)</td>
<td>62.5% (20)</td>
<td></td>
</tr>
<tr>
<td>General recidivism risk</td>
<td></td>
<td></td>
<td></td>
<td>9.76 (0.045)†</td>
</tr>
<tr>
<td>Low</td>
<td>68.0% (17)</td>
<td>28.6% (8)</td>
<td>41.0% (13)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12.0% (3)</td>
<td>39.3% (11)</td>
<td>37.5% (12)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>20.0% (5)</td>
<td>32.1% (9)</td>
<td>21.9% (7)</td>
<td></td>
</tr>
</tbody>
</table>

¹: significant post-hoc difference. *= p < 0.05. † = trend, p < 0.10

Taking into account the substantial medium to large Cohen’s d’s that were found, all results, including trends, were deemed clinically relevant (Dunst, Hamby, & Trivette, 2004).
Chapter 2

DISCUSSION

The aim of this study was to assess the recidivism risk and criminogenic treatment needs of sexually transgressive juveniles treated by MST-PSB, SYC, or FYC in the Netherlands, in order to ascertain whether treatment allocation was performed in accordance with RNR principles. All three treatment modalities are aimed at juveniles with complex issues and show similar allocation criteria. However, two distinctive allocation processes are in use to allocate sexually transgressive juveniles to treatment. In both pathways (‘civil law’ or ‘criminal justice’) an evaluation of treatment needs and risks is performed, but different professionals use different assessment techniques (non-structured clinical to actuarial). Notably, the (severity of) problem behavior does not differentiate between pathways. The response of people surrounding the juvenile or the decision whether or not to prosecute a case does (see Fig. 1). Only juveniles in need of specialized treatment (i.e., moderate to high criminogenic treatment needs score on the J-SOAP-D) were expected to receive MST-PSB, SYC, or FYC. Juveniles deemed unsuitable for outpatient treatment (high score on community (in)stability) and assessed at moderate to high recidivism risk were expected to receive SYC. Only juveniles assessed at ‘high risk’ for recidivism were expected to receive treatment in an FYC. To test the hypotheses, the J-SOAP-D scale scores (criminogenic treatment need) and the assessed recidivism risk were compared.

The comparison of treatment needs, as based on the J-SOAP-D scale scores, revealed a number of differences. MST-PSB clients showed a reduced treatment need in comparison with FYC or SYC clients. MST-PSB clients particularly scored less problematic than FYC and SYC clients on impulsive and antisocial behavior (i.e., caregiver consistency, school behavior problems, previous arrests, and history of physical assault/exposure to domestic violence), intervention items (i.e., accepting responsibility, internal motivation for change, empathy, shame and guilt, and cognitive distortions) and community stability and adjustment items (i.e., management of sexual urges and desires, stability of living situation, stability in school, and positive support system). On sexual drive and preoccupation (i.e., previous sex offenses, number of victims, duration of sex offense history, and sexual victimization history) a trend was found for MST-PSB clients to score less problematic than FYC clients. Since there are no cut off scores available for the criminogenic needs scale, it was not possible to determine if scores of MST-PSB clients should be considered ‘low’ and the clients therefore were not in need of specialized treatment. Regarding the etiology of the sexual problem behavior, the clients of intensive specialized treatment in the Netherlands seem more ‘generalist / antisocial’ in SYC, and more ‘sexually deviant’ in FYC, underscoring the viability of both typologies for sexually transgressive juveniles.
On their assessed recidivism risk, groups also differed at the time of allocation. More juveniles with an estimated low recidivism risk were referred to MST-PSB. Transgressive youth with a more ‘generalist’ pattern of offending (multiple type of rule breaking) generally tend to reoffend sooner and at an increased rate, sexually as well as violently (Chu & Thomas, 2010; Drew, 2013; Hissel et al., 2006; Parks & Bard, 2006), so a target group at greater recidivism risk appears to be treated in SYC and FYC. In FYC, however, only juveniles assessed as ‘high at sexual recidivism risk’ were expected to be treated. Based on this premise, 38% (low and medium at sexual recidivism risk) of FYC juveniles still seem to have received a too restrictive type of treatment. Seven percent of the juveniles in SYC, and 24% in MST-PSB, were assessed to be at low sexual recidivism risk, whereas only medium to high at sexual recidivism risk juveniles were expected to be treated here. Therefore, if estimated sexual recidivism risk is perceived a valid indicator for receiving intensive specialized treatment, 38% of the juveniles in FYC, 7% in SYC, and 24% in MST-PSB appear to have been misallocated.

In sum, youth referred to MST-PSB seem less problematic (i.e., on treatment needs and estimated recidivism risk levels) than juveniles referred to SYC and FYC. Low at recidivism risk youth are not in need of specialized intensive treatment and may even suffer adverse effects if allocated as such. Therefore, both referrers and MST-PSB providers should more stringently adhere to admission criteria and only refer juveniles with more (specific) treatment needs, and consequently, a higher (sexual) recidivism risk.

The populations of SYC and FYC seem to better ‘fit’ the treatment objectives of an intensive treatment form (they show a relatively high treatment need). However, their risk levels did also not always support the need for risk reduction by (long-term) limitation in freedom of movement and maximum supervision offered.

Thus, the current allocation process used, despite the use of professional judgement of risks and needs, may lead to poor allocation decisions. As a result, juveniles may have received treatment that is too restrictive or intensive, which may in turn have adverse effects on development, behavior, and motivation (Bonta & Andrews 2007; Bonta et al., 2000; Cecile & Born, 2009; De Valk, Kuiper, Van der Helm, Maas, & Stams, 2016; Hendriks & Bijleveld, 2008). Taking into account the low prevalence of actual sexual recidivism (seven to 13% after 59 months at risk), and very low rates of persistence of sexual offending into adulthood by adolescents (Lobanov-Rostovsky, 2015; Lussier et al., 2012), these are consequences to be taken very seriously.

Allocation to treatment of sexually transgressive juveniles is to be based on specific knowledge, so the authors deem it important that a structured assessment of treatment needs and recidivism risk is administered and given more weight in the allocation to
mandatory forms of treatment. The validity of the J-SOAP-D is restricted, but, in the Netherlands, currently “as good as it gets” for treatment allocation purposes based on RNR principles; actuarial assessment certainly outperforms clinical judgments (Dawes, Faust, & Meehl, 1998; Hanson & Morton-Bourgon, 2009). Because of the restricted validity in predicting (sexual) recidivism, a focus on the assessed treatment needs (i.e. etiology or typology), might provide the most valid base for treatment allocation purposes.

For example, the practical implications of the assessed background characteristics (e.g., highest age, more prepubescent victims, more paraphilia, pervasive and personality disorders) and treatment needs (relatively high scores on sexual drive, intervention items, and community instability) in FYC appear to be that stimulating moral development and learning to cope with a deviant sexual interest should be specific targets. Lord (2016) recently reviewed literature regarding the treatment of mentally disordered sexual offenders (a target group similar to the one here identified in FYC) and strongly advocated the combined use of the RNR and the Good Lives model (GLM; Collie, Ward, Ayland, & West, 2007) principles, as mentally disordered sexual offenders are especially in need of an emphasis on approach goals, enhanced responsivity, and skills acquisition. GLM is a therapy model that focuses beyond risks and needs on creating meaning and life fulfilment. It has recently been suggested that the GLM therefore may more effectively motivate people to make a difficult change, as the change advocated is (also) meaningful to themselves, in contrast to making a change for which society’s reward is non-persecution (Andrews, Bonta, & Wormith, 2011; Ward, Yates, & Willis, 2012).

In contrast, the practical implications of background characteristics (e.g., more neglect, prior out of home placements, and disruptive disorders) and treatment needs assessed in SYC (relatively high scores on impulsive / antisocial behavior and community instability) may be that their program should focus less specifically on deviant sexual behavior and attitudes, and more on a broader set of (generalist) criminogenic needs. Improving personal adjustment (taking classes/qualifying for employment, anger management, and social skills training) and building (new) social networks could be key elements for effective treatment. The latter especially may currently receive too little attention in the specialist treatment offered by SYC.

**LIMITATIONS**

Some aspects of this study merit further reflection. Firstly, this study was limited by restrictions in the validity of the instrument used (J-SOAP-D). Insufficient research is available into the prediction of sexual recidivism by adolescents. Further research into the field of relapse into
sexual transgressions by youth is needed in order to obtain a fuller understanding of which juveniles persist and which desist. Secondly, the J-SOAP was retrospectively coded based on case file information, which could have restricted our knowledge of the cases. Allocation, however, usually takes place based on case file information, and case files were extensive. Both assessors of the J-SOAP-D were affiliated with one of the treatment modalities (SYC). Therefore, files were coded blind, and it was ensured that the coders had never had any therapeutic contact with the subjects coded, restricting the knowledge of the coders regarding the clients to the case files presented. Thirdly, the sample on which the research was performed was small (although it contained all participants from a period of two years). This meant that some analyses were underpowered. Fourthly, at the time of the study, MST-PSB had only recently been implemented as a new treatment in the Netherlands. This may have influenced allocation due to an unfamiliarity with the treatment and its referral criteria. In addition, it may have influenced the acceptance of less problematic cases for MST-PSB treatment due to an insufficient number of referrals. Finally, allocation results have been studied in the Netherlands, where a broad spectrum of possible responses to sexually transgressive behavior by adolescents (from criminal prosecution to voluntary enrolment in treatment) is available. In other countries (e.g., Belgium), the criminal prosecution of juveniles does not exist, and, in contrast, in some states of the United States, a firm judicial response (lifetime registration) of all types of sexual misbehavior, including that by adolescents, is common practice. In light of this, the results of this study should be interpreted with caution, as transferability may be limited by contextual differences. The administration of mandated specialized residential or community-based therapy, however, is general practice. Therefore, the results of this study may very well prove useful in any situation in which allocation to community-based versus residential treatment, based on RNR principles, occurs.

CONCLUSION

Allocation of sexually transgressive juveniles to treatment is a complex task. Juveniles assessed as low at (sexual) recidivism risk are not in need of intensive treatment, and may even suffer adverse effects if allocated as such. For juveniles assessed as medium to high at recidivism risk, with a high community stability, an out of home placement can be avoided by allocation to intensive -specialized- community-based treatment. Only juveniles in unsafe living situations may need to be placed out of home, and only assessed as ‘high at risk’ sexually (re)offensive juveniles may need to be treated in a maximum security setting, in order to protect society. At the same time, while allocating youth, it is important to keep in mind that assessment of sexual recidivism risk in juveniles might sooner overestimate, than underestimate real risks, since sexual recidivism is rare and most juvenile sexual transgressions remain adolescence limited.
Therefore, to improve the current allocation process, a structured (actuarial) assessment of treatment needs (i.e., etiology) of sexually transgressive juveniles should, in the opinion of the authors, be administered prior to allocation to specialized mandatory forms of treatment. Broad usage of one instrument might contribute to a more matched care for all sexually transgressive juveniles, whichever allocation pathway they follow.
Treatment Effect on Recidivism for Juveniles who have Sexually Offended: a Multilevel Meta-Analysis

Ter Beek, E.
Spruit, A.
Kuiper, C.H.Z.
Van der Rijken, R.E.A.
Hendriks, J.
Stams, G.J.J.M.

Journal of Abnormal Child Psychology.
Advance Online Publication. doi: 10.1007/s10802-017-0308-3
Chapter 3

**ABSTRACT**

The current study investigated the effect on recidivism of treatment aimed at juveniles who have sexually offended. It also assessed the potential moderating effect of type of recidivism, and several treatment, participant and study characteristics. In total, 14 published and unpublished primary studies, making use of a comparison group and reporting on official recidivism rates, were included in a multilevel meta-analysis. This resulted in the use of 77 effect sizes, and 1,726 participants. A three-level meta-analytic model was used to calculate the combined effect sizes (Cohen's $d$) and to perform moderator analyses. Study quality was assessed with the EPHPP Quality Assessment Tool for Quantitative Studies. A moderate effect size was found ($d = 0.37$), indicating that the treatment groups achieved an estimated relative reduction in recidivism of 20.5% as compared to comparison groups. However, after controlling for publication bias, a significant treatment effect was no longer found.

Type of recidivism did not moderate the effect of treatment, indicating that treatment groups were equally effective for all types of recidivism. Also, no moderating effects of participant or treatment characteristics were found. Regarding study characteristics, a shorter follow up time showed a trend for larger effect sizes, and the effect size calculation based on proportions yielded larger effect sizes than calculation via mean frequency of offending. Implications for future research and clinical practice are discussed.
INTRODUCTION

An interest in treatment for juveniles who have sexually offended was first sparked in the late 1970’s. Research on adults who had sexually offended showed that they had often started their abusing careers in adolescence (Groth, Longo, & McFaddin, 1982). In accordance with findings from research on adult offenders, it was assumed that dynamic risk factors explaining juvenile sex offending would differ from dynamic risk factors explaining non-sexual juvenile delinquency, and that specific treatment for juveniles who have sexually offended would be needed to target the dynamic risk factors for sexual recidivism (Barbaree & Marshall, 2006). Recent research, however, shows that juveniles who have sexually offended differ psychologically from adult sex offenders (Adjorlolo & Egbenya, 2016). Notably, in general, dynamic risk factors for juvenile sex-offense recidivism do not always differ from risk factors for juvenile general offense recidivism (Carpentier & Proulx, 2011; Christiansen & Vincent, 2013; Worling & Långström, 2003).

Sexual (re)offending by juveniles remains mainly adolescence limited (Beaudry-Cyr, Jennings, Zgoba, & Tewksbury, 2015; Loeber, Hoeve, Slot, & Van Der Laan, 2012; Lussier, Van Den Berg, Bijleveld, & Hendriks, 2012). In 2016, Caldwell conducted a meta-analysis, and found a weighted mean of 5% official recidivism by juveniles. This probably is an underestimation of the true recidivism rate due to the rarity of documented sexual offending (Wittebrood, 2006). Interestingly, the reoffending base rate has significantly declined to a weighted mean of 3% in the last 15 years. Improved quality of treatment was hypothesized to be the most likely explanation (Caldwell, 2016). Research on recidivism reduction through treatment of juveniles who have sexually offended, however, is highly diverse, differing in study design, outcome, treatment and participant characteristics, which is thought to affect study findings (Dopp, Borduin, & Brown, 2015; Hanson, Bourgon, Helmus, & Hodgson, 2009). The current study aims to synthesize the results of earlier studies by testing the hypothesis that specialized treatment for juveniles is effective.

Treatment Components

The Risk-Need-Responsivity (RNR) model (Bonta & Andrews, 2007), prominently states that treatment allocation and delivery should be guided by specific criminogenic treatment needs and recidivism risk levels, and should be delivered responsively (i.e., attuned to personal live goals, learning and motivational style) in order to be effective. As a consequence, aiming treatment at established risk factors for (sexual) reoffending is considered to be imperative (Ter Beek et al., 2017a).
Juveniles who generally display antisocial tendencies, also known as mixed offenders, who have also displayed other forms of offending behavior, seem most at risk for recidivism, including sexual reoffending (Drew, 2013; Hendriks & Bijleveld, 2005; Lawing, Frick, & Cruise, 2010; Van Wijk, Vreugdenhil, Van Horn, Vermeiren, & Doreleijers, 2007). Overall, juveniles who have sexually offended display a fair amount of non-sexual recidivism (Caldwell, 2016; Lobanov-Rostovsky, 2015). Specialized juvenile treatment may therefore need to address risk factors for general delinquency, such as immaturity, impulsiveness, and antisocial beliefs, to effectively reduce sexual as well as general recidivism risk. Seto and Lalumière (2010) and Fanniff and Kimonis (2014), comparing juveniles who have sexually offended with juveniles who have committed non-sexual offenses only, found both similarities (mainly their level of antisocial thoughts and behavior) and differences. Juveniles who had sexually offended showed more atypical sexual interests, more social and emotional problems, and more experiences of sexual victimization, abuse and neglect than did non-sex offenders. These findings suggest that general etiological theories of antisocial behavior might only partially explain juvenile sexual misconduct, and that specific or additional explanations remain relevant.

Christiansen and Vincent’s study (2013) of 39,249 juveniles, of whom 695 juveniles had offended sexually, showed prior non-sexual offending, non-contact sexual offenses, offending against a child, early age of onset, poor school performance, and non-school attendance to be predictive of sexual recidivism in juveniles. In 2011, Carpentier and Proulx examined 351 male adolescents who had sexually offended, and they found paternal abandonment, childhood sexual victimization, association with younger children, and having victimized a stranger to be associated with a higher risk for sexual recidivism. Prior to these studies, Worling and Långström (2003) conducted a literature review of juvenile sex offending, which provided empirical evidence for the following risk factors: sexual deviation (e.g., interest in prepubescent children or sexual violence), prior criminal sanctions for sexual assault(s), having two or more victims, victimization of strangers, lack of intimate peer relationships/social isolation, and incomplete offense-specific treatment. Notably, the dynamic (changeable) risk factors, such as sexual deviation, are of most interest for recidivism risk reduction through treatment (Beggs & Grace, 2011; Hanson & Morton-Bourgon, 2005; Pedersen, Rasmussen, & Elsass, 2010).

Although support is found for dynamic risk factors that need to be addressed during treatment, most of the empirically established risk factors for juvenile sexual recidivism (e.g., number of previous victims, prior offending behavior, having offended against a child or a stranger, early age of onset) are considered static (unchangeable). Successful treatment however, may also focus on reduction of negative consequences of static risk
factors (for example, by means of trauma treatment to alleviate the consequences of earlier victimization). Also, non-criminogenic (mental health) treatment needs should be addressed (Lord, 2016).

Studies into treatment needs of juveniles who have sexually offended, have substantiated the existence of several psychological typologies. Studies into specific mental illnesses, executive functioning, and personality profiles have, for example, concluded that juveniles with child victims display different treatment needs than juvenile peer abusers (Drew, 2013; Glowacz & Born, 2012; Hart-Kerkhoffs, Doreleijers, Jansen, Van Wijk, & Bullens, 2009; Hendriks, 2006; Kjellgren, Wassberg, Carlberg, Långström, & Svedin, 2006). Also, Butler and Seto (2002), and Hissel, Bijleveld, Hendriks, Jansen, and Collot d’Escursy-Koenigs (2006) have found sex-only offenders to differ from mixed offenders (juveniles having also committed non-sexual offenses). Hendriks (2006) also differentiated group offenders from solo offenders. The diversity in the population of juveniles who have sexually offended has become increasingly clear, especially because to date no psychological typology has been found to represent a subgroup at increased risk for sexual recidivism (Cale, Smallbone, Rayment-McHugh, and Dowling, 2016; Lussier, et al., 2012).

Effectiveness of Treatment for Juveniles who have Sexually Offended

Treatments for juveniles who have sexually offended make use of different treatment strategies (i.e., wilderness therapy, cognitive and/or behavioral based group therapy, cognitive and/or behavioral based individual therapy, family therapy and psychoeducation) and are delivered in both residential and community settings (Ryan, Leversee, & Lane, 2010; Veneziano & Veneziano, 2002). Meta-analyses, predominantly including studies that combine adult and juvenile samples, have shown cognitive behavioral based treatments and multisystemic therapy to be most effective in reducing recidivism (Dopp, Borduin, & Brown, 2015; Hanson, Bourgon, Helmus, & Hodgson, 2009; Lösel & Schmucker, 2005; Schmucker & Lösel, 2015; Walker, McGovern, Poey, & Otis, 2004). Treatment types that incorporate these standards are, therefore, generally considered established treatment.

To date, only two meta-analyses have been performed that focus specifically on treatment for juveniles. Walker and colleagues (2004) performed a meta-analysis of ten published and unpublished studies; two studies with and eight without a comparison group. Reitzel and Carbonell (2006) based their findings on nine studies, all using a comparison or control group, and including five unpublished studies. These meta-analyses demonstrated a moderate effect of treatment on sexual recidivism by juveniles. Moderator analyses showed that outcomes measured via official recidivism data (rearrests or reconvictions) resulted in smaller effect sizes compared to measurement via self-report or penile plethysmography (i.e., the measurement of physical arousal to deviant stimuli). Furthermore, random
assignment to treatment and comparison groups yielded larger effect sizes. The hypotheses that cognitive behaviorally based treatments are more effective than other treatment types, was not substantiated in this meta-analysis (Reitzel & Carbonell, 2006).

Reitzel and Carbonell (2006), as well as Walker and colleagues (2004), assessed the effects of treatment on sexual recidivism only, and based their findings on a restricted number of studies. In the last decade more and more robust studies (making use of a comparison or control group and examining the effectiveness of treatment for multiple types of recidivism) have been conducted. In addition, multilevel statistical techniques have become more refined, making it possible to analyze moderators with a restricted number of studies, accounting for both within and between study variability in effect sizes. A multilevel approach enables comprehensive moderator analyses to examine the influence of study, sample, and intervention characteristics (Van den Noortgate, López-López, Marin-Martínez, & Sánchez-Meca, 2013; Van den Noortgate & Onghena, 2003). This may be useful in the complex field of juvenile sex offending, where multiple causes may contribute to recidivism. Furthermore, multilevel meta-analytical techniques enable the use of all available effect sizes in the primary studies, thus preserving all information and generating maximum statistical power (Assink et al., 2015). This is especially important when examining the results of studies with small samples, as is the case with many of the primary studies in this field.

**STUDY AIM**

The aim of this study is to review contemporary research on the effect on officially measured recidivism (re-arrest or reconviction) of treatment for juveniles who have sexually offended. In addition, the potential moderating effects of participant, treatment, and study characteristics were investigated. This provided an opportunity to detect factors that may influence the effect of treatment on the reduction of recidivism.

**METHOD**

To assess the effect of treatment on recidivism, and the factors moderating this effect, a multilevel meta-analysis was carried out. The term meta-analysis refers to a stepwise procedure and a set of statistical techniques, combining results of independent primary studies into effect sizes, so that overall conclusions can be drawn. An important requirement for traditional univariate meta-analytic approaches is that no dependency between effect sizes is allowed, so that only one effect size per primary study can be included. By stepping away from the traditional univariate approach, it becomes possible to deal with dependency of effect sizes, so all information can be preserved and a maximum of statistical power is
Treatment effect on Recidivism

achieved. In this study, applying a three-level structure statistically considers the three different variance components distributed over three levels: differences among all effect sizes or random sampling error (level 1), differences in effect sizes within studies (level 2), and differences in effect sizes between studies (level 3). If there is evidence for heterogeneity in effect sizes, moderator analyses can be conducted to test variables that may explain within-study or between-study heterogeneity. For these analyses, the three-level random effects model can be extended with study and effect size characteristics, making the model a three-level mixed effects model (Assink & Wibbelink, 2016).

Inclusion Criteria

Multiple inclusion criteria were formulated to select the studies. First, the treatment condition had to be aimed at treating sexually offending behavior. Second, the study sample had to exclusively contain juveniles. Therefore, the mean age of the researched group had to lie between 12 and 18 years of age, and/or the study had to specifically report on juveniles or adolescents. Third, the studies had to report treatment and comparison group recidivism rates via officially obtained data for sexual and/or other types of recidivism. And finally, the comparison group had to comprise juveniles who have sexually offended as to ensure basic comparability between groups.

Selection of Studies

All studies published before February 2016 that met the inclusion criteria were to be included in the current meta-analysis. Firstly, several electronic databases were searched, including Campbell library, PubMed, OVID (Medline, PsycINFO, ERIC), and Proquest (Sociological Abstracts, Social Services Abstracts, Proquest Dissertations). Secondly, Google Scholar was searched. The following English search string was used: (sex* offender) AND (juvenile OR adolescent) AND (treatment OR therapy) AND (effect* OR efficacy) AND (recidivism). Finally, the references of other meta-analyses and reviews were checked for eligible studies, and authors of non-published work were contacted. Regrettably the contacted authors did not respond, so a few non-published studies could not be included. A flow chart is presented as Figure 1 in the results section.

The initial search and screening resulted in 23 studies that met the basic criterion of examining the effectiveness of an intervention for juveniles who have sexually offended with the use of a comparison group. After exclusion, mainly because of the use of other outcome variables (see also Table 2), 14 studies remained, with 77 effect sizes, 1,726 participants, and 13 independent samples. Table 1 presents the study characteristics of the included studies. Table 2 specifies the excluded studies, and our reasons for excluding them in italic.
Chapter 3

Identification
443 records through database search

Screening
453 records screened

Eligibility
23 full text articles assessed

Included
14 studies included in quantitative synthesis

14 additional records through other sources, 4 not retrievable (unpub.)
430 records excluded

9 full text articles excluded:
- 5 outcome variable (non-official recidivism / other)
- 1 non-treatment (registration)
- 1 non-eligible experimental group (NJSO)
- 1 non-eligible comparison group (NJSO)
- 1 age of respondents (-12)

Figure 1. Flowchart of study selection

Coding the Studies

The dependent variable in this meta-analysis was recidivism. Three types of recidivism were distinguished: any recidivism, sexual recidivism, and non-sexual recidivism. In all studies, recidivism was measured by means of official data as to ensure comparability among the effect sizes of the primary studies, because the type of measurement used may influence the effect measured (Reitzel & Carbonell, 2006).

The independent variable was the treatment offered. Type of recidivism, participant, treatment, and study characteristics were coded to assess whether treatment effects varied among the possible moderator variables. In order to reduce the problem of multiple testing (Tabachnik & Fidell, 2013), only moderators of possible theoretical importance were used. Studies with overlapping samples (Worling & Långström, 2000; Worling, Littlejohn, & Bookalam, 2010) were coded with corresponding study identification numbers.
Table 1. Characteristics of Included Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Study characteristics</th>
<th>Sample characteristics</th>
<th>Treatment characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First auth. &amp; year of pub.</td>
<td>Treat. coh.</td>
<td>N</td>
</tr>
<tr>
<td>Barlow 1998</td>
<td>1995</td>
<td>44</td>
<td>2 (.328) N Weak QE Y</td>
</tr>
<tr>
<td>Borduin 1990</td>
<td>1983-1985</td>
<td>16</td>
<td>6 (.1130) Y Weak RCT N</td>
</tr>
<tr>
<td>Borduin 2009</td>
<td>1990-2001</td>
<td>48</td>
<td>5 (.915) Y Mod RCT N</td>
</tr>
<tr>
<td>Gillis 2010</td>
<td>1990-2005</td>
<td>190</td>
<td>6 (.341) Y Mod QE Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holly 2000</td>
<td>1990-1998</td>
<td>64</td>
<td>14 (.383) N Mod QE Y</td>
</tr>
<tr>
<td>Johnson 2000</td>
<td>2000-2007</td>
<td>78</td>
<td>4 (.267) N Strong QE Y</td>
</tr>
<tr>
<td>Letourneau 2013</td>
<td>2004-2006</td>
<td>131</td>
<td>2 (.051) Y Strong RCT N</td>
</tr>
<tr>
<td>Schram 1991</td>
<td>1984-1991</td>
<td>196</td>
<td>6 (.055) N Weak QE Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worling 2000</td>
<td>1987-1995</td>
<td>148</td>
<td>12 (.575) Y Mod QE N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worling 2010</td>
<td>1987-1995</td>
<td>148</td>
<td>6 (.352) Y Mod QE N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Treat. coh. = Years in which treatment was delivered; N = number of participants; # r (M) = number of effect sizes (mean d); Pb. = published in peer reviewed article yes/no; Study Quality = strong/moderate/weak according to EP-PP quality assessment tool; Design = RCT (randomized controlled trial) or QE (quasi experimental); IA = Independent author: yes/no; Type of recidivism measured: 1 = all recidivism, 2 = sexual, 3 = non-sexual; Age = mean age of sample; % Male = percentage of males in sample; % Caucasian = percentage of Caucasian ethnicity; % Contact offense = percentage contact offense; % Child molester = percentage child molestation; Setting = outpatient, community based, residential or a mix; EXP type = Type of experimental intervention; Experimental Condition = name of experimental intervention; Comparison or Control condition = Type of control intervention(s); # = studies using the same participants; * = treatment comparison group; * = no treatment control group. The allocation of control group status was based on available information, sometimes control group juveniles may have still have received alternate forms of treatment elsewhere.
Table 2. Excluded Studies and Reasons for Exclusion

<table>
<thead>
<tr>
<th>Study</th>
<th>First author &amp; year of publication</th>
<th>Treat. cohort</th>
<th>N</th>
<th>Setting</th>
<th>Outcome measure</th>
<th>Experimental treatment</th>
<th>Comparison or Control condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Carpentier 2006</td>
<td>1992-1995</td>
<td>135</td>
<td>Residential</td>
<td>Psychosocial Functioning &amp; Official Recidivism Rates</td>
<td>Play therapy (5 to 12 year olds)*</td>
<td>CBT (5 to 12 year olds)**</td>
</tr>
<tr>
<td>Kahn 1991</td>
<td>1984</td>
<td>221</td>
<td>Mixed</td>
<td>Official Recidivism Rates (total group)*</td>
<td>8 Outpatient programs</td>
<td>2 Residential programs*</td>
<td></td>
</tr>
<tr>
<td>Seabloom 2003</td>
<td>1977-1986*</td>
<td>122</td>
<td>Outpatient</td>
<td>Official Recidivism Rates</td>
<td>Personal Social Awareness Program completers</td>
<td>Personal Social Awareness Program withdrawers*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Letourneau 2008</td>
<td>1995-2005</td>
<td>222</td>
<td>-</td>
<td>Official Recidivism Rates</td>
<td>Registration*</td>
<td>Non-Registration*</td>
</tr>
<tr>
<td></td>
<td>Letourneau 2009</td>
<td>2004-2006</td>
<td>127</td>
<td>Community</td>
<td>Psychosocial Functioning*</td>
<td>JSO Multi Systemic Therapy</td>
<td>JSO Usual Community Services*</td>
</tr>
<tr>
<td>Henggeler 2009</td>
<td>2004-2006</td>
<td>127</td>
<td>Community</td>
<td>Psychosocial Functioning*</td>
<td>JSO Multi Systemic Therapy</td>
<td>JSO Usual Community Services*</td>
<td></td>
</tr>
<tr>
<td>Van Outsem 2009</td>
<td>-</td>
<td>113</td>
<td>Outpatient</td>
<td>Psychological Functioning*</td>
<td>Treatment completers</td>
<td>Treatment non-completers and ‘normal’ controls*</td>
<td></td>
</tr>
<tr>
<td>Weinrott 1997</td>
<td>NR</td>
<td>69</td>
<td>Outpatient</td>
<td>Phallometric measurement and Selfreport*</td>
<td>Vicarious Sensitization</td>
<td>Cognitive Therapy*</td>
<td></td>
</tr>
</tbody>
</table>

Note: JSO = Juvenile Sex offender; JNSO = Juvenile Non Sex Offender; CBT = Cognitive Behavioral Therapy.

* The characteristic in italic font specifies reason(s) for exclusion. Mostly the outcome variable was other than a measure of recidivism (Letourneau, 2009, Henggeler, Van Outsem, and Weinrott). In the Brannon study a comparison group of non-sex offenders was used. In the Carpentier study the respondents were too young to be included, in the Kahn study the official recidivism rates were not available for the experimental and control group, and in the Letourneau 2008 study registration was researched which, to our judgement, is not a form of treatment. The Seabloom study included transgressive acts (i.e., homosexuality, transvestitism) that were then considered to be a sexual transgression, but are now considered a variation in normal human sexuality, so their sample was deemed not comparable.
Firstly, outcome characteristics were coded as the type of recidivism reported; recidivism in general (including all types of recidivism), sexual recidivism, or non-sexual recidivism.

The following participant characteristics were coded. The criminal status of the clients was coded as all adjudicated or as also suspected of a sex crime, as not all studies only included formally adjudicated offenders. The cultural background of the juveniles was coded as the percentage of Caucasians in the researched group. Furthermore, the percentage of juveniles with child victims (< 12 years of age and ≥ 5 years younger than their assailant) was coded, as was the percentage of contact offenders (physical sexual offending, such as rape or sexual assault; not including, for example, exhibitionism). In most reports information on the estimated recidivism risk level of the respondents was available (groups were designated as relatively high at risk or low at risk), so also the mean estimated recidivism risk level of clients was coded as high or low. Finally, the living condition of participants was coded as living at home versus living in a residential facility or mixed (both types of living situation present in the researched group).

Several treatment characteristics were coded: firstly, whether the type of treatment was cognitive behavioral, systemic, or another form of (non-established) treatment as usual. Mean treatment duration was coded in months. Finally, the exclusion of clients with a lower IQ was coded as yes or no, for this may influence learnability of the researched group.

As for study characteristics, we coded whether the authors were independent researchers or whether they were involved in the development or implementation of the intervention. It was also coded whether the study was peer reviewed. Further, the design of the study (randomized controlled trial versus quasi-experimental) was coded. The type of comparison treatment used was coded as treatment as usual (TAU) or established treatment (incorporating cognitive behavioral treatment and/or systemic therapy). The comparability of the treatment and comparison group(s) was coded as high (including statistically controlling for any potential influential differences found, for example, differences in group recidivism risk levels) or low (when significant differences were reported between groups and this was not controlled for in results). The mean length of follow up was coded in months. The type of effect size was coded as proportion based (the difference in proportion of reoffenders in the treatment and comparison group) or based on mean number of reoffenses in each group. Type of recidivism measured was coded as re-arrest or reconviction. Lastly, study quality was coded by use of the EPHPP Quality Assessment Tool for Quantitative Studies (http://www.ephpp.ca/tools.html). This tool assesses the quality of a study as weak, moderate or strong, providing a comprehensive and structured assessment of the concept of study quality (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012). The EPHPP was developed by the Effective Public Health Practice Project in Canada. It is a generic tool used
to evaluate a variety of intervention study designs. It has been judged suitable to be used in systematic reviews of effectiveness (Deeks et al., 2003) and has been reported to have content and construct validity (Jackson & Waters, 2005; Thomas, Ciliska, Dobbins, & Micucci, 2004). The tool assesses six domains: selection bias, study design (including appropriateness of the design), confounders, blinding, data collection methods, and withdrawals and dropouts, see Table 1 for the results of the assessment.

The first author of this article coded the included studies according to the suggestions of Lipsey and Wilson (2001). All studies were double coded by the second author. Following the guidelines by Shrout (1989), for the continuous variables single ICCs were calculated for the two-way random effects model, with absolute agreement as a criterion. On all 18 variables single measures showed an ICC of > .99 (i.e., substantial reliability). For 23 categorical variables kappa was calculated (Landis & Koch, 1977). Two variables reached substantial interrater reliability (kappa = .61-.80), all others reached almost perfect interrater reliability (kappa > .81).

Calculations and Analyses

Effect sizes were transformed into Cohen’s $d$ by using the calculator of Wilson (2013) and formulas of Lipsey and Wilson (2001). Most $d$-values were calculated based on reported means and percentages of recidivism. A positive effect size indicated that the treatment group benefited more than the comparison group, whereas a negative effect size indicated that the comparison group benefitted more than the treatment group. If a study only mentioned that an effect was not significant (as was the case in 9% of all effect sizes), the effect size was coded as zero (Lipsey & Wilson, 2001). The continuous variables (percentage immigrants, percentage with child victims, percentage contact, percentage mixed offenders, and mean treatment length) were centered around their mean, and all other (categorical) variables were recoded into dummy variables. We checked for the presence of extreme outliers using box plots (Tabachnik & Fidell, 2013); no extreme outliers were identified. Standard errors were estimated using formulas of Lipsey and Wilson (2001).

In all studies, we were able to calculate more than one effect size. Most studies reported on multiple outcome variables (i.e., all types of recidivism, sex only recidivism, non-sex recidivism, and sometimes violent non-sexual recidivism and non-violent non-sexual recidivism). Effect sizes from the same study may prove more alike than effect sizes from different studies. Therefore, the assumption of statistical independency, which underlies classical meta-analytic strategies, was violated (Hox, 2002; Lipsey & Wilson, 2001). In line with recently conducted meta-analyses, we applied a multilevel approach in order to deal with the interdependency of effect sizes (Assink et al., 2015; Houben, Van Den Noortgate, &
Kuppens, 2015; Spruit, Assink, Van Vugt, Van Der Put, & Stams, 2016; Weisz, et.al, 2013). The multilevel approach accounts for the hierarchical structure of the data in which effect sizes are nested within the studies (Van den Noortgate & Onghena, 2003).

We used a three-level meta-analytic model to calculate the combined effect sizes and to perform moderator analyses. Three sources of variance were modelled, including the sampling variance for the observed effect sizes (level 1), the variance between effect sizes from the same study (level 2), and the variance between the studies (level 3) (Cheung, 2014; Van den Noortgate et al, 2013; Wibbelink & Assink, 2015). The sampling variance of observed effect sizes (level 1) was estimated by using the formula of Cheung (2014). Log-likelihood-ratio-tests were performed to compare the deviance of the full model to the deviance of the models excluding one of the variance parameters, making it possible to determine whether significant variance is present at the second and third level (Wibbelink & Assink, 2015). Significant variance at level 2 or 3 indicates a heterogeneous effect size distribution, meaning that the effect sizes cannot be treated as estimates of a common effect size. In that case, we proceed to moderator analyses, because the differences between the effect sizes may be explained by outcome, study, sample, and/or intervention characteristics. Moderator analyses were only performed when each category of the potential moderator was filled with at least three studies. As a result, Study Quality was collapsed into a dichotomous variable (weak versus medium to strong) because only two studies were considered to be of strong quality. Exclusion of IQ could not be tested because only two studies reported on the exclusion of lower IQ respondents.

The multilevel meta-analysis was conducted in R (version 3.2.0) with the metafor-package, using a multilevel random effects model (Viechtbauer, 2010; Wibbelink & Assink, 2015). The restricted maximum likelihood estimate was used to estimate all model parameters, and the Knapp and Hartung-method (2003) was used for testing individual regression coefficients of the meta-analytic models and for calculating the corresponding confidence intervals (see also Assink et al., 2015; Houben et al., 2015; Spruit et al, 2016; Wibbelink & Assink, 2015).

**Publication Bias**

In systematic reviews, the aim is to include all studies previously conducted that meet the inclusion criteria (Lipsey & Wilson, 2001). However, a common problem is that studies may not have been published due to non-significant or unfavorable findings, and, therefore, are difficult to locate. Not including these studies may lead to an overestimation of the true effect size, the so called “publication bias” (Rosenthal, 1979). In order to check the presence of publication bias in our meta-analysis, we performed a trim and fill procedure (Duval & Tweedie, 2000) by drawing a trim-and-fill plot in R (Version 3.2.0) using the function “trimfill”
of the metafor package (Viechtbauer, 2010). We tested whether effect sizes were missing on the left side of the distribution, since publication bias would only be likely to occur in case of non-significant or unfavorable (i.e., negative) results.

RESULTS

Overall, a significant, small to moderate effect, \( d = 0.37, p < .001 \), of treatment on recidivism was found, indicating that the treatment groups achieved an estimated reduction in recidivism of 20.5% as compared to comparison groups. The likelihood ratio test comparing models with and without between-study variance (level 3) showed that significant variance was present at the between-study level, \( \sigma^2_{\text{level 3}} = 0.11, \chi^2(1) = 37.12; p < .000 \). The variance between the effect sizes within studies (level 2) was not significant, \( \sigma^2_{\text{level 2}} = 0.01, \chi^2(1) = 3.32; p = .068 \). About 24% of the total effect size variance was accounted for by the sampling variance (level 1), 7% for the variance between effect sizes within studies (level 2), and 69% for the variance between studies (level 3). Because the variance on the third level was significant, we proceeded to moderator analyses to assess factors that could possibly explain variance in treatment effects (see Table 3).

As presented in the last column of Table 3, only the calculation type of effect size moderated the effect of treatment on recidivism in adolescents who have sexually offended. Stronger treatment effects were found for effect sizes that were proportion-based compared to effect sizes based on the mean number of re-offenses. The proportion of recidivism was significantly lower in the treatment groups compared to the comparison groups, but the mean number of offenses for every reoffender in the two groups did not significantly differ. Furthermore, a moderating trend was found for the follow up period, indicating that stronger treatment effects were found in shorter follow up periods. None of the participant characteristics and none of the treatment characteristics moderated the effect of treatment on recidivism in adolescents who have sexually offended.

The trim-and-fill plot imputed estimations of effect sizes of missing studies, indicating the presence of publication bias (see Figure 2). We included the imputed estimations (the open circles) and performed the meta-analysis again to compute an overall effect size that takes the influence of publication bias into account (Duval & Tweedie, 2000). After controlling for publication bias, no significant treatment effect was found (\( d = 0.15, p = .176 \)).
DISCUSSION

A multilevel meta-analysis was performed to assess the strength of the effect on the reduction of recidivism of specific treatment for juveniles who have sexually offended, and to assess what variables have a moderating influence on the effect measured. An overall significant and small to moderate effect ($d = 0.37$) was found, indicating treatment to be more effective in reducing recidivism than comparison or control conditions, a finding similar to that of previous meta-analyses (Reitzel & Carbonell, 2006; Walker et al., 2004). However, indications of publication bias were present, and after statistically controlling for this bias, the effect of treatment on recidivism was no longer significant. Studies with less favorable outcomes may indeed not have been published, and the small to moderate effect found may therefore represent an overestimation of the true overall effect size. Thus, in the current meta-analysis, after correcting for publication bias, an effect of treatment on recidivism, as compared to comparison or control conditions, was not demonstrated. The treatment conditions in the primary studies, however, were mainly compared to other forms of treatment ($N=10$), or to groups of assessment-only juveniles, who may have received...
**Table 3.** Overall Results and Moderator Effects of the Relation between Treatment and Recidivism

<table>
<thead>
<tr>
<th>Moderator variables</th>
<th># Studies</th>
<th># ES</th>
<th>$\beta_0$ (mean $d$)</th>
<th>$t$</th>
<th>$r$</th>
<th>F(df1,df2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall relation</td>
<td>13</td>
<td>77</td>
<td>0.367</td>
<td>3.713***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After trim and fill</td>
<td>18</td>
<td>99</td>
<td>0.154</td>
<td>1.364</td>
<td>(p=.176)</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Recidivism</td>
<td>13</td>
<td>77</td>
<td>F(2,74) = 0.543</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any offenses (RC)</td>
<td>13</td>
<td>29</td>
<td>0.401</td>
<td>3.764***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex offenses</td>
<td>11</td>
<td>19</td>
<td>0.348</td>
<td>3.100**</td>
<td>-0.054</td>
<td>-0.735</td>
</tr>
<tr>
<td>Non-sex offenses</td>
<td>9</td>
<td>28</td>
<td>0.332</td>
<td>3.012**</td>
<td>-0.069</td>
<td>-0.999</td>
</tr>
<tr>
<td><strong>Participant characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criminal Status</td>
<td>13</td>
<td>77</td>
<td>F(1,75) = 1.135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All convicted (RC)</td>
<td>9</td>
<td>58</td>
<td>0.299</td>
<td>2.506*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Also suspects</td>
<td>4</td>
<td>19</td>
<td>0.533</td>
<td>2.880**</td>
<td>0.235</td>
<td>1.065</td>
</tr>
<tr>
<td>Percentage Caucasian</td>
<td>11</td>
<td>46</td>
<td>0.333</td>
<td>2.694*</td>
<td>0.406</td>
<td>0.473</td>
</tr>
<tr>
<td>Proportion Child Molestation</td>
<td>8</td>
<td>44</td>
<td>0.337</td>
<td>2.105*</td>
<td>0.395</td>
<td>0.451</td>
</tr>
<tr>
<td>Proportion Contact Offenses</td>
<td>10</td>
<td>64</td>
<td>0.332</td>
<td>2.554*</td>
<td>0.084</td>
<td>0.185</td>
</tr>
<tr>
<td>Recidivism Risk</td>
<td>12</td>
<td>72</td>
<td>F(1,70) = 1.162</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High risk (RC)</td>
<td>8</td>
<td>49</td>
<td>0.423</td>
<td>3.284**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>4</td>
<td>23</td>
<td>0.186</td>
<td>1.044</td>
<td>-0.237</td>
<td>-1.078</td>
</tr>
<tr>
<td>Living Conditions</td>
<td>13</td>
<td>77</td>
<td>F(2,74) = 0.085</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential (RC)</td>
<td>3</td>
<td>12</td>
<td>0.327</td>
<td>1.432</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living at home</td>
<td>5</td>
<td>31</td>
<td>0.427</td>
<td>2.393*</td>
<td>0.100</td>
<td>0.345</td>
</tr>
<tr>
<td>Mixed</td>
<td>5</td>
<td>34</td>
<td>0.339</td>
<td>1.971*</td>
<td>0.012</td>
<td>0.041</td>
</tr>
<tr>
<td><strong>Treatment characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Exp. treatment</td>
<td>13</td>
<td>77</td>
<td>F(2,74) = 0.372</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAU (RC)</td>
<td>3</td>
<td>14</td>
<td>0.213</td>
<td>0.993</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBT</td>
<td>6</td>
<td>46</td>
<td>0.438</td>
<td>2.864**</td>
<td>0.226</td>
<td>0.857</td>
</tr>
<tr>
<td>Systemic</td>
<td>4</td>
<td>17</td>
<td>0.384</td>
<td>1.961*</td>
<td>0.171</td>
<td>0.590</td>
</tr>
<tr>
<td>Duration EXP</td>
<td>9</td>
<td>49</td>
<td>0.400*</td>
<td>2.554</td>
<td>-0.002</td>
<td>-0.116</td>
</tr>
<tr>
<td><strong>Study characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>13</td>
<td>77</td>
<td>F(1,75) = 1.262</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent (RC)</td>
<td>6</td>
<td>38</td>
<td>0.488</td>
<td>3.342**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent</td>
<td>7</td>
<td>39</td>
<td>0.265</td>
<td>1.969*</td>
<td>0.223</td>
<td>1.123</td>
</tr>
<tr>
<td>Peer Reviewed</td>
<td>13</td>
<td>77</td>
<td>F(1,75) = 2.456</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (RC)</td>
<td>9</td>
<td>51</td>
<td>0.463</td>
<td>4.125***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderator variables</td>
<td># Studies</td>
<td># ES</td>
<td>$\beta_0$ (mean $d$)</td>
<td>$t_0$</td>
<td>$\beta_1$</td>
<td>$t_1$</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------</td>
<td>-----</td>
<td>----------------------</td>
<td>-------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>No</td>
<td>4</td>
<td>26</td>
<td>0.144</td>
<td>0.850</td>
<td>-0.319</td>
<td>-1.567</td>
</tr>
<tr>
<td>Study Design</td>
<td>13</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomized (RC)</td>
<td>3</td>
<td>13</td>
<td>0.653</td>
<td>3.131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quasi-experimental</td>
<td>10</td>
<td>64</td>
<td>0.292</td>
<td>2.739</td>
<td>-0.361</td>
<td>-1.542</td>
</tr>
<tr>
<td>Type of Control Treatment</td>
<td>12</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment as usual (RC)</td>
<td>5</td>
<td>23</td>
<td>0.423</td>
<td>2.895</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Established treatment</td>
<td>7</td>
<td>34</td>
<td>0.389</td>
<td>3.057</td>
<td>-0.035</td>
<td>-0.233</td>
</tr>
<tr>
<td>Comparability of Groups</td>
<td>10</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High (RC)</td>
<td>5</td>
<td>35</td>
<td>0.374</td>
<td>2.867</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>5</td>
<td>21</td>
<td>0.351</td>
<td>0.162</td>
<td>-0.023</td>
<td>-0.162</td>
</tr>
<tr>
<td>Length Follow Up</td>
<td>10</td>
<td>68</td>
<td>0.336</td>
<td>2.460</td>
<td>-0.002</td>
<td>-1.919</td>
</tr>
<tr>
<td>Type effect</td>
<td>13</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion (RC)</td>
<td>13</td>
<td>68</td>
<td>0.396</td>
<td>3.728</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>4</td>
<td>9</td>
<td>0.162</td>
<td>1.167</td>
<td>-0.234</td>
<td>-2.291</td>
</tr>
<tr>
<td>Recidivism Measure</td>
<td>13</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reconviction (RC)</td>
<td>3</td>
<td>20</td>
<td>0.178</td>
<td>0.841</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rearrest</td>
<td>10</td>
<td>57</td>
<td>0.421</td>
<td>3.731</td>
<td>0.243</td>
<td>1.011</td>
</tr>
<tr>
<td>Study Quality</td>
<td>13</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong/Moderate (RC)</td>
<td>7</td>
<td>54</td>
<td>0.380</td>
<td>2.739</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td>6</td>
<td>23</td>
<td>0.353</td>
<td>2.270</td>
<td>-0.028</td>
<td>-0.134</td>
</tr>
</tbody>
</table>

Note: # studies = number of independent studies; # ES = number of effect sizes; $t_0 = $ difference in mean $r$ with zero; $t_1 = $ difference in mean $d$ with reference category; mean $d = $ mean effect size ($d$); $F(df_1, df_2)$ = omnibus test; (RC) = reference category. * = trend, significant at a 0.1 level, ** = significant at a 0.05 level, *** = significant at a 0.01 level, **** = significant at a 0.001 level.
treatment elsewhere ($N = 5$). Therefore, the difference in effect on recidivism between the treatment conditions and the comparison conditions might have been reduced. Comparing treatment to a control condition (a group that receives no treatment at all), is considered unethical in this population as it requires withholding treatment. Such research, therefore, will remain scarce. Few studies ($N=3$) made use of drop out control groups. These studies, in contrast, may have resulted in an overestimation of the effect of treatment because drop outs are at higher risk for reoffending (Hendriks & Bijleveld, 2005).

Significant variance was present at the between-study level. This indicates that the effect size distribution was heterogeneous, and that moderating variables might explain differences in strength of the effect sizes. Therefore, moderator analyses were performed.

The type of recidivism measured did not moderate the effects significantly; on all types of recidivism a small to moderate effect was achieved. This result could signify that, in juveniles, while treating sexual offending behavior, other offending behavior is also treated. This effect has also previously been established by Lösel and Schmucker (2005) and Schmucker and Lösel (2015). Treatment usually entails a confrontation with the sexual offending behavior, thoughts, and feelings prior to, during, and after the offense, the development of victim empathy, the development of anger management and stress coping skills, social skills training, treatment of the juveniles own victimization experiences, and relapse prevention strategies (Becker & Hunter, 1997; Becker & Johnson, 2001). Moreover, contemporary treatments have been developed to take into account broader issues. The Good Lives Model, for example, was presented as an addition to the relatively risk focused RNR model (Lord, 2016). Contextual treatment forms, such as Multisystemic Therapy-Problem Sexual Behavior, are explicitly holistic in nature (Swenson, Henggeler, Taylor, & Addison, 2005). Thus, risk factors for juvenile delinquency in general also seem to be more and more addressed via treatment aimed at sex offending. This is considered expedient, because while sexual recidivism amongst juveniles is rare, other types of recidivism are relatively frequent (Caldwell, 2010; Caldwell, 2016; Lobanov-Rostovsky, 2015).

None of the tested participant characteristics moderated the effect of treatment on recidivism in juveniles. Thus, the currently coded subgroups within the sample did not account for differences in effects on reducing recidivism between the treatment and comparison or control groups. Notably, few studies defined their participants more specifically than juveniles who have sexually offended, and as a result, in the included studies the heterogeneous group of juveniles who commit sexual offenses was presented as relatively homogenous. This hampered research into moderating effects of participant characteristics. Many studies on specific mental illnesses, traits, executive functioning, and personality profiles of adolescents who have sexually offended have been conducted.
Treatment effect on Recidivism

(Adjorlolo & Egbenya, 2016; Glowacz & Born, 2012; Hart-Kerhoffs et al., 2009; Lawing et al., 2010; Margari et al., 2015; Purcel, 2010; Seto & Lalumière, 2010). None of the studies included in the current meta-analysis, however, reported on psychological profiles or the specific treatment needs juveniles presented at admission that deemed them eligible or in need of specific treatment (other than their type of offending behavior). Also, several typologies are commonly in use: non-contact versus contact offenders, sex-only offenders versus mixed or sex-plus offenders, peer abusers versus child abusers, and group offenders versus solo offenders (Butler & Seto, 2002; Drew, 2013; Hendriks, 2006; Hissel et al., 2006; Kjellgren et al., 2006), and are described as in need of different treatment approaches. Of these typologies, only two (victim age and type of offending behavior) could be tested in this study. A more extensive moderator analysis could have been performed if studies had reported on treatment needs or typologies of their respondents.

Further, none of the coded treatment characteristics moderated the effect of treatment. The treatment conditions were grouped as predominantly cognitive behavioral based (e.g., SAFE-T, TBASOP), systemic (i.e., Multi Systemic Therapy and Functional Family Therapy), or as treatment as usual for juveniles who have sexually offended (TAU). The first two categories are established treatments, the third category contained a compilation of non-established treatment forms (i.e., adventure or wilderness therapy, individual treatment, or family treatment). The established treatments were expected to achieve better results in reducing recidivism in juveniles based on previous research (Hanson et al., 2009; Lösel & Schmucker, 2005; Schmucker & Lösel, 2015; Walker et al., 2004).

In study characteristics, one moderator significantly explained variance (calculation type of effect size), while for another (follow up time) a moderating trend was found, indicating stronger treatment effects in shorter follow up periods. These results are in line with expectations. Calculation of Cohen’s $d$ via proportions (the relative number of recidivists per group) yielded a more pronounced effect (difference measured) than calculation via mean number of re-offenses (how many times recidivists reoffended per group). This indicates that specialist treatment showed a stronger reduction in the number of juveniles reoffending than comparison conditions did, but that reoffenders recidivated as many times in both groups. Regarding follow up time, most recidivism by juveniles who have sexually offended was reported in the first years after release (Hendriks & Bijleveld, 2005), while the majority of adolescent offending proves to be adolescence limited (Lussier, et al., 2012; Moffit, 1993). Therefore, differences in percentages of recidivism were expected to be most pronounced in the first years after treatment. Lastly, comparing treatment with a non-established treatment comparison group did not generate larger effects than comparing treatment with an established treatment comparison group. The inclusion of drop-out respondents in an established treatment comparison group (by most of the studies), may
have reduced their effect, since drop-outs were not able to profit fully from the established treatment administered. Drop-outs also may have been more distressed or less motivated than treatment completers; a reason for dropping out and achieving less positive results (Hendriks & Bijleveld, 2005).

Notably, not many moderating influences were detected. However, non-significant or borderline significant results may still prove clinically relevant (Man-Son-Hing et al., 2002), especially when an analysis is conducted with a relatively small sample size. We refer to the limitations section for an overview of our borderline significant findings.

**LIMITATIONS**

This study was limited by only using studies that relied on official measures of recidivism. Thereby, undoubtedly, recidivistic behavior was missed that was not officially reported. Levels of unreported sexual problem behavior ('dark numbers') are known to be quite high (Wittebrood, 2006). Also, official reports involve getting caught, which results in a specific subgroup of treated juveniles who were caught sexually offending, and again getting caught reoffending (Yun & Lee, 2013). The use of a conservative measure, however, to our opinion generated greater comparability between studies and, therefore, more reliable results.

The restrictive inclusion criteria used, and, therefore, the relatively small number of included studies (mostly performed on small samples), presented us with underpowered analyses. However, to assess the effect of treatment the inclusion of studies with a comparison group is imperative (Weisburd, Lum, & Petrosino, 2001). Some results previously detected in other reviews might have been replicated, if a larger sample size would have been available (i.e., the moderating effect of established treatment, high at recidivism risk respondent groups, dependent authors, publication, study design, and re-arrest as outcome measure), because some of these variables did show rather pronounced, yet non-significant, differences in mean effect size (see Table 3).

Furthermore, there are methodological difficulties in the use of trim and fill procedures in multilevel meta-analytic data (Nakagawa & Santos, 2012; Peters, Sutton, Jones, Abrams, & Rushton, 2007; Terrin, Schmid, Lau, & Olkin, 2003). Therefore, the difference between the adjusted and observed mean effect size in this study should be interpreted as indicative of (the effects of) publication bias; the adjusted effect size is not to be considered the true effect size.
In the included studies, most likely due to ethical considerations, one type of treatment was usually compared to another type of treatment, that is, treatment as usual. Only one comparison of a treatment group to an explicit non-treatment (refuser) group was available (Worling & Curwen, 2000). As a result, the moderating effect of no treatment could not be tested and no statement about the effects of therapy as compared to administering no therapy could be made. Also, we could not test whether treatment specifically aimed at juveniles who have sexually offended proved as effective or less effective as non-specific treatment (treatment aimed at juveniles with all types of behavioral problems) in reducing recidivism, for only one study addressed this issue (Guarino-Ghezzi & Kimball, 1998).

Lastly, recidivism is just one aspect relevant to assess the quality of treatment. In addition to reduction of recidivism, treatment aims at improving the quality of life and psychosocial functioning of clients. This dual focus is important to bear in mind, since some treatments may excel in reducing recidivism (the treatment need of society), while others might more primarily serve treatment needs as formulated by their clients (quality of life). A meta-analysis on the effect of treatment on psychosocial improvement would, therefore, complement the findings of the current study.

**CONCLUSION**

Juveniles who have sexually offended constitute a notoriously heterogeneous group regarding treatment needs and offending behavior patterns. Various treatment forms are aimed at them, although sexual recidivism rates are generally low. Relatively low sexual recidivism rates, however, do not imply that juveniles who have sexually offended are not in need of therapy to reduce recidivism risk, if only because their general recidivism rates are relatively high. To date, recidivism risk factors that are specific to juveniles who have sexually offended mainly indicate at what specific criminogenic treatment needs treatment should be aimed (i.e., childhood sexual victimization, sexual deviation, association with younger children, poor school performance, not attending school and lack of intimate peer relationships/social isolation). Also, mixed offenders, who display more antisocial tendencies, have been shown to be at higher risk for recidivism, and therefore, in greater need of treatment. The current study has shown sex offense specific treatment to have an equal effect on other types of recidivism. Thus, juveniles who have sexually offended and who display specific as well as general criminogenic needs may be best off in treatment aimed at reducing sexual recidivism. Not all juveniles who have sexually offended, however, display these features equally, so not all juveniles who have sexually offended may be in need of sex offense specific treatment as to reduce recidivism.
Chapter 3

The authors would like to encourage more robust (independent, high quality) studies on the effect of treatment for juveniles who have sexually offended, possibly making use of a non-specific treatment comparison group. Reporting on specific treatment needs might also shed more light onto what treatment works, for what (type of) sexually offending juvenile. This could improve allocation to treatment and contribute to more positive treatment effects.
Treatment Effect on Psychosocial Functioning of Juveniles with Harmful Sexual Behavior: a Multilevel Meta-Analysis

Ter Beek, E.
Kuiper, C.H.Z.
Van der Rijken, R.E.A
Spruit, A.
Stams, G.J.J.M.
Hendriks, J.

Manuscript submitted for publication
ABSTRACT

This multilevel meta-analysis examined the effects of treatment for juveniles with harmful sexual behavior on psychosocial functioning, and the potential moderating effects of outcome, treatment, participant, and study characteristics. In total, 23 studies, comprising 31 independent samples and 1,342 participants, yielded 362 effect sizes (Cohen's $d$). A moderate overall effect size was found of $d = 0.60$, indicating that groups receiving treatment achieved an estimated relative improvement in psychosocial functioning of 33%. Type of outcome did moderate the effect of treatment, indicating that effects on atypical sexual arousal and empathy (a trend) were smaller, compared to effects on other outcomes. Most prominently, studies of weak quality produced larger effect sizes. Unexpectedly, non-established treatments had more effect than did established treatments, which may be explained by the use of less rigorous study designs. Treatment groups with a higher percentage of juveniles with similar age victims or mixed type problem behavior also yielded larger effect sizes. Lastly, evaluation of treatment effects by professionals produced higher effect sizes, compared to other sources of information (e.g., adolescent self-report). Although only a marginal to no indication was found for publication bias by means of funnel plot analysis of the distribution of effect sizes, articles published in peer reviewed journals showed relatively large effect sizes. Implications for future research and clinical practice are discussed.
INTRODUCTION

Most studies on the effect of treatment for juveniles with harmful sexual behavior use recidivism as their primary outcome measure. Several meta-analyses have shown that the treatment effect on juvenile recidivism reduction is only moderate (Reitzel & Carbonell, 2006; Walker, McGovern, Poey, & Otis, 2004), or even small and non-significant after controlling for possible publication bias (Ter Beek et al., 2017). Notably, sexual recidivism is relatively rare amongst juveniles with sexually harmful behavior (Caldwell, 2016), which can result in ceiling effects (and therefore small effect sizes) when testing the effects of treatment on recidivism, whilst their psychosocial functioning often is considered to be highly problematic (Barbaree & Marshall, 2006; Seto & Lalumiere, 2010; Ryan, Leversee & Lane, 2010). Recently, several scholars specifically advocated the importance of improving the general well-being of juveniles with harmful sexual behavior, amongst others by more prominently targeting psychosocial treatment needs (Ward, 2012; Worling, 2013). In line with Self Determination Theory (Ryan & Deci, 2017), improving general well-being (i.e., a state of good mental health and social adaptation) is postulated to also reduce recidivism (See also, Willis, Yates, Gannon, & Ward, 2012), in particular with respect to externalizing disorders (Wibbelink, Hoeve, Stams, & Oort, 2017).

The effect treatment is thought to have on psychosocial functioning has not been the topic of meta-analytic research yet. Research on the improvement of psychosocial functioning of juveniles with harmful sexual behavior differs significantly in study design, type of outcome measure, type of treatment, and participant characteristics, which affects study findings and limits generalizability (Dopp, Borduin, & Brown, 2015; Hanson, Bourgon, Helmus, & Hodgson, 2009). The current study is the first to synthesize (quasi-) experimental studies evaluating the results of treatments targeting the improvement of psychosocial functioning of juveniles with harmful sexual behavior. A multilevel approach is used to explore potential moderating effects of outcome, participant, treatment, and study characteristics.

Psychosocial treatment aims of juveniles with harmful sexual behavior

Etiological theories usually provide a foundation for determining treatment goals. Different views, however, exist on the etiology of harmful sexual behavior in juveniles. A ‘specialist view’ focusses on determining and treating psychosocial dysfunctions (e.g., an atypical sexual interest or intimacy deficits) specific to juveniles with harmful sexual behavior (Van Wijk & Boonmann, 2017). Harmful sexual behavior, however, has also been explained by the presence of a more general antisocial development pattern; a ‘generalist view’ on the development of harmful sexual behavior (Dopp, Borduin, & Brown, 2015). From this viewpoint, treatment focusses on psychosocial issues linked to general conduct problems.
Recent research on the ‘specialist’ and ‘generalist’ view supports both perspectives. Juveniles with harmful sexual behavior differ from juveniles with non-sexual problem behavior by presenting more extensive histories of early sexual exposure/abuse and physical and emotional abuse or neglect, more atypical sexual interests, poorer social relationships, higher levels of anxiety, and lower self-esteem (Seto & Lalumiere, 2010). Fanniff and Kimonis (2014) did not replicate a difference in anxiety levels, but found a lower level of callous unemotional traits in juveniles with harmful sexual behavior. In general, juveniles with harmful sexual behavior indeed seem to suffer from fewer conduct problems than do non-sexually offending juveniles (Seto & Lalumiere, 2006, 2010). However, similarities between juveniles with harmful sexual behavior and juveniles with non-sexual problem behavior have also been found. McCuish, Lussier, and Corrado (2015), for example, found similar antisocial behavior patterns in sexually transgressive versus non-sexually transgressive adolescents. Seto and Lalumiere (2010) found a similar early onset of antisocial behavior, (self-reported) antisocial personality traits, exposure to non-sexual violence, family problems, interpersonal problems, general psychopathology, and IQ scores in both groups.

Typology research provides a partial explanation for these conflicting findings. Juveniles with harmful sexual behavior form a heterogeneous group regarding treatment needs and offending patterns. Juveniles with a similar age or older victim and those with a ‘mixed offending pattern’, which includes non-sexual problem behavior, seem to have relatively more in common with juveniles with non-sexual problem behavior than do juveniles with a younger victim (≥ 5 years younger and below the age of 12). Juveniles with mixed type offending and similar age victim groups show higher levels of conduct problems (Drew, 2013; Hendriks, 2006; Leroux, Pullman, Montayne, & Seto, 2016). Intrapsychic problems seem more present in juveniles with younger victims (Hendriks & Bijleveld, 2004; Van Wijk & Boonmann, 2017). Notably, a relatively large group of juveniles with harmful sexual behavior does not report any problems; many juveniles with harmful sexual behavior score within the normal range on psychosocial measures, pointing to situational or developmental phase-bound explanations for harmful sexual behavior (Ryan et al., 2010; Van Outsem, 2009) rather than to an existing dysfunction or disorder.

Most juveniles who have displayed harmful sexual behavior do not reoffend sexually (Cale, Smallbone, Rayment-McHugh, & Downling, 2016). Adolescence-limited sexual transgression and desistance are norm rather than exception (Lussier, Van den Berg, Bijleveld, & Hendriks, 2012). Only a small percentage of juveniles, notably found across all researched typologies, persist. If treatment is deemed necessary, at least three types of treatment goals may need to be addressed (specific psychosocial problems, general conduct problems and issues within the juveniles context).
The dominant paradigm for offender rehabilitation, based on its conceptual coherence and empirical support, is the Risk-Need-Responsivity (RNR) model (Andrews & Bonta, 2010; Newsome & Cullen, 2017). Hanson and colleagues (2009) have, via meta-analysis, shown its principles to apply to adults and juveniles with harmful sexual behavior, providing the (assessed) level of recidivism risk / criminogenic treatment needs (characteristics empirically associated with reoffending) a prominent place in allocation to treatment and the determination of treatment goals. This prominence, however, has recently received critique (Ward & Maruna, 2007). The professional assessment of treatment needs through the RNR paradigm, would limit treatment to ‘avoidance goals’ (relapse prevention), instead of addressing, more intrinsically motivating, ‘approach goals’. In the case of juveniles with sexual harmful behavior, low rates of sexual reoffending make sexual recidivism hard to predict. Caldwell’s (2016) most recent meta-analytic study found a weighted mean base rate for sexual recidivism by juveniles of 4.92% over a mean follow-up time of 58.98 months. Their rate of general reoffending over the same period was 30.0%. A systematic review of recidivism risk assessment tools for juveniles with harmful sexual behavior found none of the instruments reviewed undisputed, and therefore the risk factors currently in use lack predictive validity (Hempel, 2013).

In their comprehensive theory of motivation, development, and wellness, Ryan and Deci (2017) postulate that the fulfillment of three basic psychosocial human needs (relatedness, autonomy, and competence) leads to psychological well-being and adaptive social behavior. The thwarting of these needs could lead to psychosocial dysfunction, among which offending behavior. Offending behavior is thus conceptualized as a means of fulfilling a thwarted basic need, that is, functional behavior for reaching well-being under difficult circumstances. Aiming treatment at enhancing well-being by fulfilling basic psychological needs, therefore, is advocated. Ryan and Deci (2017) define well-being as a state of good mental health, social adaptation, or a combination of both, underscoring the importance of treatments successfully improving juvenile psychosocial functioning in general. Empirical evidence for this idea is found in meta-analyses by Wibbelink et al. (2017) and Van Langen, Wissink, Van Vugt, Van der Stouwe, & Stams (2014).

**Treatments for juveniles with harmful sexual behavior**

Treatments for juveniles with harmful sexual behavior make use of several treatment strategies, including behavioral therapy, cognitive therapy, psycho-education, family therapy, contextual therapy or combinations of these, mainly presented as cognitive-behavioral therapy. Such treatments are delivered in both residential and community settings, and are provided in individual and in (family)group contexts (Ryan et al., 2010; Veneziano & Veneziano, 2002). Previous meta-analyses, including studies that combine adult...
and juvenile samples, have shown cognitive behavioral based treatments and multisystemic (contextual) therapy to be most effective in reducing sexual recidivism (Dopp et al., 2015; Hanson et al., 2009; Lösel & Schmucker, 2005; Schmucker & Lösel, 2015). Treatment types that incorporate these standards are, therefore, generally considered established treatment.

Two recent meta-analyses on the effects of treatments for juveniles with harmful sexual behavior showed no moderating effects for type of treatment (Reitzel & Carbonell, 2006; Ter Beek et al., 2017), indicating all included types of treatment to be equally effective in reducing sexual as well as non-sexual recidivism.

Studies focusing on the effect of treatment on psychosocial well-being of juveniles with harmful sexual behavior report results on varying categories of psychological or social functioning as obtained by a single treatment form, predominantly not including a comparison or control group, hampering the generalization of study findings. To our knowledge no prior meta-analysis on the effect of treatment on psychosocial measures has been conducted. Thus, the question remains whether treatment in general has an effect on overall psychosocial functioning and, if so, which specific psychosocial treatment needs are most influenced.

**STUDY AIM**

The aim of this study is to review the available research on the effect of treatment on psychosocial functioning of juveniles with harmful sexual behavior. In addition, the potential moderating effects of outcome, participant, treatment, and study characteristics are investigated. This provides an opportunity to detect factors that may influence the effect of treatment on the psychosocial functioning of juveniles with harmful sexual behavior.

**METHOD**

To assess the effect of treatment on psychosocial functioning and the factors moderating this effect, a multilevel meta-analysis was carried out. The term meta-analysis refers to a stepwise procedure and a set of statistical techniques, combining results of independent primary studies into effect sizes, so that overall conclusions can be drawn. An important requirement for traditional univariate meta-analytic approaches is that no dependency between effect sizes is allowed, so that only one effect size per primary study can be included. By stepping away from the traditional univariate approach, it becomes possible to deal with dependency of effect sizes, so that all information can be preserved and a maximum of statistical power is achieved. In the current multilevel study, we distinguish between variance components distributed over three levels: differences among all effect
sizes or random sampling error (level 1), differences in effect sizes within studies (level 2),
and differences in effect sizes between studies (level 3). If there is evidence for heterogeneity
in effect sizes, moderator analyses can be conducted to test variables that may explain
within-study or between-study heterogeneity. For these analyses, the three-level random
effects model can be extended with study and effect size characteristics, making the model
a three-level mixed effects model (Assink & Wibbelink, 2016).

**Inclusion Criteria**

Multiple inclusion criteria were formulated to select the studies. First, the treatment
condition had to be aimed at improving psychosocial functioning. Second, the study
sample had to exclusively contain juveniles with harmful sexual behavior. Therefore, the
mean age of the researched group had to lie between 12 and 18 years and/or the study had
to specifically report on juveniles or adolescents referred to treatment because of harmful
sexual behavior. Third, the studies had to report on treatment results, either by reporting on
pre- and posttest measurements of a treatment group, or by comparing an experimental
treatment group with a comparison treatment group at post-test. Outcome, participant,
treatment, and study characteristics were coded as reported below (see Coding the studies).

**Selection of Studies**

All studies published before April 2017 that met the inclusion criteria were to be included
in the current meta-analysis. Firstly, several electronic databases were searched, including
Campbell library, PubMed, OVID (Medline, PsycINFO, ERIC), and Proquest (Sociological
Abstracts, Social Services Abstracts, Proquest Dissertations). Secondly, Google Scholar was
searched. The following English search string was used: (sex*) AND (offen* OR harmful OR
transgressive) AND (juvenile OR adolescent) AND (treatment OR therapy OR program OR
intervention OR training OR rehab* OR prevention OR management) AND (evaluat* OR
follow up OR outcome* OR effect* OR efficacy OR success*). No limits were used. Finally, the
references of other meta-analyses and reviews were checked for eligible studies and authors
of non-published work were contacted. Not all of the contacted authors did respond, so a
few non-published studies could not be included. A flow chart of the selection of studies is
presented in Figure 1.

The initial search and screening resulted in 50 studies that met the basic criterion of
examining the effect of an intervention on psychosocial functioning of juveniles with
harmful sexual behavior. After exclusion, 23 manuscripts remained, with 362 effect sizes,
1,342 participants, and 31 independent samples. Table 1 presents the study characteristics
of the included studies. Table 2 specifies the excluded studies and our reasons for excluding
them in italic.
### Table 1. Characteristics of Included Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Study characteristics</th>
<th>Sample characteristics</th>
<th>Treatment characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year of pub.</strong></td>
<td><strong>Author(s)</strong></td>
<td><strong>N</strong></td>
<td><strong>IS</strong></td>
</tr>
<tr>
<td>1986</td>
<td>Hains et al.</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>1988</td>
<td>Becker et al.</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>1990</td>
<td>Hunter &amp; Santos</td>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td>1990</td>
<td>Hunter &amp; Goodwin</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>1992</td>
<td>Graves, Ophenshaw, Adams</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>1993</td>
<td>Kaplan et al.</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>1993</td>
<td>Pittero</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>1994</td>
<td>Knox</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>1997</td>
<td>Weinrott et al.</td>
<td>118</td>
<td>3</td>
</tr>
<tr>
<td>1998</td>
<td>Guarino-Ghezzi &amp; Kimball</td>
<td>75</td>
<td>3</td>
</tr>
<tr>
<td>2000</td>
<td>Schuck</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>2004</td>
<td>Apsche et al.</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>2004</td>
<td>Eastman</td>
<td>100</td>
<td>1</td>
</tr>
</tbody>
</table>
### Treatment effect on Psychosocial Functioning

<table>
<thead>
<tr>
<th>Study</th>
<th>Study characteristics</th>
<th>Sample characteristics</th>
<th>Treatment characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of pub.</td>
<td>Author(s)</td>
<td>N</td>
<td># IS</td>
</tr>
<tr>
<td>2005</td>
<td>Heran</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>2008</td>
<td>Erickson</td>
<td>80</td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>Clift et al.</td>
<td>120</td>
<td>2</td>
</tr>
<tr>
<td>2009</td>
<td>Van Outsem</td>
<td>122</td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>Jones et al.</td>
<td>58</td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>Borduin et al.</td>
<td>48</td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>Letourneau et al.</td>
<td>127</td>
<td>1</td>
</tr>
<tr>
<td>2013</td>
<td>Letourneau et al.</td>
<td>124</td>
<td>1</td>
</tr>
<tr>
<td>2014</td>
<td>Greaves &amp; Salloum</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>2017</td>
<td>Ter Beek et al.</td>
<td>34</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note:** N = number of participants; # IS = number of independent samples; # r (M) = number of effect sizes (mean d); Pub. = published in peer reviewed article yes/no; Study Quality = strong/moderate/weak according to EPHPP quality assessment tool; Design = RCT (randomized controlled trial) or QE (quasi experimental); Independent author = yes/no; Type of psychosocial functioning measured: 1 = overall functioning, 2 = rule breaking & aggression, 3 = impulse control, 4 = social & coping skills, 5 = empathy, 6 = emotions & self-image, 7 = cognitions and knowledge on sexuality, 8 = arsical sexual arousal, 9 = family functioning; M. Age = mean age of sample; % Male = percentage of males in sample; % Caucasian = percentage of Caucasian ethnicity; % Child vic. = percentage with child victims; % SA vic. = percentage with same age or older victims; Setting = outpatient, residential or a mix; Treatment type = Type of intervention: CT = cognitive therapy, BT = behavioral therapy, CBT = cognitive-behavioral therapy, ST = systemic therapy, non-established (NEST) or established (EST); Treatment Condition = name of intervention, Comparison or Control condition = name of control intervention, none = no treatment comparison group; FFT = Functional Family Therapy, MST-PSB = Multi Systemic Therapy-Problem Sexual Behavior. **these categories are not mutually exclusive, juveniles in some studies were reported as both types creating a higher than 100% total score in other studies juveniles with younger and same age victims were left out of both categories.**
Chapter 4

Identification

177 records through database search

Screening

24 additional records through other sources, 4 not retrievable (unpub.)

197 records screened

Eligibility

147 records excluded

50 full text articles assessed

27 full text articles excluded:
- 25 non-eligible outcome measurements
- 2 non-eligible samples

Included

23 studies included in quantitative synthesis

Figure 1. Flowchart of study selection

Coding the Studies

Table 4 shows all variables that were coded in this study. The dependent variable in this meta-analysis was psychosocial functioning. The independent variable was the treatment offered. Type of psychosocial functioning, participant, treatment, and study characteristics were coded to assess whether treatment effects varied among the possible moderator variables. In order to reduce the problem of multiple testing (Tabachnik & Fidell, 2013), only moderators of possible theoretical importance were used. Studies using multiple independent samples were coded with separate study identification numbers. Two research assistants coded the included studies according to the suggestions of Lipsey and Wilson (2001). Five studies (22%) were double coded. Following the guidelines by Shrout (1989), for the continuous variables ICCs were calculated for the single measure two-way random effects model, with absolute agreement as a criterion. The mean ICC for all 5 variables was 1.00 (i.e., perfect reliability). For the categorical variables kappa was calculated (Landis & Koch, 1977) yielding almost perfect interrater reliabilities (i.e., kappa .96). One variable (i.e., type of placement) reached substantial reliability (.69).
Table 2. Excluded Studies and Reasons for Exclusion

<table>
<thead>
<tr>
<th>Study</th>
<th>Year of pub.</th>
<th>Author(s)</th>
<th>N</th>
<th>Group</th>
<th>Measurements</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989 McConaghy et al.</td>
<td>1989</td>
<td>ASO &amp; JHSB</td>
<td>45</td>
<td>None (descriptive)</td>
<td>None (descriptive)</td>
<td>Medication &amp; Covert Sensitization</td>
</tr>
<tr>
<td>1990 Borduin et al.</td>
<td>1990</td>
<td>JHSB</td>
<td>16</td>
<td>None (descriptive)</td>
<td>None (descriptive)</td>
<td>MST-PSB</td>
</tr>
<tr>
<td>1991 Bromberg</td>
<td>1991</td>
<td>JHSB</td>
<td>199</td>
<td>Non-validated test</td>
<td>Non-validated test</td>
<td>Outpatient JSO Specific Treatment</td>
</tr>
<tr>
<td>1992 Becker et al.</td>
<td>1992</td>
<td>JHSB</td>
<td>160</td>
<td>T1 measurements only</td>
<td>Covert Sensitization</td>
<td></td>
</tr>
<tr>
<td>1993 Emmerick &amp; Dutton</td>
<td>1993</td>
<td>JHSB</td>
<td>67</td>
<td>T1 measurements only</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>1995 Hunter et al.</td>
<td>1995</td>
<td>JHSB</td>
<td>76</td>
<td>None (test validation)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>1998 Simpson et al.</td>
<td>1998</td>
<td>JHSB</td>
<td>12</td>
<td>No post measurements</td>
<td>Adventure based treatment</td>
<td></td>
</tr>
<tr>
<td>2000 Worling &amp; Curwen</td>
<td>2000</td>
<td>JHSB</td>
<td>148</td>
<td>T1 measurements only &amp; Recidivism</td>
<td>SAFE-T</td>
<td></td>
</tr>
<tr>
<td>2000 Derezotes</td>
<td>2000</td>
<td>JHSB</td>
<td>14</td>
<td>None (descriptive)</td>
<td>None (descriptive)</td>
<td>Yoga &amp; Meditation</td>
</tr>
<tr>
<td>2000 Seto et al.</td>
<td>2000</td>
<td>JHSB</td>
<td>150</td>
<td>T1 measurements only</td>
<td>Sensitization</td>
<td></td>
</tr>
<tr>
<td>2000 Lambie et al.</td>
<td>2000</td>
<td>JHSB</td>
<td>14</td>
<td>None (descriptive)</td>
<td>Wilderness Community Treatment</td>
<td></td>
</tr>
<tr>
<td>2000 Cooper</td>
<td>2000</td>
<td>JHSB</td>
<td>89</td>
<td>Recidivism</td>
<td>TBASOP</td>
<td></td>
</tr>
<tr>
<td>2002 Myklebust &amp; Kay</td>
<td>2002</td>
<td>JHSB</td>
<td>100</td>
<td>T1 measurements only</td>
<td>Juvenile Correctional Facility</td>
<td></td>
</tr>
<tr>
<td>2004 Saleh et al.</td>
<td>2004</td>
<td>JHSB</td>
<td>6</td>
<td>None (descriptive)</td>
<td>Residential Treatment</td>
<td></td>
</tr>
<tr>
<td>2004 Ryback</td>
<td>2004</td>
<td>JHSB</td>
<td>21</td>
<td>None (descriptive)</td>
<td>Residential JSO Treatment</td>
<td></td>
</tr>
<tr>
<td>2005 Aylwin et al.</td>
<td>2005</td>
<td>JHSB</td>
<td>87</td>
<td>No post measurements</td>
<td>Covert Sensitization</td>
<td></td>
</tr>
<tr>
<td>2005 Apsche et al.</td>
<td>2005</td>
<td>JNSO &amp; JHSB</td>
<td>60</td>
<td>T1 &amp; T2</td>
<td>CBT, Social Skills Training, Mode Deactivation Therapy</td>
<td></td>
</tr>
<tr>
<td>2006 Worling &amp; Längstrom</td>
<td>2006</td>
<td>JHSB</td>
<td>78</td>
<td>T1 measurements only</td>
<td>Residential and Community based treatment</td>
<td></td>
</tr>
<tr>
<td>2006 Van Outsem et al.</td>
<td>2006</td>
<td>JHSB &amp; JNSO</td>
<td>196</td>
<td>T1 measurements only</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2007 McCoy</td>
<td>2007</td>
<td>JHSB</td>
<td>128</td>
<td>Recidivism</td>
<td>Outpatient JSO treatment</td>
<td></td>
</tr>
<tr>
<td>2008 Hendriks &amp; Bijleveld</td>
<td>2008</td>
<td>JHSB</td>
<td>114</td>
<td>Recidivism</td>
<td>Residential JSO treatment</td>
<td></td>
</tr>
<tr>
<td>2008 Letourneau et al.</td>
<td>2008</td>
<td>JSP &amp; JHSB</td>
<td>696</td>
<td>T1 &amp; T2</td>
<td>MST-PSB</td>
<td></td>
</tr>
<tr>
<td>2009 Vlijoen et al.</td>
<td>2009</td>
<td>JHSB</td>
<td>193</td>
<td>Recidivism</td>
<td>Residential Treatment</td>
<td></td>
</tr>
<tr>
<td>2010 Worling et al.</td>
<td>2010</td>
<td>JHSB</td>
<td>148</td>
<td>Recidivism</td>
<td>SAFE-T</td>
<td></td>
</tr>
<tr>
<td>2010 Hart-Kerkhoffs</td>
<td>2010</td>
<td>JHSB</td>
<td>226</td>
<td>T1 measurements only</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2011 Halse et al.</td>
<td>2011</td>
<td>JHSB</td>
<td>12</td>
<td>None (descriptive)</td>
<td>Community Based Treatment</td>
<td></td>
</tr>
<tr>
<td>2016 Newland</td>
<td>2016</td>
<td>JHSB</td>
<td>129</td>
<td>T1 measurements only</td>
<td>Residential Treatment</td>
<td></td>
</tr>
</tbody>
</table>

Note. The characteristic in italic font specifies reason(s) for exclusion. JHSB = Juveniles with Harmful Sexual Behavior; JNSO = Juvenile Non Sex Offenders; ASO = Adult Sex offenders; JSP = Juveniles with Sexual Problems (also inappropriate sexual behavior).
Characteristics of the psychosocial outcome measure

The type of psychosocial functioning measured was coded into several broad categories. Nine psychosocial constructs were distinguished through review of the research literature, as to aggregate the different measures used: overall psychosocial functioning (e.g., CGAS, CAFAS total scores), rule breaking and aggression (e.g., ASEBA, ASAP-D scale scores), impulse control (e.g., ASAP-D, MESSY scale scores), social skills and coping (e.g., ASSET, SPSI scale scores), empathy (e.g., ASAP-D, IRI scale scores), (negative) emotions and self-image (ACLSA-II, OQ-45 scale scores), cognitive distortions and sexual knowledge (MSI, PAA scale scores), atypical sexual arousal1 (e.g., ASAP-D, ASIC scale scores, but mostly measured by penile plethysmography and operationalized as being sexually aroused by young children of coerced / sadistic sexual activity) and, finally, family functioning (e.g., PSI, PARI scale scores). An overview of which psychometric measurements were used in the included studies, and to what constructs they contributed is offered in Table 3. Not all studies reported on all outcomes. An overview of how many independent samples and effect sizes contributed to each outcome is presented in Table 4.

Participant characteristics

The cultural background of the juveniles was coded as the percentage of Caucasians in the researched group. Furthermore, the percentage of juveniles with younger victims (< 12 years of age and ≥ 5 years younger) was coded, as was the percentage of juveniles with peer (similar age) victims, and the percentage of juveniles with mixed type problem behavior (also displaying non-sexual transgressions). Finally, the type of harmful sexual behavior was coded as with physical contact (such as rape) or as also non-contact sexual behavior (such as voyeurism).

1. What constitutes atypical sexual arousal in juveniles is much debated. The relative normalcy of feeling sexually aroused by relatively atypical stimuli in adolescence, a developmental stage defined by its flexibility and high levels of hormonal imbalances, has been established. Therefore, this construct is no longer in use as an outcome measure in contemporary research.
Table 3. Overview of psychometric measurements used

<table>
<thead>
<tr>
<th>Study</th>
<th>Year of pub.</th>
<th>Author(s)</th>
<th>Measurements, with reference as mentioned in the original manuscript</th>
<th>Type of Functioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986 Hains et al.</td>
<td>1986</td>
<td></td>
<td>Adolescent Problems Inventory (Freedman et al., 1978), Sexual Knowledge Questionnaire &amp; Psychological Inventory (Kirby et al., 1979), Defining Issues Test (Rest, 1974).</td>
<td>4, 7</td>
</tr>
<tr>
<td>1988 Becker et al.</td>
<td>1988</td>
<td></td>
<td>Penile Plethysmography (mercury-in-rubber strain gauge)</td>
<td>8</td>
</tr>
<tr>
<td>1990 Hunter &amp; Santos</td>
<td>1990</td>
<td></td>
<td>Penile Plethysmography (indium-gallium strain gauge)</td>
<td>8</td>
</tr>
<tr>
<td>1992 Hunter &amp; Goodwin</td>
<td>1992</td>
<td></td>
<td>Penile Plethysmography (indium-gallium strain gauge)</td>
<td>8</td>
</tr>
<tr>
<td>1992 Graves, Openshaw, Adams</td>
<td>1992</td>
<td></td>
<td>CBC (Achenbach &amp; Edelbrock, 1978), ASSET pre-post training checklist (Adams et al., 1988), Piers Harris self-concept scales, Parent Adolescent Relationship Inventory (Robin, Koepke &amp; Mayor, 1984)</td>
<td>1, 2, 4, 6, 9</td>
</tr>
<tr>
<td>1993 Kaplan et al.</td>
<td>1993</td>
<td></td>
<td>Penile Plethysmography (mercury-in-rubber strain gauge)</td>
<td>8</td>
</tr>
<tr>
<td>1994 Piliero</td>
<td>1994</td>
<td></td>
<td>Multiphasic Sexual Inventory (Nichols &amp; Molinder, 1984)</td>
<td>7, 8</td>
</tr>
<tr>
<td>1994 Knox</td>
<td>1994</td>
<td></td>
<td>Achenbach System of Empirically Based Assessment - ASEBA (Achenbach &amp; Edelbrock, 1983), Matson Evaluation of Social Skills Youth (Matson, 1990), Social Problem Solving Index (D’Zurilla &amp; Nezu, 1990)</td>
<td>1, 2, 3, 4, 6, 7</td>
</tr>
<tr>
<td>1997 Weinrott et al.</td>
<td>1997</td>
<td></td>
<td>Adolescent Sexual Interest Cardsort (Becker &amp; Kaplan, 1988), Penile Plethysmography</td>
<td>8</td>
</tr>
<tr>
<td>1998 Guarino-Ghezzi &amp; Kimball</td>
<td>1998</td>
<td></td>
<td>Attitudes towards Sex, Rape Myth Scale, Adolescent Cognition Scale (no references)</td>
<td>4, 7</td>
</tr>
<tr>
<td>2000 Schuck</td>
<td>2000</td>
<td></td>
<td># behavioral incidents, Ansell-Casey Life Skills Assessment (no reference), ASEBA (Achenbach, 1991)</td>
<td>2, 4, 6</td>
</tr>
<tr>
<td>2004 Apsche et al.</td>
<td>2004</td>
<td></td>
<td>Devereux Scale of Mental Disorders (Devereux Foundation, 1984*), ASEBA (Achenbach, 1991), J-SOAP (Prentky et al., 2000), Beliefs Assessment (Apsche, 2000*)</td>
<td>1, 7</td>
</tr>
<tr>
<td>2004 Eastman</td>
<td>2004</td>
<td></td>
<td>Interpersonal Reactivity Index (Davis, 1983), Index of Self Esteem (Hudson, 1987), Sexual Knowledge Questionnaire (Kirby et al., 1979), Attitudes and Values Inventory (Kirby, 1984), Bumby Cognitive Distortions Scales (Bumby, 1996)</td>
<td>5, 6, 7</td>
</tr>
<tr>
<td>2005 Heran</td>
<td>2005</td>
<td></td>
<td>Teenage Inventory of Social Skills (Inderbitzen &amp; Foster, 1992), Interpersonal Reactivity Index (Davis, 1980), Child Molester Empathy Measure (Davis, 1983), empathy logs</td>
<td>4, 5</td>
</tr>
<tr>
<td>2008 Erickson</td>
<td>2008</td>
<td></td>
<td>Outcome Questionnaire-45 (Lambert et al., 1996), Youth Outcome Questionnaire (Burlingame et al, 1996)</td>
<td>2, 4, 6</td>
</tr>
<tr>
<td>2009 Clift et al.</td>
<td>2009</td>
<td></td>
<td>Penile Plethysmography (mercury strain gauges)</td>
<td>8</td>
</tr>
</tbody>
</table>
### Study

<table>
<thead>
<tr>
<th>Year of pub.</th>
<th>Author(s)</th>
<th>Measurements, with reference as mentioned in the original manuscript</th>
<th>Type of Functioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Van Outsem</td>
<td>Adolescent Sexoffender Assessment Pack-Dutch (Van Outsem et al., 2006)</td>
<td>2, 3, 4, 5, 6, 7, 8</td>
</tr>
<tr>
<td>2009</td>
<td>Jones et al.</td>
<td>Child and Adolescent Functional Assessment Scale (Hodges 1995), Abel Assessment for Sexual Interest (Abel et al., 2004)</td>
<td>1, 4, 8</td>
</tr>
<tr>
<td>2009</td>
<td>Borduin et al.</td>
<td>Brief Symptoms Inventory-Global Severity Index youth (Derogatis, 1993), Revised Behavior Problem Checklist (Quay &amp; Peterson, 1987), Missouri Peer Relations Inventory (Borduin et al., 1989), Family Adaptability and Cohesion Evaluation Scales II (Olson et al., 1982)</td>
<td>1, 2, 4, 9</td>
</tr>
<tr>
<td>2009</td>
<td>Letourneau et al.</td>
<td>ASEBA (Achenbach, 1995; Achenbach 2001), Self-Report Delinquency scale (Elliot et al., 1985), Personal Experience Inventory (Winters &amp; Henly, 1989), Adolescent Sexual Behavior Inventory (Friedrich et al., 2004),</td>
<td>1, 2, 8</td>
</tr>
<tr>
<td>2013</td>
<td>Letourneau et al.</td>
<td>Services Utilization Tracking Form (Henggler et al., 1997), Self-Report Delinquency scale (Elliot et al., 1985), Personal Experience Inventory (Winters &amp; Henly, 1989), Adolescent Sexual Behavior Inventory (Friedrich et al., 2004),</td>
<td>1, 2, 8</td>
</tr>
<tr>
<td>2014</td>
<td>Greaves &amp; Salloum</td>
<td>Estimate of Risk of Adolescent Sexual Offense Recidivism (Worling, 2004), Child Global Assessment Scale (Shaffer et al., 1983), Parent Stress Index-Short Form (Abidin, 1995)</td>
<td>1, 4, 8, 9</td>
</tr>
<tr>
<td>2017</td>
<td>Ter Beek et al.</td>
<td>Adolescent Sexoffender Assessment Pack-Dutch (Van Outsem et al., 2006)</td>
<td>2, 3, 4, 5, 6, 7, 8</td>
</tr>
</tbody>
</table>

**Note:** Type of psychosocial functioning designated: 1 = overall functioning, 2 = rule breaking & aggression, 3 = impulse control, 4 = social & coping skills, 5 = empathy, 6 = emotions & self-image, 7 = cognitions and knowledge on sexuality, 8 = atypical sexual arousal, 9 = family functioning. * = reference not found in original reference list.

---

**Treatment characteristics**

Firstly, the duration of treatment was coded in months. The exclusion of respondents with a (borderline) intellectual disability was coded as yes or no. It was coded whether treatment was specifically designed as treatment of juveniles with harmful sexual behavior, or whether the same treatment was also offered to juveniles without harmful sexual (but otherwise problematic) behavior. The timeframe in which treatment was offered was coded as before 2000 and after 2000, because in Caldwell’s recent meta-analysis (2016) it was hypothesized that after the year 2000 treatment might have become more effective. Treatment status was coded as non-established treatment (NEST) or established treatment (treatment that has been referred to in previous research as effective, i.e., incorporating cognitive behavioral treatment and/or systemic therapy). In addition, it was coded whether the type of treatment was cognitive behavioral, behavioral, cognitive, or contextual. The type of placement was coded as following a conviction, mandatory treatment (without conviction), voluntary
Treatment effect on Psychosocial Functioning

enrollment, and mixed (several types of placements). Treatment context was coded as residential or outpatient. Treatment method was coded as group therapy, individual therapy, family therapy, or a mix of these. If reported on, treatment integrity was coded as high (adhering to the protocol, having supervision and training), medium (adhering to a manual), or low (none of the aforementioned). Finally, if reported on, also the level of responsivity of the treatment offered (the reported flexibility in adjusting the treatment to the individual’s learning and coping style, motivation, and individual treatment needs) was coded as high (fully adjusting treatment to the individuals’ preferences/needs), medium (responding to individual characteristics of the juvenile), or low (following the protocol/prescribed treatment modules for all juveniles alike).

Study characteristics

It was coded whether a comparison treatment was used and on what continent (North America or Europe) the study was performed. Intention to treat was coded as yes (including all juveniles in posttreatment measurements) or no. For example, the classification intent to treat was not awarded when some juveniles had refused to complete posttreatment measurements or had dropped out of treatment, which was relatively often the case. It was coded whether the authors were independent researchers or whether they were involved in the development or implementation of the intervention. It was also coded whether the study was published in a peer reviewed journal. Further, the design of the study (randomized controlled trial versus quasi-experimental) was coded. The type of effect size calculation was coded as mean gain score (calculation based on pre- and posttest values of the same group, accounting for test-retest reliability), means and standard deviations of posttest values of two groups, a T or F test value, proportions (percentages) or significance levels. The type of informant was coded as professional (e.g. a type of measurement filled in by the therapist about the juvenile), penile plethysmography (the measurement of physical arousal to atypical sexual stimuli), parents (e.g. a type of measurement filled in by the parents about the juvenile such as the CBCL), or self-report (e.g. a type of measurement filled in by the juvenile about himself such as the YSR). Lastly, study quality was coded by use of the EPHPP Quality Assessment Tool for Quantitative Studies (http://www.ephpp.ca/tools.html). This tool assesses the quality of a study as weak, moderate or strong, providing a comprehensive and structured assessment of the concept of study quality (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012). It has been judged suitable to be used in systematic reviews of effectiveness (Deeks et al., 2003) and has been reported to have sufficient content and construct validity (Jackson & Waters, 2005; Thomas, Ciliska, Dobbins, & Micucci, 2004). The tool assesses six domains: selection bias, study design (including appropriateness of the design), confounders, blinding, data collection methods, and withdrawals and dropouts. Table 1 shows the results of the assessment.
Chapter 4

Calculations

Effect sizes were transformed into Cohen’s d by using the calculator of Wilson (2013) and formulas of Lipsey and Wilson (2001). A positive effect size indicated that the treatment group benefited from treatment, whereas a negative effect size indicated that there was a negative effect of treatment as compared to a comparison group or compared to the treatment group itself at admission. To account for differences in effect sizes between pre-posttest measurement and posttest measurements, a mean gain score was calculated for pre-posttest measures, accounting for test-retest reliability (Morris & DeShon, 2002). If a study only mentioned that an effect was not significant (as was the case in 2.5% of all effect sizes), the effect size was coded as zero (Lipsey & Wilson, 2001). The continuous variables (percentage Caucasian, percentage with younger victims, percentage with similar age victims, percentage with mixed problem behavior, and mean treatment length) were centered around their mean, and all other (categorical) variables were recoded into dummy variables. We checked for the presence of extreme outliers using Z scores (Tabachnik & Fidell, 2013); no extreme outliers were detected. Standard errors were estimated using formulas of Lipsey and Wilson (2001).

In all studies we were able to calculate more than one effect size. Most studies reported on multiple outcome variables. Effect sizes from the same study may prove more alike than effect sizes from different studies. Therefore, the assumption of statistical independency, which underlies classical meta-analytic strategies, was violated (Hox, 2002; Lipsey & Wilson, 2001). In line with recently conducted meta-analyses, we applied a multilevel approach in order to deal with the interdependency of effect sizes (Assink et al., 2015; Houben, Van den Noortgate, & Kuppens, 2015; Spruit, Assink, Van Vugt, Van der Put, & Stams, 2016; Weisz et al., 2013). The multilevel approach accounts for the hierarchical structure of the data in which effect sizes are nested within the studies (Van den Noortgate & Onghena, 2003; Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013).

We used a three-level meta-analytic model to calculate the combined effect sizes and to perform moderator analyses. The sampling variance of observed effect sizes (level 1) was estimated by using the formula of Cheung (2014). Log-likelihood-ratio-tests were performed to compare the deviance of the full model to the deviance of the models excluding one of the variance parameters, making it possible to determine whether significant variance is present at the second and third level (Wibbelink & Assink, 2015). Significant variance at level 2 or 3 indicates a heterogeneous effect size distribution, meaning that the effect sizes cannot be treated as estimates of a common effect size. In that case, we proceeded to moderator analyses, because the differences between the effect sizes may be explained by outcome, study, sample, and/or intervention characteristics. Moderator analyses were only performed when each category of the potential moderator was filled with at least
three studies. As a result, cognitive distortions and sexual knowledge were collapsed into one moderator (cognitions & knowledge), because only two studies reported on the latter. Behavioral treatment and cognitive treatment were recoded into ‘behavioral or cognitive treatment’, because only two treatments were considered cognitive. Voluntary treatment was recoded into voluntary & mixed, because only one treatment mentioned strictly voluntary enrollment. Treatment integrity was recoded into a dichotomous variable (high versus medium & weak), because only one study was considered to have a weak treatment integrity. Lastly, in type of effect size calculation, proportion-based and significance-based calculations were collapsed into proportion & significance because significance measures were only used in one study.

The multilevel meta-analysis was conducted in R (version 3.2.0) with the metafor-package, using a multilevel random effects model (Viechtbauer, 2010; Wibbelink & Assink, 2015). The restricted maximum likelihood estimate was used to estimate all model parameters, and the Knapp and Hartung-method (2003) was used for testing individual regression coefficients of the meta-analytic models and for calculating the corresponding confidence intervals (see also Assink et al., 2015; Houben et al., 2015; Spruit et al, 2016; Wibbelink & Assink, 2015).

**Publication Bias**

In systematic reviews, the aim is to include all studies previously conducted that meet the inclusion criteria (Lipsey & Wilson, 2001). However, a common problem is that studies may not have been published due to non-significant or unfavorable findings and, therefore, are difficult to locate. Not including these studies may lead to an overestimation of the true effect size, the so called ‘publication bias’ (Rosenthal, 1979). In order to check the presence of publication bias in our meta-analysis, we performed a trim and fill procedure (Duval & Tweedie, 2000) after averaging all effect sizes at the third between study level by drawing a trim-and-fill plot in R (Version 3.2.0) using the function ‘trimfill’ of the metafor package (Viechtbauer, 2010). Notably, publication bias does not have bearing on the within study level effects. Therefore, publication bias was examined by means of a traditional random effects model with only one mean effect size per study. We tested whether effect sizes were missing on the left side of the distribution, since publication bias would only be likely to occur in case of non-significant or unfavorable (i.e., negative) results.

**RESULTS**

Overall, a significant, moderate effect ($d = 0.60$, $p < .001$) of treatment on psychosocial functioning was found, indicating that the treatment groups achieved an estimated improvement in psychosocial functioning of 33%. Publication bias was examined by using
the aggregated effect sizes per study, with seven trim-and-fill plot imputed estimations of effect sizes of missing studies at the left side of the funnel plot, indicating the presence of possible publication bias (see Figure 2). We included the imputed estimations (the open circles) and performed the meta-analysis again to compute an overall effect size that takes the influence of publication bias into account (Duval & Tweedie, 2000). After controlling for publication bias, the overall aggregated effect size of Cohen’s $d = 0.59$ ($p < .001$) remained moderate and significant (effect size of Cohen’s $d = 0.48$, $p < .001$), with overlapping confidence intervals of the original and unbiased estimates of $0.45 < \text{Cohen's } d < 0.74$ and $0.33 < \text{Cohen's } d < 0.63$, respectively. Although visual inspection of the funnel plot (asymmetry at the left side) and the drop in overall effect size ($\Delta \text{Cohen's } d = -0.11$) suggest a minor effect of publication bias, overlapping confidence intervals show that the results of this meta-analysis are not significantly affected by publication bias.

Figure 2. Results of trim-and-fill procedure testing for publication bias.

Note: The closed circles represent the primary studies included. The open circles represent forecasted missing effect sizes, pointing out possible publication bias.
The likelihood ratio test comparing models with and without between-study variance (level 3) showed that significant variance was present at the between-study level ($\sigma^2_{\text{level 3}} = 0.151$, $\chi^2(1) = 129.92; p < .0001$). The variance between the effect sizes within studies (level 2) was also significant ($\sigma^2_{\text{level 2}} = 0.119$, $\chi^2(1) = 337.15; p = < .0001$). About 17% of the total effect size variance was accounted for by the sampling variance (level 1), 36% by the variance between effect sizes within studies (level 2), and 46% by the variance between studies (level 3). Because the variance on the third level was significant, we proceeded to moderator analyses to assess factors that could possibly explain variance in treatment effects (see Table 4).

As presented in the last column of Table 4, the type of psychosocial functioning measured moderated the effect of treatment on psychosocial functioning in juveniles with harmful sexual behavior. Weaker treatment effects were found for changes in atypical sexual arousal, as compared to the reference category. Also a moderating trend was found for measurement of the improvement of empathy, indicating that regarding the improvement of empathy, treatment tended to be less effective. On all other psychosocial constructs treatment was found to be equally effective. Within participant characteristics the percentage of juveniles with similar age victims was found to moderate the effect of treatment, indicating that in samples with higher percentages of juveniles with similar age victims (peers), better treatment results were obtained. Also the percentage of juveniles with a mixed offending pattern moderated the effect of treatment, indicating that for samples with higher percentages of juveniles with also non-sexual problem behavior, better treatment results were obtained. For all other participant characteristics treatment was equally effective on improving psychosocial functioning. Of the treatment characteristics, treatment status moderated the effect of treatment. Treatment as usual yielded higher effect sizes than did established treatment. Moderating trends were found for type of treatment, the years in which treatment was administered, and for treatment integrity, indicating that behavioral or cognitive treatments, treatments before 2000, and treatments with a medium to low treatment integrity tended to yield larger effect sizes. Of the study characteristics, peer reviewed publishing moderated the effect of treatment, generating higher effect sizes in published studies. Also the type of informant moderated treatment effect: professional judgments of improvement yielded larger results than did penile plethysmography, parental judgment, and self-report. Lastly, study quality moderated the effect of treatment on psychosocial functioning. Studies with a weak quality generated larger effect sizes than studies with a strong study quality.
Table 4. Overall Results and Moderator Effects of the Relation between Treatment and Psychosocial Functioning

<table>
<thead>
<tr>
<th>Moderator variables</th>
<th># ES</th>
<th># ES</th>
<th>β₀ (mean d)</th>
<th>t₁₀</th>
<th>β₁</th>
<th>t₁</th>
<th>F(df₁, df₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall relation</td>
<td>31</td>
<td>362</td>
<td>0.600</td>
<td>7.679***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Outcome characteristics**

<table>
<thead>
<tr>
<th>Type of Functioning</th>
<th># ES</th>
<th># ES</th>
<th>β₀ (mean d)</th>
<th>t₁₀</th>
<th>β₁</th>
<th>t₁</th>
<th>F(df₁, df₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Functioning (RC)</td>
<td>8</td>
<td>30</td>
<td>0.761</td>
<td>6.184**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule Breaking &amp; Aggression</td>
<td>11</td>
<td>26</td>
<td>0.626</td>
<td>5.273*** -0.136 -1.100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impulse control</td>
<td>4</td>
<td>7</td>
<td>1.002</td>
<td>5.531*** 0.240 1.253</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Skills &amp; Coping</td>
<td>14</td>
<td>60</td>
<td>0.870</td>
<td>8.425*** 0.109 0.968</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empathy</td>
<td>5</td>
<td>28</td>
<td>0.484</td>
<td>3.539*** -0.278 -1.810*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotion &amp; Self-image</td>
<td>9</td>
<td>29</td>
<td>0.637</td>
<td>5.184*** -0.124 -0.939</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitions &amp; Knowledge</td>
<td>10</td>
<td>56</td>
<td>0.803</td>
<td>7.175*** 0.042 0.322</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical Sexual Arousal</td>
<td>18</td>
<td>118</td>
<td>0.351</td>
<td>3.617*** -0.410 -3.658***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Functioning</td>
<td>4</td>
<td>12</td>
<td>0.557</td>
<td>3.314* 0.205 -1.188</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Participant characteristics**

| Percentage Caucasian | 24  | 276  | 0.630       | 6.677*** 0.002 0.644 0.414 (1, 274) |
| Percentage Younger Vict. | 19   | 191  | 0.523       | 6.052*** 0.005 1.275 1.625 (1, 189) |
| Percentage Similar Age Vict. | 10   | 113  | 0.360       | 6.379*** 0.015 3.335* 11.124 (1, 111)* |
| Percentage Mixed Prob. Beh. | 5    | 94   | 0.338       | 3.332** 0.007 2.587* 6.693 (1, 92)* |
| Type of Sex. Behavior | 27  | 313  |             | 0.117 (1, 311) |
| Contact (RC)         | 24  | 261  | 0.552       | 7.007*** |
| Also non-contact      | 3   | 52   | 0.474       | 2.193* -0.079 -0.343 |

**Treatment characteristics**

| Duration Treatment | 28  | 347  | 0.600       | 7.154*** -0.012 -1.000 0.999 (1, 345) |
| Exclusion Low IQ    | 29  | 345  |             | 1.856 (1, 343) |
| No (RC)             | 23  | 247  | 0.627       | 7.453*** |
| Yes                 | 6   | 98   | 0.393       | 2.611* -0.235 -1.362 |
| JSO Specific Treatment | 31  | 362  |             | 0.523 (1, 360) |
| Yes (RC)            | 26  | 288  | 0.575       | 6.652** |
| No                  | 5   | 74   | 0.725       | 3.814*** 0.151 0.723 |
| Treatment Administration | 28  | 321  |             | 5.558 (1, 319)* |
| < 2000 (RC)         | 21  | 229  | 0.692       | 8.150*** |
| > 2000              | 7   | 92   | 0.323       | 2.451 -0.369 -2.357* |
| Treatment Status    | 31  | 362  |             | 5.136 (1, 360)* |
| NEST (RC)           | 13  | 143  | 0.814       | 6.813*** |
| EST                 | 18  | 219  | 0.473       | 5.198*** -0.340 -2.266* |
| Type of treatment   | 31  | 362  |             | 2.997 (2,359)* |
| Cognitive Behavioral (RC) | 16  | 138  | 0.536       | 5.172*** |
### Treatment effect on Psychosocial Functioning

#### Moderator variables

<table>
<thead>
<tr>
<th>Moderator variables</th>
<th># IS</th>
<th># ES</th>
<th>$\beta_0$ (mean $d$)</th>
<th>$t_0$</th>
<th>$\beta_1$</th>
<th>$t_1$</th>
<th>$F(df_1,df_2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral or Cognitive</td>
<td>10</td>
<td>140</td>
<td>0.778</td>
<td>6.518$^{***}$</td>
<td>0.305</td>
<td>1.842$^{*}$</td>
<td></td>
</tr>
<tr>
<td>Contextual</td>
<td>5</td>
<td>67</td>
<td>0.390</td>
<td>2.687$^{**}$</td>
<td>-0.181</td>
<td>-0.918</td>
<td></td>
</tr>
<tr>
<td><strong>Type of Placement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convicted (RC)</td>
<td>10</td>
<td>93</td>
<td>0.570</td>
<td>4.427$^{***}$</td>
<td></td>
<td></td>
<td>0.0815 (2, 322)</td>
</tr>
<tr>
<td>Mandatory Treatment</td>
<td>4</td>
<td>69</td>
<td>0.660</td>
<td>3.212$^{***}$</td>
<td>0.090</td>
<td>0.370</td>
<td></td>
</tr>
<tr>
<td>Mix &amp; Voluntary</td>
<td>13</td>
<td>163</td>
<td>0.569</td>
<td>4.538$^{***}$</td>
<td>-0.001</td>
<td>-0.007</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential (RC)</td>
<td>16</td>
<td>165</td>
<td>0.660</td>
<td>0.660$^{**}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatient</td>
<td>15</td>
<td>197</td>
<td>0.547</td>
<td>5.054$^{***}$</td>
<td>-0.113</td>
<td>-0.712</td>
<td></td>
</tr>
<tr>
<td><strong>Method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mix (RC)</td>
<td>16</td>
<td>163</td>
<td>0.734</td>
<td>6.761$^{***}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group therapy</td>
<td>7</td>
<td>112</td>
<td>0.515</td>
<td>3.308$^{**}$</td>
<td>-0.219</td>
<td>-1.153</td>
<td></td>
</tr>
<tr>
<td>Individual therapy</td>
<td>3</td>
<td>20</td>
<td>0.588</td>
<td>2.077$^{*}$</td>
<td>-0.146</td>
<td>-0.482</td>
<td></td>
</tr>
<tr>
<td>Family therapy</td>
<td>5</td>
<td>67</td>
<td>0.354</td>
<td>1.989$^{*}$</td>
<td>-0.380</td>
<td>-0.823</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment Integrity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High (RC)</td>
<td>6</td>
<td>91</td>
<td>0.343</td>
<td>2.221$^{*}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium &amp; Low</td>
<td>8</td>
<td>105</td>
<td>0.721</td>
<td>5.229$^{***}$</td>
<td>0.378</td>
<td>1.824$^{*}$</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment Responsivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High (RC)</td>
<td>5</td>
<td>65</td>
<td>0.582</td>
<td>2.929$^{*}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>17</td>
<td>189</td>
<td>0.540</td>
<td>4.849$^{***}$</td>
<td>-0.042</td>
<td>-0.185</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>5</td>
<td>89</td>
<td>0.740</td>
<td>3.830$^{***}$</td>
<td>0.158</td>
<td>0.569</td>
<td></td>
</tr>
</tbody>
</table>

#### Study characteristics

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th># IS</th>
<th># ES</th>
<th>$\beta_0$ (mean $d$)</th>
<th>$t_0$</th>
<th>$\beta_1$</th>
<th>$t_1$</th>
<th>$F(df_1,df_2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
<td>31</td>
<td>362</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.871 (1, 360)</td>
</tr>
<tr>
<td>Yes (RC)</td>
<td>10</td>
<td>147</td>
<td>0.501</td>
<td>3.793$^{***}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>21</td>
<td>215</td>
<td>0.654</td>
<td>6.692$^{***}$</td>
<td>0.153</td>
<td>0.933</td>
<td></td>
</tr>
<tr>
<td><strong>Continent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America (RC)</td>
<td>28</td>
<td>308</td>
<td>0.639</td>
<td>7.832$^{***}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>3</td>
<td>54</td>
<td>0.296</td>
<td>1.302</td>
<td>-0.343</td>
<td>-1.419</td>
<td></td>
</tr>
<tr>
<td><strong>Intention to Treat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.640 (1, 360)</td>
</tr>
<tr>
<td>Yes (RC)</td>
<td>6</td>
<td>78</td>
<td>0.725</td>
<td>4.132$^{***}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>25</td>
<td>284</td>
<td>0.568</td>
<td>6.454$^{***}$</td>
<td>-0.157</td>
<td>-0.800</td>
<td></td>
</tr>
<tr>
<td><strong>Authors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent (RC)</td>
<td>17</td>
<td>230</td>
<td>0.608</td>
<td>5.757$^{***}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent</td>
<td>14</td>
<td>132</td>
<td>0.590</td>
<td>4.894$^{***}$</td>
<td>-0.018</td>
<td>-0.112</td>
<td></td>
</tr>
<tr>
<td><strong>Peer Reviewed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.612 (1, 360)$^{**}$</td>
</tr>
<tr>
<td>Yes (RC)</td>
<td>31</td>
<td>362</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23</td>
<td>234</td>
<td>0.754</td>
<td>9.909$^{***}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>128</td>
<td>0.214</td>
<td>1.803</td>
<td>-0.539</td>
<td>-3.823$^{***}$</td>
<td></td>
</tr>
</tbody>
</table>
**DISCUSSION**

A multilevel meta-analysis was performed to assess the effect of treatment on psychosocial functioning in juveniles with harmful sexual behavior, and to assess what variables have a moderating influence on treatment effects. An overall significant and moderate effect ($d = 0.60$) was found, indicating treatment to be effective in improving psychosocial functioning of juveniles with harmful sexual behavior. Although there were indications of publication bias, this did not significantly affect the overall results. Moderator analyses showed that treatment effects on atypical sexual arousal and empathy (a trend, $p < .10$) were smaller than treatment effects on other outcomes. Samples that contained more juveniles with similar age victims or a mixed type problem behavior pattern, including non-sexual problem behavior, showed relatively large effect sizes compared to samples with less juveniles with similar age victims. Non-established treatment yielded larger effect sizes than did established treatments. Articles published in peer reviewed journals showed relatively large
effect sizes compared to non-published articles. Finally, evaluation of treatment effects by professionals, compared to other types of assessment (e.g., self-report), and studies of weak quality yielded larger effect sizes.

The type of psychosocial construct measured needs discussion, since smaller effect sizes were found for atypical sexual arousal, which in 48% percent of the cases was measured by penile plethysmography, possibly of influence on this finding. Penile plethysmography has received critique on its validity and is considered unethical, because it involves violation of physical integrity and the use of illegal audio or imagery (Hunter & Lexier, 1998). To date, it is widely accepted that penile plethysmography does not adequately measure (reduction of) atypical sexual arousal, given that respondents without an atypical interest also respond physically to atypical sexual stimuli (Plaud & Blackstone, 2014). Furthermore, for stable forms of (harmful) atypical sexual interest (a very small subgroup among juveniles with harmful sexual behavior, see also Hunter, 2008 and Worling, 2013), it has been concluded that treatment should focus on learning to cope with the atypical sexual arousal pattern, because the successful remediation of atypical sexual interests has been found to be difficult (McManus, Hargreaves, Rainbow, & Alison, 2013; Wakefield, 2011). Notably, in the developmental stage of adolescence, sexual preferences are still fluid (Hunter, Figueredo, Malamuth, & Becker, 2003), and with time -not therapy- juveniles may also become aroused by other than atypical stimuli.

Enhancing moral development (i.e., learning not to harm others) is an important treatment goal for juveniles whose harmful sexual behavior was influenced by a paucity of (developmental on-target) moral development. However, only a relatively small impact was found on the improvement of empathy through treatment in this meta-analysis. Van Vugt (2011) found moral judgment, rather than (affective) empathy, to constitute a dynamic treatment need for juveniles with harmful sexual behavior. If an outcome measure assessed the (innate) ability to sense what someone else is feeling, that is, affective empathy instead of cognitive empathy (Van Outsem, 2009), it may represent a more trait-like, static factor, explaining lesser results of treatment on this construct. When a potentially harmful arousal pattern proves to be relatively stable, targeting the juvenile’s moral cognitions, such as moral judgment and cognitive empathy, might yield more positive treatment results (Van Langen et al., 2014; Van Vugt et al., 2011). The relatively larger effects of treatment on, for example, social and coping skills, emotion and self-image, and family functioning that were found in this meta-analysis are encouraging, in particular because these are considered protective factors for (sexually) harmful behavior through their positive effect on the well-being of a juvenile (Ward, 2012; Worling, 2013).
Two characteristics of the juveniles moderated the effect of treatment. Treatment groups with a higher percentage of juveniles with similar age victims showed larger effects of treatment on psychosocial functioning. Also, treatment groups with a higher percentage of juveniles with a mixed offending pattern (also non-sexual problem behavior) resulted in larger effect sizes. Juveniles with mixed type problem behavior and those with similar age victims generally show higher levels of conduct problems (Drew, 2013; Hendriks, 2006; Leroux et al., 2016). In contrast, intrapsychic/externalizing problems seem dominant in juveniles with younger victims (≥ 5 years younger and below the age of 12) (Hendriks & Bijleveld, 2004). The treatment of intrapsychic problems might take more time, since these problems may be linked to more pervasive problems, such as an insecure attachment (Miner, 2006, 2008), or developmental problems like a Pervasive Developmental Disorder (Hendriks, 2006). Additionally, the measurement of change in intrapsychic constructs is challenging (Tak, Bosch, Begeer, & Albrecht, 2014) which may have influenced the reported levels of change.

We found a moderator effect for treatment status. Non-established treatments yielded larger effects than established treatments, which is contrary to research findings on the effectiveness of treatment on reducing sexual recidivism. The latter has shown established treatments (i.e., cognitive behavioral based treatments, CBT) to be more effective than non-established treatment, often designated as treatment as usual (Dopp et al., 2015; Hanson et al., 2009; Lösel & Schmucker, 2005; Schmucker & Lösel, 2015; Walker et al., 2004). Treatments that were coded as non-established in the current study comprised cognitive therapy (including psycho-education), behavioral therapy (i.e., satiation and desensitization), and treatments making use of both cognitive and behavioral techniques, but next to each other instead of integrated. For example, one treatment offered psychotherapy alongside satiation. This was coded as CBT, but non-established, since the concept of established CBT includes social learning and a more holistic view on the origins of harmful sexual behavior; Ward, Polaschek, & Beech, 2006). Furthermore, psychotherapy and satiation are both non-established types of treatment.

As the effect of treatment status and type of treatment (a moderating trend; p < .10) may be related to study quality, we post-hoc tested the possible explanation of study quality being responsible for the unexpected moderator effects. A multivariate analysis with study quality as a covariate showed that the effects for treatment status and treatment type disappeared when controlling for study quality, which was the only significant moderator in the multivariate model. Therefore, it is concluded that study quality was the most important moderator of treatment effect. The use of more rigorous study designs in contemporary research may (partly) explain why the hypothesis of Caldwell (2016), stating that modern established treatments have become more effective, was not substantiated by the current
study. Becoming more effective in preventing recidivism, however, may not fully overlap with becoming more effective in the improvement of well-being. Both constructs seem related, but are not identical. For most treatments both aims are important and, therefore, should be addressed. The Good Lives Model (Fortune, Ward, & Print, in preparation), offers a promising paradigm which prominently addresses this dual aim of treatment for juveniles with harmful sexual behavior.

The judgment of treatment effects by professionals proved to result in larger effect sizes than did penile plethysmography, parent-report, or adolescent self-report. An explanation might be that the professionals involved in treatment, judging psychosocial functioning of the juveniles they treated, were biased by wanting their efforts to render an effect. Also, professionals might have been influenced by socially desirable behavior of the juveniles in treatment, and the results might represent a restricted (more positive) view of their clients’ behavior (Bryman, 2012). Our findings support earlier statements arguing against the (erroneous) assumption that these juveniles are deceitful (Worling, 2013), by yielding a smaller effect of treatment via parent report and self-report, arguably two methods usually thought to overestimate treatment effect in comparison to professional judgment. Study status also proved to have a moderating influence on treatment effects. Published studies reported better treatment results than did non published studies. This is in agreement with the file drawer effect (Rosenthal, 1979), which entails that studies with unfavorable results are published less frequently than studies with positive (treatment) results.

Study quality moderated the effects measured, proving studies of weak quality to yield larger effect sizes than studies of strong quality, which is in line with previous research findings (see Weisburd, Lum, & Petrosino, 2001). The study quality index that was used accounts for more than ‘just’ study design. Sample size and other study characteristics are part of the evaluation of study quality, providing a more comprehensive assessment of quality than study design alone (Armijo-Olivo et al., 2012), and rendering some quasi-experimental studies to be of strong quality. These studies also produced more modest effect sizes. Therefore, when isolating study design, randomized controlled trials did not produce significantly smaller results than did quasi-experimental research designs. The mean effect of all quasi-experimental studies may have been reduced by quasi-experimental studies of strong quality (rendering relatively smaller effects).
LIMITATIONS

In this study, it proved not possible to conduct a multivariate analysis with all significant moderators to examine the unique impact that moderators may have had, due to missing values. Where post-hoc testing was possible, this was conducted. Because of the diverse measurements of psychosocial constructs, and the differences in study quality of the included studies, this hampers the current research effort. More strict inclusion criteria could have reduced the impact of this limitation, but would have also much restricted the inclusion of all previous studies on a subject, a prerequisite for review studies. Since the current study is the first on this subject, the choice was made to include studies of lesser quality and a broad set of psychosocial outcome measures. This offers the reader a first, albeit exploratory, insight into the effects of treatment in general and an indication of the most promising psychosocial treatment goals to improve the well-being of juveniles with harmful sexual behavior.

The inclusion of older studies and the inclusion of mostly (83%) North American studies, limits generalizability (Bijleveld, 2015). Studies are conducted within a certain time frame and context, which especially influences studies on sexual problem behavior. What is considered atypical in some parts of the world may not be considered so in other parts. Also, time alters perceptions on normalcy of sexual behavior (in adolescence). Results, therefore, should be cautiously applied to other (especially non-Western) parts of the world.

The total sample size used in the current meta-analysis is restricted, since the included studies mainly consisted of small samples. Usage of small sample sizes is frequent in studies on juveniles with harmful sexual behavior (Fanniff & Kimonis, 2014) and a limited amount of studies are performed on this subject. Therefore, a thorough literature search was conducted that also included non-published studies. In addition, a three-level mixed effects model (Assink & Wibbelink, 2016) was used to maximize statistical power. It may be assumed that a relatively large amount of juveniles with harmful sexual behavior was included in the current analysis, creating substantial representativeness.

Finally, only few participant characteristics could be included in the moderator analyses, because not many studies reported on specific sample characteristics. The heterogeneity of juveniles with harmful sexual behavior demands a comprehensive reporting of sample characteristics to enable assessment of external validity of study results and to conduct moderator analyses to test intervention effects in subgroups of juveniles with harmful sexual behavior (Bijleveld, 2015).
CONCLUSION

Treatment aimed at psychosocial functioning of juveniles with harmful sexual behavior proved to be moderately effective. Surprisingly, we found no empirical evidence supporting that ‘modern’ CBT is most effective, possibly partly due to the use of more rigorous study designs in contemporary research. Treatment aimed at overall functioning, rule breaking and aggression, impulse control, social and coping skills, emotion and self-image, cognitions and sexual knowledge, and family functioning seems promising. Even if some of these factors have not (yet) been established as criminogenic factors, they represent real life problems of juveniles. The dominant RNR model has been critiqued for providing a too restrictive, risk focused, view on offender rehabilitation (Newsome & Cullen, 2017; Ward, Yates, & Willis, 2012). Recent developments in treatment methods describe a return to more holistic treatment frameworks (Dopp et al., 2015; Ward & Maruna, 2007; Worling, 2013). For juveniles with harmful sexual behavior this might prove especially relevant, since (sexual) recidivism rates are low, recidivism risk factors are hard to establish, and risk assessment instruments often overestimate the actual recidivism risk of juveniles (Hempel, 2013). Future research into the improvement of psychosocial functioning (i.e., well-being) of juveniles with harmful sexual behavior should further operationalize well-being as an outcome measure, if possible establish its link with desistance of problem behavior, and distinguish between relevant typologies. This will contribute to general knowledge on what treatment might prove the best fit for what (type of) sexually harmful juvenile.
Primum Non Nocere: A Comparative Study of Differences and Changes in Treatment Needs of Juveniles with Harmful Sexual Behavior Receiving Residential or Non-Residential Treatment

Ter Beek, E.
Kuiper, C.H.Z.
Van der Rijken, R.E.A.
Van Horn, J. E.
Hendriks, J.
Stams, G.J.J.M.

Manuscript submitted for publication
ABSTRACT

To date, research comparing the differential impact of residential and non-residential treatment on juveniles with problem behavior has not included adolescents with harmful sexual behavior. This study compares cognitive-behavioral based residential to non-residential (multi-systemic contextual) treatment in the Netherlands by administration of the Adolescent Sexual abuser Assessment Pack – Dutch version (ASAP-D) to 36 adolescents (M = 14.3 years; SD = 1.5) with harmful sexual behavior at admittance and after treatment completion. Through calculation of Reliable Change Indices, it was examined whether changes in psychosocial treatment needs after residential treatment differed from changes after non-residential treatment. Results indicate that juveniles in residential treatment represented a group at higher risk for recidivism, with greater (self-reported) treatment needs compared to juveniles in non-residential treatment. For those juveniles with equal levels of treatment needs, residential cognitive behavioral based treatment and contextual multi-systemic treatment obtained equal positive changes in psychosocial treatment needs of juveniles with harmful sexual behavior. Some juveniles, however, showed significant deterioration. Suggestions for the improvement of treatment results are discussed.
INTRODUCTION

Juveniles placed in residential (24-hour) youth care have been found to compare unfavorably to juveniles in non-residential youth care regarding their mental health and behavioral problems (Baker, Kurland, Curtis, Alexander, & Papa-Lentini, 2007; Curtis, Alexander, & Lunghofer, 2001). Most juveniles in residential care have been diagnosed with severe problems, such as serious delays in moral development (Stams et al., 2006), aggressive behavior, and mental disorders (Colins et al., 2010; Fairchild, Van Goozen, Calder, & Goodyer, 2013; Hoeve et al., 2015; Vachon, Kruger, Rogosch, & Cicchetti, 2015). Although there is consensus that these juveniles need treatment, the effectiveness of residential treatment targeting problematic functioning of children and adolescents, in particular in secure (judicial) institutions, is a subject of continuous debate (Parhar, Wormith, Derkzen, & Beauregard, 2008; Souverein, Van der Helm, & Stams, 2013).

From the perspective of the Risk Need Responsivity model (RNR-model; Bonta & Andrews, 2007), treatment is most effective if the following basic principles are met: (1) the intensity of care or treatment matches the risk for criminal offense recidivism (risk principle), (2) it targets dynamic criminogenic needs specific to the offenders (need principle), and (3) the delivered care or treatment is responsive to the individual characteristics (e.g., cultural background, gender, age, and cognitive abilities) and treatment motivation (responsivity principle). In general, the severity of problems of juvenile offenders (i.e., criminogenic treatment needs) warrants residential placement, in accordance with the risk principle. Residential treatment for juvenile offenders following the RNR principles has, however, been shown to be notably less effective than non-residential community-based treatment following the RNR principles. Residential treatment that does not adhere to any of the RNR principles, may even result in increased recidivism rates (Andrews & Bonta, 2010).

Recently, Mathys (2017) conducted a narrative review of the effectiveness of secure (residential) youth facilities and concluded that coercion into treatment, deprivation of the natural social context through residential placement, along with detrimental effects of sanctioning (e.g., learned helplessness, retaliatory aggression), are negative aspects of mandated residential placements (Andrews & Bonta, 2006; Gatti, Tremblay, & Vitaro, 2009; Mendel, 2011; Parhar et al., 2008). De Valk and colleagues (2016) have shown the mandated aspect of residential treatment to be able to transform into repression under particular circumstances. Also, iatrogenic effects (e.g., deviancy training) may occur when juveniles with similar problems are grouped for therapeutic reasons (Dishion, Poulin, & Burraston, 2001; Mathys, Hyde, Shaw, & Born, 2013).
In contrast, a recent meta-analysis of 27 controlled studies, including 17,038 juveniles comparing the effectiveness of institutional to non-institutional residential youth care (De Swart et al., 2012) found a similar positive effect for residential and non-residential treatment on delinquency, behavior problems, skills and a mixed category of miscellaneous problems. Additionally, De Swart and colleagues (2012) showed evidence-based residential treatment (i.e., a structured and manual-guided treatment that is based on empirical evidence), to be more effective than residential ‘treatment as usual’. A meta-analysis by Strijbosch and others (2015) including 19 controlled studies (15,526 participants), and 63 effect sizes on behavior problems (externalizing, internalizing, and total), skills (social and cognitive) and delinquency, also showed no differences between the effects of evidence-based residential and non-residential care. Non-residential treatment as usual (TAU) did gain better results as compared to residential TAU.

An explanation for these relatively positive results of some forms of residential care is also offered by Mathys (2017). She has postulated three components of effective mandated residential care: (1) the use of (established) cognitive behavioral interventions that build on strengths, (2) a positive group climate based on supportive relationships, and (3) a focus on motivating and improving the internal locus of control of juveniles. This explanation concurs with the conclusion of a narrative review on the effects of residential versus non-residential treatment by Souverein, Van der Helm, and Stams (2013). They argued that possible negative side effects of residential care specifically might be overcome by a therapeutic group climate, high responsivity of staff members, fostering positive peer relationships, and the use of evidence-based treatments and aftercare (see James, Stams, Asscher, Van der Laan, & De Roo, 2013).

To date, research comparing positive changes in juveniles receiving residential versus non-residential treatment has not included juveniles with harmful sexual behavior. This study aims to contribute to the literature on the residential treatment of juveniles with harmful sexual behavior by a) examining their level of psychosocial treatment needs as measured by the Adolescent Sexual abuser Assessment Pack – Dutch version (ASAP-D; see also Table 1) at admittance to treatment, and b) examining the treatment outcomes on the reduction of the aforementioned needs, obtained by cognitive-behavioral based residential treatment in comparison with non-residential multi-systemic (contextual) treatment, both considered established treatments.

**Treatment foci for juveniles with harmful sexual behavior**

Fanniff and Kimonis (2014) and Seto and Lalumiere (2010) have recently shown juveniles with harmful sexual behavior in general to differ from juveniles who display non-sexual problem behavior regarding their treatment needs. Therefore, specific etiological
explanations may apply. Empirically driven etiological theories suggest that varied factors can contribute to sexually harmful behaviors, including adverse (early) life experiences, atypical social development, disturbed sexual development, cognitive dysfunctions, or maladaptive schemes about self and other (Ryan et al., 2010), which convert into specific treatment needs. Notably, however, a fair amount of juveniles with harmful sexual behavior do not show personality or behavior characteristics that set them apart from non-problematic juveniles (Ryan et al., 2010; Van Outsem, 2009), and there are several similar antisocial tendencies between juvenile sexual and non-sexual offenders (Fanniff & Kimonis, 2014; Ronis & Borduin, 2007; Seto & Lalumiere, 2010).

Regarding treatment needs, Ryan and colleagues (2010) describe four groups: 1) juveniles who seem relatively non-problematic, 2) juveniles who mainly show internalizing problems and social deficits; 3) juveniles presenting with more generally antisocial and externalizing problem behavior and, 4) a very small group of juveniles who seem to portray an early onset pedophilic interest (Ryan et al., 2010). The first and third group, respectively, resemble the adolescent limited and life course persistent juvenile delinquent as described in Moffit’s developmental taxonomy (Moffit, 1993). Research has established a rather large subgroup of juveniles with harmful sexual behavior to fit into the adolescent limited subgroup (Becker & Kaplan, 1988; Cale et al., 2016; Hunter et al., 2003; Lussier et al., 2012; Richardson, Kelly, Graham, & Bhate, 2004). Marshall and Barbaree (2008) have previously stated that sexually transgressive behavior in adolescence could be considered ‘norm’ rather than ‘deviance’ since (sexual) experimentation and rule breaking are a normal aspect of adolescence. Wilson and Daly (1985) have previously called this ‘the young male syndrome’. Van Outsem (2009) also postulated that harmful sexual behavior in adolescence may stem from regular adolescent sexual impulses. Absent contextual inhibitions and opportunity may facilitate actual acting on these impulses instead of (severe) socio-psychological dysfunction.

In contrast, the much smaller life course persistent subgroup (Becker & Kaplan, 1988; Lussier et al., 2012) seems more fundamentally antisocial, callous-unemotional, aggressive and/or coercive. Sexual abuse of peers seems more common in this group than the abuse of small children (Richardson et al., 2004; Worling, 2001). Non-sexual recidivism by these more ‘generalist’ transgressive juveniles has been shown to be relatively high and stable until adulthood (Butler & Seto, 2002; Chu & Thomas, 2010; Drew, 2013; Hissel et al., 2006; Parks & Bard, 2006).

A subgroup of more internalizing and socially impaired juveniles seems to more often molest young children, as their fundamental drive seems intimacy, acceptance, safety, and connection, which they are not able to derive from peer contacts (Becker, 1988; Hunter, Figueredo, Malamuth, & Becker, 2003; Hunter, 2008; Miner et al., 2016). These juveniles
usually present a high self-reported and specific treatment need, as they experience social isolation, inadequacy, anxiety, and/or depression (Carpenter, Peed, & Eastman, 1995; Hunter et al., 2003; Ryan et al., 2010).

Finally, Hunter (2008) has described a very small subgroup of juveniles displaying an early onset pedophilic interest. He also categorizes these juveniles either as ‘antisocial’ or as ‘non-antisocial’. The first with a greater risk for (adolescent and adult) sexual recidivism, especially in child abuse, and the latter with a more favorable prognoses.

In sum, juveniles with harmful sexual behavior are considered a heterogeneous group with highly individualized levels and types of treatment needs (Fanniff & Kimonis, 2014; Barbaree & Marshall, 2008; Ryan, Leversee, & Lane, 2010; Hendriks, 2006; Seto & Lalumiere, 2010). For that reason, and in line with the RNR model (Bonta & Andrews, 2007), an individual assessment of treatment needs is necessary to determine if (specific) treatment is warranted, how it should be delivered (e.g., residential or non-residential), and on what psychosocial aspects treatment should mainly focus.

Modern cognitive behavioral based treatments (CBT) for juveniles with harmful sexual behavior incorporate different theories and strategies, adjusting their treatment strategy to the individual treatment needs as assessed. CBT has, via meta-analyses, been proven effective in reducing sexual and non-sexual recidivism in juveniles with harmful sexual behavior (Hanson, Bourgon, Helmus, & Hodgson, 2009; Lösel & Schmucker, 2005; Schmucker & Lösel 2015; Walker, McGovern, Poey, & Otis, 2004; Ter Beek et al., 2017). Because (sexual) recidivism in juveniles has been linked to several psychosocial treatment needs (i.e., aggression, impulsiveness, and antisocial beliefs), childhood trauma (i.e., paternal abandonment and childhood sexual victimization), sexual deviation (i.e., interest in prepubescent children or sexual violence), lack of intimate peer relationships and social isolation, and poor school performance (Carpentier & Proulx, 2011; Christiansen & Vincent, 2013; Fanniff & Kimonis, 2014; Seto & Lalumière, 2010; Worling & Långström, 2003), the assumption is made that CBT based treatments are also (successfully) targeting these criminogenic psychosocial treatment needs.

**STUDY AIM**

This study aims to contribute to the literature on the treatment of juveniles with harmful sexual behavior in residential youth care by describing treatment needs at admittance as measured by the ASAP-D (see also Table 1) and treatment outcomes, obtained by means of cognitive-behavioral based residential treatment compared to non-residential multi-systemic (contextual) treatment. By offering an analysis of typology, diagnosis, and
Treatment Needs; differences and changes

Psychosocial treatment needs at admission, this study first examines the treatment needs of juveniles allocated to residential versus contextual care for juveniles with harmful sexual behavior in the Netherlands. Second, after matching those juveniles with a similar (high) treatment need, the observed reliable change is calculated to explore the results of treatments offered. We specifically examine whether changes in treatment needs after residential treatment differ from changes after non-residential treatment. Notably, differences in these changes may provide important information to increase the effectiveness of treatment.

**Hypotheses**

Both of the researched treatment programs are specifically aimed at treating juveniles with harmful sexual behavior who are at risk for recidivism. The assumption is made that these juveniles suffer from complex underlying psychosocial problems, and that therefore intensive specialized treatment is necessary. We expect the level of treatment needs of juveniles at admission to differ, in accordance with the need and risk principles of the RNR-model (Bonta & Andrews, 2007), with higher levels of (criminogenic) treatment needs in residential care (see also, Ter Beek, Van der Rijken, Kuiper, Hendriks, & Stams, 2016). Therefore, we will match juveniles with equal levels of self-reported treatment needs in the comparison of treatment results. Based on de Swart and colleagues (2012) and Strijbosch and colleagues (2015) our hypothesis is that secure youth care will treat a more problematic target group than multi systemic therapy, but that after both treatments, equal positive changes in psychosocial treatment needs will be found.

**METHOD**

**Participants**

In total, 36 male juveniles who were allocated to the mandatory treatments researched, were included in this study. Their average age at admittance was 14.3 years (SD = 1.5). Their main treatment goals were linked to their sexual problem behavior. All participants and their parents gave their informed consent prior to participation. Juveniles received treatment from February 2012 (when measurements started) up until January 2016. From February 2012 up until November 2014, a total of 20 juveniles were allocated to treatment in residential secure youth care and asked to participate. One juvenile refused participation and two were excluded due to language difficulties (being unable to complete the measurements in Dutch). Of the remaining 17 juveniles, two did not complete a final measurement. These juveniles, therefore, could not be included in the calculation of the effect of treatment. A matched size comparison group was recruited from March 2014 until February 2015. In this period, a total of 21 juveniles were allocated to non-residential multi...
systemic therapy and asked to participate. One juvenile refused participation and one was excluded due to language difficulties (being unable to complete the measurements in Dutch), leaving 19 juveniles in the subsample. Two juveniles refused to complete a final assessment. They, therefore, were not included in the calculation of the effect of treatment. For an overview of the characteristics of the (sexual) problem behavior and diagnoses of the participants, see Table 2.

Researched Treatment Programs

Secure Youth Care for juveniles with Severe Sexual Problems (SYC-SSP)

For juveniles with harmful sexual behavior (aged 12 to 18 years), who have not been charged or have not been adjudicated, specialized treatment in a Secure Youth Care (SYC) facility is an option. Dutch law on (secure) youth care (Hirsch Ballin, 2007) states that restricting freedom of movement in a secure setting can only become an option if so advised by an independent psychologist, based on an assessment of recidivism risk and treatment needs of the juvenile. Therefore, a placement in SYC is only possible via a civil court order. In the Netherlands, one SYC-SSP treatment group (10 beds) is available specifically for juveniles with harmful sexual behavior. During their stay, cognitive behavioral based group treatment is offered by youth care professionals, and monitored by a behavioral specialist. Social learning strategies are explicitly used to stimulate change, and the specific interventions applied are tailored to the juvenile's treatment needs. At admittance, (criminogenic) treatment needs are assessed by means of the J-SOAP-D (Bullens, Van Horn, & Van Eck, 2012), and used for treatment planning purposes. Group therapy 'Out of the Circle' (Koster & Tel, 2010), a central part of the treatment offered, focusses on the sexual misconduct. It assesses the cycle of sexual aggression (triggers, thoughts, behaviors, and consequences), and creates an individualized risk management plan. This plan is presented by the juvenile to his parents/social network, providing insight into the harmful sexual behavior and its origins. Individual therapy (as to address possible victimization experiences or trauma) may supplement the group treatment offered. Contraindication for placement is an IQ < 70; these juveniles receive treatment in a secure institution for the mentally disabled.

Multi Systemic Therapy - Problem Sexual Behavior (MST-PSB)

MST-PSB is an adaptation of multi systemic therapy (Dwyer & Letourneau, 2011). It is aimed at adjudicated and non-adjudicated juveniles aged 10 to 18 years, who have exhibited harmful sexual behavior, have complex (family) issues, and pose a risk to society (Boonstra & Van der Rijken, 2010). Since MST-PSB is a community-based type of treatment, the family system needs to actively take responsibility for safety issues during the at home treatment. MST-PSB is based on regular MST, and targets youth and family problems within and between the multiple systems in which the family members are embedded. The overriding
Treatment Needs; differences and changes

The goal of MST is to empower parents and adolescents with the skills and resources needed to cope with their familial and extra-familial problems. Using treatment strategies derived from strategic family therapy, structural family therapy, behavioral parent training, and cognitive-behavioral therapy, MST addresses intrapersonal (e.g., cognitive), familial, and extra familial (i.e., peer, school, neighborhood) factors that are known to be linked with a youths’ serious antisocial behavior, including sexual offending. If, and when, specific individual interventions are used to modify the juvenile’s social perspective-taking skills, belief system, or attitudes believed to contribute to offending and the sexual assault cycle, the parent is reinforced as the change agent, and is directly involved in the intervention. The exact nature of interventions applied varies for each family, depending on their strengths and weaknesses (Borduin & Schaeffer, 2001). Contraindications for the Dutch MST-PSB program are an IQ < 70, acute psychiatric problems, and severe forms of autism (Boonstra & Van der Rijken, 2010).

Procedure

Using a pre-post evaluation design, juveniles were asked to complete pen and paper self-report measures at admission, and again at their completion of the program. A test assistant or counselor asked them to fill in the forms and collected the results. The adult supervised the juvenile during the tests / remained available for questions. The mean treatment exposure for SYC-SSP was 16 months (SD = 6.8), for MST-PSB this was 6 months (SD = 1.1), a significant difference (F = 12.989, p = .001).

Measures

For the comparison of the two subsamples on typology and diagnosis, it was coded if at admission juveniles had shown repeat sexually harmful behavior or not, what type of sexual problem behavior they had shown (contact, non-contact or both), whether they had also shown other types of transgressive behavior (i.e., sexplus problem behavior), the age range of their victim(s) and, finally, their DSM-classification(s). Case file information was used to code these items.

In the United Kingdom, the Adolescent Sexual Abuser Project (ASAP) developed the A-SOAP (Adolescent Sex Offender Assessment Pack), which allows for the measurement of (aspects of) specific individual factors that are considered relevant for the development of harmful sexual behavior in adolescents (i.e., hypersexuality and sexual deviance, anti-social behavior, callous-unemotional traits, intimacy and friendship deficits, and victimization; Beckett, Brown, & Gerhold, 1997; Bromberg & O’Donohue, 2014). In cooperation with the European Society for Working with Sexually Abusive Youth (ESSAY), a Dutch version of the A-SOAP, the ASAP-D, was developed (Van Outsem et al., 2006). Because of the availability of a
Dutch validated version, the ASAP-D was used in this study. It consists of nine psychometric instruments, offering a broad measure of individual psychosocial functioning. Table 1 provides an overview of the characteristics it measures.

Table 1. Overview of the psychometric ability of the ASAP-D

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Instrumentsa</th>
<th>Dimensions measured</th>
<th># of items &amp; response options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Desirability</td>
<td>Leiden Social Desirability Scale</td>
<td>-</td>
<td>11, 5 point scale</td>
</tr>
<tr>
<td>Self Esteem</td>
<td>Thornton’s Self-esteem / Self Derogation Scale</td>
<td>-</td>
<td>12, dichotomous</td>
</tr>
<tr>
<td>Emotional Loneliness</td>
<td>Emotional Loneliness Scale by Russel and colleagues</td>
<td>-</td>
<td>20, 4 point scale</td>
</tr>
<tr>
<td>Empathy</td>
<td>Interpersonal Reactivity Index by Davis</td>
<td>perspective taking, fantasy and emotional comprehension</td>
<td>28, 5 point scale</td>
</tr>
<tr>
<td>Locus of Control</td>
<td>Locus of Control Scale by Nowicky</td>
<td>-</td>
<td>40, dichotomous</td>
</tr>
<tr>
<td>Aggression</td>
<td>Leiden Aggression Questionnaire</td>
<td>reactive aggression and character aggression</td>
<td>14, 3 point scale</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>Eysenck Personality Questionnaire / Impulsivity Scale</td>
<td>-</td>
<td>22, dichotomous</td>
</tr>
<tr>
<td>Cognitive Distortions</td>
<td>Beckett’s Children and Sex Scales</td>
<td>total quantity of cognitive distortions, justifications, child sexual maturity, self-identification as a child, mutual special relations and ideation of attractiveness</td>
<td>84, 5 point scale</td>
</tr>
<tr>
<td>Sexuality</td>
<td>Leiden Sexuality Questionnaire</td>
<td>egodystonic hypersexuality and positive orientation towards sexuality</td>
<td>17, dichotomous</td>
</tr>
</tbody>
</table>

a (revised) Dutch versions by Van Outsem et al., 2006

The ASAP-D is deemed suitable for adolescents aged 12-20 years of age, with an IQ of 80 and higher, therefore, in the case of borderline intellectual functioning (IQ 70-80), the ability to complete the forms was to be assessed and supported by a professional. The ASAP-D was validated making use of various samples: juveniles who had offended sexually (N = 146), non-sexually but violent offenders (N = 153), and a representative sample of non-delinquent Dutch youth (N = 500). Validity and reliability of the ASAP-D were considered sufficient (Cronbach’s α internal consistency reliabilities between .70 and .93; test-retest correlations between .70 and .98; Van Outsem et al., 2006).
Calculations

Testing the significance of change at the individual level is difficult. Many authors have highlighted a range of issues. The Reliable Change Index (Jacobson & Truax, 1991) focuses on individual change over time and, in the context of treatment, refers to whether an individual improved, whether that improvement may be considered reliable and, whether that change puts the client’s functioning closer to the functioning of ‘normal controls’. Reliable change indices (RCI’s) with a magnitude of $z \geq 1.96$ can be considered statistically significant at the $p < 0.05$ level (two-tailed). They represent changes over and above what might be accounted for by the variability of the measure. An RCI smaller than $z = 1.96$ may reflect a change, but this change may also be accounted for by measurement variability (Jacobson & Truax, 1991). Thus, to statistically assess individuals’ improvement or deterioration, the reliable change index (RCI) was computed:

$$\text{RCI} = \frac{T_2 - T_1}{S_{\text{diff}}}$$

$$S_{\text{diff}} = \sqrt{2 (S_e)^2}$$

$$S_e = Sd \sqrt{1 - r_{xx}}$$

Test-retest reliability coefficients of the ASAP-D varied from $\alpha = .70$ to .98 (Van Outsem et al., 2006). Because Van Outsem did not report reliability coefficients per scale, the most conservative available measure of measurement variability ($r_{xx} = .70$) was used to compute the RCI. The standard deviation ($SD$) of the norm group was available per scale, and was used.

Notably, for the investigated juvenile to be able to become more similar to the normal functioning group, the first measurement must have placed the juvenile in a (sub)clinical range of functioning. As described, not all sexually transgressive juveniles statistically differ psychologically from the normal population (Ryan et al., 2010). Therefore, only juveniles scoring in the (sub)clinical range of a scale (> 70th or < 30th percentile) were included in the analyses of individual improvement (via calculation of the RCI). Hypothetically, juveniles may have also reported a statistically significant deterioration. In the analysis of deterioration (via calculation of the RCI) all subjects were included, because significant deterioration is considered detrimental to all respondents.
To assess group differences in treatment needs at admission across the two intervention groups, a group comparison of typology, diagnosis, and subclinical scores was offered first. We expected a relatively more problematic group to be treated in SYC-SSP, so the hypotheses were tested one-sided. Whenever sample size proved too small for Chi Square, the more conservative Fischer’s Exact Test was used ( Brace, Kemp, & Snelgar, 2000).

RESULTS

In Table 2, a comparison of typology and diagnosis is presented. In SYC-SSC more repeat offending, $p = .001$, and more sex-plus offending, $p = .002$, juveniles were treated. Also, as compared to MST-PSB, in SYC-SCC more juveniles with a primary diagnosis of a pervasive or reactive attachment disorder were treated, whereas in MST-PSB more juveniles with a primary diagnosis of disruptive disorders were treated, exact $p = .003$.

Additionally, Table 3 presents a group comparison of self-reported subclinical treatment needs at admission. Juveniles allocated to SYC-SSP reported a positive orientation towards sex more often, exact $p = .008$ (indicating a higher than average interest in sexuality), to be more emotionally lonely, $p = .043$, and more impulsive, $p = .032$. On the other scales no significant differences in (sub)clinical values were reported. Most juveniles in treatment reported low self-esteem issues (72%). In SYC-SSP specifically, emotional loneliness (65%) and impulsivity (59%) also proved relatively often reported treatment needs.

On the effect of treatment for juveniles with self-reported (sub)clinical treatment needs, Table 4 shows no differences between both programs in changes achieved. Notably, however, many juveniles who reported a (sub)clinical treatment need reported no significant change, and thus, no improvement. Most of the juveniles reporting themselves as functioning within the normal range at admission (see Table 5) reported no significant change, which is to be interpreted as staying within the normal range of functioning. However, some negative side-effects were also discovered, as some of these juveniles reported a significant deterioration, placing them more towards the level of functioning of juveniles with psychosocial problems.
Table 2. Typology and Diagnosis at admission

<table>
<thead>
<tr>
<th></th>
<th>MST-PSB (N = 19)</th>
<th>SYC-SSP (N = 17)</th>
<th>Significance T or χ² or E (p, one sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age (Sd)</td>
<td>13.7 (1.4)</td>
<td>14.8 (1.4)</td>
<td>.357 (544)²</td>
</tr>
<tr>
<td>% First offenders</td>
<td>63.2 (12)</td>
<td>11.8 (2)</td>
<td>9.972 (.001)***</td>
</tr>
<tr>
<td>% Repeated sexual problem behavior</td>
<td>36.8 (7)</td>
<td>88.2 (15)</td>
<td></td>
</tr>
<tr>
<td>% Contact sexual problem behavior</td>
<td>68.4 (13)</td>
<td>82.4 (14)</td>
<td>3.911 (.084)²</td>
</tr>
<tr>
<td>% Non-contact sex. problem behavior</td>
<td>241.1 (4)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>% Mixed sexual problem behavior</td>
<td>10.5 (2)</td>
<td>17.6 (3)</td>
<td></td>
</tr>
<tr>
<td>% Sex only problem behavior</td>
<td>84.2 (16)</td>
<td>35.3 (6)</td>
<td>9.034 (.002)***</td>
</tr>
<tr>
<td>% Sex plus problem behavior</td>
<td>15.8 (3)</td>
<td>64.7 (11)</td>
<td></td>
</tr>
<tr>
<td>% Only child victims &lt; 12y and 5y ↓</td>
<td>47.4 (9)</td>
<td>29.4 (5)</td>
<td>2.074 (.174)²</td>
</tr>
<tr>
<td>% Only peer victims</td>
<td>36.8 (7)</td>
<td>35.3 (6)</td>
<td></td>
</tr>
<tr>
<td>% Peer and child victims</td>
<td>15.8 (3)</td>
<td>35.3 (6)</td>
<td></td>
</tr>
<tr>
<td>% Pervasive disorder (incl. PDD-nos)a</td>
<td>0 (0)</td>
<td>17.6 (3)</td>
<td>9.286 (.003)**</td>
</tr>
<tr>
<td>% Disruptive disorder (incl. ADHD /CD)a</td>
<td>89.5 (17)</td>
<td>41.2 (7)</td>
<td></td>
</tr>
<tr>
<td>% Reactive attachment disordera</td>
<td>10.5 (2)</td>
<td>41.2 (7)</td>
<td></td>
</tr>
</tbody>
</table>

* the primary diagnosis was coded, several youths were classified with more than one diagnosis. ** Significant at a p .05 level. *** Significant at a p .01 level. "" Significant at a p .001 level. T-test. E Fishers Exact Test.

Table 3. Group comparison percentage subclinical scores at admission

<table>
<thead>
<tr>
<th>ASAP-D scale</th>
<th>MST (19) % (N)</th>
<th>SYC (17) % (N)</th>
<th>Total (36) % (N)</th>
<th>Significance χ² (p, one sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Desirability¹</td>
<td>26.3 (5)</td>
<td>23.5 (4)</td>
<td>25.0 (9)</td>
<td>0.037 (577)³</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>73.7 (14)</td>
<td>70.6 (12)</td>
<td>72.2 (26)</td>
<td>.043 (564)³</td>
</tr>
<tr>
<td>Emotional Loneliness</td>
<td>35.5 (6)</td>
<td>64.6 (11)</td>
<td>-</td>
<td>2.941 (043)²</td>
</tr>
<tr>
<td>Locus of Control</td>
<td>29.4 (5)</td>
<td>41.2 (7)</td>
<td>34.4 (11)</td>
<td>.515 (237)</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>27.8 (5)</td>
<td>58.8 (10)</td>
<td>-</td>
<td>3.441 (032)²</td>
</tr>
<tr>
<td>Empathy perspective taking</td>
<td>31.6 (6)</td>
<td>41.2 (7)</td>
<td>36.1 (13)</td>
<td>.358 (275)</td>
</tr>
<tr>
<td>Empathy fantasy</td>
<td>47.4 (9)</td>
<td>29.4 (5)</td>
<td>38.9 (14)</td>
<td>1.217 (135)</td>
</tr>
<tr>
<td>Empathy emo. comprehension</td>
<td>10.5 (2)</td>
<td>17.6 (3)</td>
<td>13.9 (5)</td>
<td>.380 (445)³</td>
</tr>
<tr>
<td>Reactive Aggression</td>
<td>21.1 (4)</td>
<td>29.4 (5)</td>
<td>25.0 (9)</td>
<td>.334 (423)²</td>
</tr>
<tr>
<td>Character Aggression</td>
<td>10.5 (2)</td>
<td>17.6 (3)</td>
<td>13.9 (5)</td>
<td>.380 (445)³</td>
</tr>
<tr>
<td>Hypersexuality</td>
<td>33.3 (6)</td>
<td>47.1 (8)</td>
<td>40.0 (14)</td>
<td>.686 (204)</td>
</tr>
<tr>
<td>Positive orientation towards Sex</td>
<td>0.0 (0)</td>
<td>35.3 (6)</td>
<td>-</td>
<td>7.667 (008) **</td>
</tr>
<tr>
<td>Cognitive Distortions</td>
<td>16.7 (3)</td>
<td>17.6 (3)</td>
<td>17.1 (6)</td>
<td>.0066 (642)²</td>
</tr>
</tbody>
</table>

¹ A high score indicates a possible effect on the measurement reliability of the other scales. ** Significant on a p .05 level. *** Significant on a p .01 level. T-test. E Fishers Exact Test.
Table 4. Reliable Change Index with Pearson Chi-Square test for association with setting in the (sub)clinical range of the population

<table>
<thead>
<tr>
<th>ASAP D Scales</th>
<th>% improvement</th>
<th>% no change</th>
<th>% deterioration</th>
<th>$\chi^2$ (p, two sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RCI &gt; 1.96 (n)</td>
<td>1.96≥RCI≥−1.96 (n)</td>
<td>RCI &lt; - 1.96 (n)</td>
<td></td>
</tr>
<tr>
<td>Selfesteem</td>
<td>MST (n=14)</td>
<td>21.4 (3)</td>
<td>78.6 (11)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SYC (n=11)</td>
<td>9.1 (1)</td>
<td>81.8 (9)</td>
<td>9.1 (1)</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>16 (4)</td>
<td>80 (20)</td>
<td>4 (1)</td>
</tr>
<tr>
<td>Emotional Loneliness</td>
<td>MST (n=5)</td>
<td>40 (2)</td>
<td>60 (3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SYC (n=11)</td>
<td>45.5 (5)</td>
<td>36.4 (4)</td>
<td>18.2 (2)</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>43.8 (7)</td>
<td>43.8% (7)</td>
<td>12.5 (2)</td>
</tr>
<tr>
<td>Locus of Control</td>
<td>MST (n=4)</td>
<td>0</td>
<td>75 (3)</td>
<td>25 (1)</td>
</tr>
<tr>
<td></td>
<td>SYC (n=6)</td>
<td>16.7 (1)</td>
<td>33.3 (2)</td>
<td>50 (3)</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>10 (1)</td>
<td>50 (5)</td>
<td>40 (4)</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>MST (n=5)</td>
<td>20 (1)</td>
<td>80 (4)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SYC (n=9)</td>
<td>44.4 (4)</td>
<td>55.6 (5)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>64.3 (9)</td>
<td>35.7 (5)</td>
<td>0</td>
</tr>
<tr>
<td>Empathy persp. tak.</td>
<td>MST (n=6)</td>
<td>16.7 (1)</td>
<td>66.7 (4)</td>
<td>16.7 (1)</td>
</tr>
<tr>
<td></td>
<td>SYC (n=6)</td>
<td>33.3 (2)</td>
<td>66.7 (4)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>25 (3)</td>
<td>66.7 (8)</td>
<td>8.3 (1)</td>
</tr>
<tr>
<td>Empathy fantasy</td>
<td>MST (n=8)</td>
<td>12.5 (1)</td>
<td>87.5 (7)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SYC (n=4)</td>
<td>25 (1)</td>
<td>75 (3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>16.7 (2)</td>
<td>83.3 (10)</td>
<td>0</td>
</tr>
<tr>
<td>Empathy emo. comp.</td>
<td>MST (n=8)</td>
<td>25 (2)</td>
<td>75 (6)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SYC (n=3)</td>
<td>0</td>
<td>100 (3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>18.2 (2)</td>
<td>81.8 (9)</td>
<td>0</td>
</tr>
<tr>
<td>Reactive Aggression</td>
<td>MST (n=4)</td>
<td>0</td>
<td>100 (4)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SYC (n=5)</td>
<td>20 (1)</td>
<td>80 (4)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>11.1 (1)</td>
<td>88.9 (8)</td>
<td>0</td>
</tr>
</tbody>
</table>
Treatment Needs; differences and changes

<table>
<thead>
<tr>
<th>ASAP D Scales</th>
<th>% improvement RCI &gt; 1.96 (n)</th>
<th>% no change 1.96 ≥ RCI ≥ −1.96 (n)</th>
<th>% deterioration RCI &lt; −1.96 (n)</th>
<th>χ² (p, two sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Character Aggression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=2)</td>
<td>50 (1)</td>
<td>50 (1)</td>
<td>0</td>
<td>1.875 (p = .400)</td>
</tr>
<tr>
<td>SYC (n=3)</td>
<td>0</td>
<td>100 (3)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>20 (1)</td>
<td>80 (4)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Hypersexuality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=6)</td>
<td>16.7 (1)</td>
<td>83.3 (5)</td>
<td>0</td>
<td>.049 (p = 1.000)</td>
</tr>
<tr>
<td>SYC (n=8)</td>
<td>12.5 (1)</td>
<td>87.5 (7)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>14.3 (2)</td>
<td>85.7 (12)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Positive Orientation towards Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=14)</td>
<td>0</td>
<td>100 (14)</td>
<td>0</td>
<td>2.767 (p = .183)</td>
</tr>
<tr>
<td>SYC (n=11)</td>
<td>18.2 (2)</td>
<td>81.8 (9)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>8 (2)</td>
<td>92 (23)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cognitive Distortions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=3)</td>
<td>66.7 (2)</td>
<td>33.3 (1)</td>
<td>0</td>
<td>.667 (p = 1.000)</td>
</tr>
<tr>
<td>SYC (n=3)</td>
<td>33.3 (1)</td>
<td>66.7 (2)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>50 (3)</td>
<td>50 (3)</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

* Fishers Exact Test
### Table 5. Reliable Change Index with Pearson Chi-Square test for association with setting in the 'normal' range of the population

<table>
<thead>
<tr>
<th>ASAP D Scales</th>
<th>% improvement RCI &gt; 1.96 (n)</th>
<th>% no change 1.96 ≥ RCI ≥ −1.96 (n)</th>
<th>% deterioration RCI &lt; −1.96 (n)</th>
<th>χ² (p, two sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selfesteem</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=3)</td>
<td>0</td>
<td>100 (3)</td>
<td>0</td>
<td>2.100 (p = .429)</td>
</tr>
<tr>
<td>SYC (n=4)</td>
<td>0</td>
<td>50 (2)</td>
<td>50 (2)</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>0</td>
<td>71.4 (5)</td>
<td>28.6 (2)</td>
<td></td>
</tr>
<tr>
<td><strong>Emotional Loneliness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=10)</td>
<td>0</td>
<td>100 (10)</td>
<td>0</td>
<td>2.692 (p = .286)</td>
</tr>
<tr>
<td>SYC (n=4)</td>
<td>25 (1)</td>
<td>75 (3)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>7.1 (1)</td>
<td>92.9 (13)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Locus of Control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=11)</td>
<td>0</td>
<td>100 (11)</td>
<td>0</td>
<td>1.287 (p = .450)</td>
</tr>
<tr>
<td>SYC (n=9)</td>
<td>0</td>
<td>88.9 (8)</td>
<td>11.1 (1)</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>0</td>
<td>95 (19)</td>
<td>5 (1)</td>
<td></td>
</tr>
<tr>
<td><strong>Impulsivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=11)</td>
<td>9.1 (1)</td>
<td>81.8 (9)</td>
<td>9.1 (1)</td>
<td>3.527 (p = .163)</td>
</tr>
<tr>
<td>SYC (n=6)</td>
<td>0</td>
<td>50 (3)</td>
<td>50 (3)</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>5.9 (1)</td>
<td>70.6 (12)</td>
<td>23.5 (4)</td>
<td></td>
</tr>
<tr>
<td><strong>Empathy persp. tak.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=11)</td>
<td>9.1 (1)</td>
<td>72.7 (8)</td>
<td>18.2 (2)</td>
<td>1.182 (p = 1.000)</td>
</tr>
<tr>
<td>SYC (n=9)</td>
<td>0</td>
<td>88.9 (8)</td>
<td>11.1 (1)</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>5 (1)</td>
<td>80 (16)</td>
<td>15 (3)</td>
<td></td>
</tr>
<tr>
<td><strong>Empathy fantasy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=9)</td>
<td>0</td>
<td>77.8 (7)</td>
<td>22.2 (2)</td>
<td>1.940 (p = .591)</td>
</tr>
<tr>
<td>SYC (n=11)</td>
<td>18.2 (2)</td>
<td>72.7 (8)</td>
<td>9.1 (1)</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>10 (2)</td>
<td>75 (15)</td>
<td>15 (3)</td>
<td></td>
</tr>
<tr>
<td><strong>Empathy emo. comp.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=8)</td>
<td>0</td>
<td>75 (6)</td>
<td>25 (2)</td>
<td>1.617 (p = .728)</td>
</tr>
<tr>
<td>SYC (n=12)</td>
<td>8.3 (1)</td>
<td>83.3 (10)</td>
<td>8.3 (1)</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>5 (1)</td>
<td>80 (16)</td>
<td>15 (3)</td>
<td></td>
</tr>
<tr>
<td><strong>Reactive Aggression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MST (n=5)</td>
<td>0</td>
<td>100 (13)</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>SYC (n=5)</td>
<td>0</td>
<td>100 (10)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>0</td>
<td>100 (23)</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
DISCUSSION

This study aimed to contribute to the literature on the treatment of juveniles with harmful sexual behavior in residential youth care by describing treatment needs at admittance and treatment outcomes obtained by means of cognitive-behavioral based residential treatment, as compared to non-residential multi-systemic (contextual) treatment. Our findings on the differences in treatment needs of juveniles in residential SYC-SSP and non-residential MST-PSB were in line with our hypotheses. A relatively more problematic target group was allocated to SYC-SSP as compared to MST-PSB. More juveniles with repeating sexual problem behavior and more juveniles also displaying other forms of antisocial behavior were allocated to SYC-SSP. The common DSM classifications of juveniles in SYC-SSP were attachment disorder and pervasive disorders, whereas in MST-PSB juveniles with a primary diagnosis of disruptive problem behavior (i.e., conduct disorder and ADHD) were treated mostly. The diagnoses of juveniles in SYC-SSP represent pervasive problems in social and relational abilities, specifically linked to the development of sexual problem behavior (Baarsma et al., 2016; Fanniff & Kimonis, 2014; Miner et al., 2008, 2010, 2016; Seto...
& Lalumière, 2010). On psychosocial treatment needs, as measured by the ASAP-D (Van Outsem et al., 2006), juveniles in SYC-SSP reported higher levels of emotional loneliness, more impulsivity, and a higher interest in sex. These results are in line with prior research into the results of allocation to these specific treatments (Ter Beek et al., 2017a). The levels of these treatment needs for residentially allocated juveniles, theoretically, place them more at risk for recidivism than the juveniles allocated to contextual MST-PSB, which is in line with the RNR principles of effective judicial interventions (Bonta & Andrews, 2007), placing juveniles with higher levels of (criminogenic) treatment needs in residential care. The current allocation of juveniles to mandated treatment in the Netherlands, therefore, seems in line with the RNR model regarding levels of risk and criminogenic treatment needs.

Secondly, both MST-PSB and SYC-SSP aim to reduce specific psychosocial problems and make use of promising treatment frameworks (Hanson, Bourgon, Helmus, & Hodgson, 2009; Lösel & Schmucker, 2005; Schmucker & Lösel 2015; Walker et al., 2004). Therefore, both programs were expected to be as effective on most treatment needs, as measured by the ASAP-D. The main hypothesis of residential SYC-SSP being as effective as non-residential MST-PSB was substantiated. No differences were found when assessing the reliable change indices of most treatment needs. The significantly longer treatment exposure in residential SYC-SSP may be associated with the more pervasive diagnoses of juveniles at admission. The fact remains, however, that non-residential treatment reached similar results with less treatment exposure. This underscores the importance of allocation to non-residential treatment, whenever this is a possibility.

The assessed positive and negative changes in psychosocial treatment needs and their subsequent implications for clinical practice are discussed below. Results for the total group (SYC-SSP and MST-PSB) are discussed since no significant differences were found. Policy and practical implications are deemed to apply equally to residential as to non-residential treatment.

Few (sub)clinically scoring juveniles reported an improvement in self-esteem. Most (80%, N = 20) reported no change. Two juveniles in the normal range of functioning (29%) reported a deterioration in self-esteem. Low scores on this scale represent a negative self-image, uncertainty, dissatisfaction, and pessimism about the future (Van Outsem et al., 2008). In general, an increase in self-esteem is considered to be an important treatment aim. Attachment theory explains a connection between low self-image, as a result of early traumatic childhood experiences (such as neglect and abuse), and the development of coercive, antisocial, behavior patterns (Ainsworth, 1989; Bowlby, 1969, 1973; Hoeve et al, 2012). Also, developmental theories (Erikson, 1968; Freud, 1965; Piaget, 1928) stress the importance of completing developmental tasks and, therefore, feeling competent, in order
to prevent (further) dysfunctional behavior. However the mandatory administration of the studied treatments, might not provide the most fertile ground to improve self-esteem (Bonta & Andrews, 2007; Bonta, Wallace-Capretta, & Rooney, 2000; Cecile & Born, 2009; De Valk et al., 2016; Hendriks & Bijleveld, 2008). Hence, consequently, mandatory treatment may need to especially invest in improving the self-esteem of its clients. The Good Lives Model (Ward, Yates, & Willis, 2012), for example, explicitly focusses on improving general quality of life for juveniles (and adults) who have shown harmful sexual behavior. Consequently, it focusses less prominently on sexual problem behavior and risks (possibly worsening self-esteem) and more prominently on life-fulfillment and the achievement of a ‘good life’. In this way, the juvenile experiences himself as someone who can secure all primary human goods in socially acceptable and personally satisfying ways (Collie, Ward, Ayland, & West, 2007). The Good Lives Model also aims to target criminogenic needs, but may prove useful to explicitly counteract universal detrimental aspects of harmful sexual behavior that juvenile perpetrators face in general (i.e., stigma and social deprivation). This seems even more important since so very few of these juveniles have been found to actually repeat sexual offending or become pervasively sexually offensive (Caldwell, 2016; Lussier et al., 2012).

An equal percentage of sub-clinically scoring juveniles on emotional loneliness reported significant improvement or no improvement, while 13% (N = 2) reported a deterioration (feeling even more lonely). This might indicate that the examined treatments were not yet effective for all juveniles. Neither group therapy (with, i.a., acknowledgement by peers) nor systemic treatment (e.g., through improving family functioning) alone seems to have sufficiently decreased feelings of loneliness. Thus, more attention should possibly be paid to building (more intimate) relationships, moreover so, because relatively many juveniles self-reported this treatment need at admission. The Good Lives Model (Collie et al., 2007) might also prove an effective tool for improving treatment of feelings of loneliness, because it is based on becoming able to fulfill all basic needs (such as social belonging and intimacy). The ways towards becoming more able to interact socially or improve relationships might not differ, since the model incorporates all of the treatment strategies mentioned earlier, but the focus on this particular need may become more pronounced when the juveniles themselves deem it an important subject for their well-being.

Few sub-clinically scoring juveniles reported a significant improvement in locus of control post-treatment. Most reported no improvement (50%, N = 5), and others (40%, N = 4) even reported a deterioration (feeling less in control). These outcomes indicate a suboptimal treatment result for most sub-clinically scoring juveniles, while the aim of creating a risk management plan in therapy (concrete thoughts and behaviors the juvenile may use to prevent recidivism) is to increase the sense of self control the juvenile experiences. Being
under an intensive and mandatory form of supervision, as provided by SYC-SSP and MST-PSB, might indeed further decrease the individual’s sense of being in control (Bonta et al., 2000; Cecile & Born, 2009; De Valk et al., 2016; Hendriks & Bijleveld, 2008; Parhar et al., 2008).

Most of the sub-clinically scoring juveniles (64%, N = 9) reported a significant improvement in impulsivity, while 36% (N = 5) reported no improvement, but showed no deterioration either. Learning to think before acting is considered to be an integral part of what juveniles should learn in treatment. Developing behavioral alternatives to offending includes learning to react less impulsively. Several juveniles were diagnosed with ADHD, implying that a reduction in impulsivity may be difficult to establish. These specific juveniles might benefit from regular administration of psychotropics, which reduces impulsivity by improving the ability to control impulses. Lastly, some (24%, N = 4) juveniles in the normal range of functioning reported more impulsivity after treatment. This may prove a side effect of gaining more insight into abilities and functioning through treatment.

Few sub-clinically scoring juveniles reported improvement in empathy, while most juveniles (about 80%, N = 8 up to 10) reported no improvement. When a lack of empathy is the result of a pervasive disorder, this condition may prove less resolvable, and, therefore, little improvement may be measured by the ASAP-D, which operationalizes empathy as the (innate) ability to see things from another person’s perspective, sympathize with fictional characters, and understand what other people are feeling by looking at them (Van Outsem et al., 2008). Treatment, however, may have offered juveniles (cognitive) insight into harming others and, therefore, will be able to more explicitly ask for consent. These (social) skills may not be reflected in post-treatment measurements of empathy. Of the juveniles in the normal range of functioning, some (15%, N = 3) reported a deterioration in empathy. Gaining more insight into abilities and functioning via therapy might explain this decrease.

On aggression (total construct), few sub-clinically scoring juveniles reported improvement, most juveniles (about 80%, N = 4 up to 8) reported no change. Of the juveniles in the normal range of functioning, almost all juveniles remained in the normal range.

Most sub-clinically scoring juveniles reported no improvement in hypersexuality (N = 12). Therefore, most juveniles seem to have remained very focused on / interested in sex. In both treatment types only one juvenile reported a significant decrease in hypersexuality. Hypersexuality is operationalized in the ASAP-D as an (obsessive) preoccupation with sex, which the juvenile himself is bothered by (Van Outsem et al., 2008). Thus, after treatment, only two juveniles reported having reduced this pre-occupation. This finding might implicate
that most juveniles still experience problems with their sexual interest levels. Becoming more confident with their own sexuality and levels of (to some extent pubescent) sex drive may still prove a challenge for these juveniles.

Fifty percent (N = 3) of the sub-clinically scoring juveniles reported a decrease in cognitive distortions, while the other fifty percent (N = 3) reported no change. The cognitions measured depict harmful sexual behavior towards peers and children as acceptable, justifiable or harmless, and therefore serve as a sexual problem sustaining factor (Van Outsem et al., 2008). Restructuring or ‘challenging’ these thoughts are a part of Out of the Circle therapy and group therapy as offered by SYC-SSP (Koster & Tel, 2010). MST-PSB also targets these convictions. Apparently not all distortions were sufficiently challenged. Giving more attention to this subject, especially in juveniles who report subclinical levels of cognitive distortions, seems advisable. However, juveniles who make use of justifications may use them to preserve a more positive self-image (Abel, Becker, & Cunningham-Rathner, 1984). Fully admitting the harmfulness of one’s previous behavior might prove very stressful. For this reason, treatment of juveniles with a fragile personality structure, justifications might be challenged less, and treatment might have been particularly focused on self-image, which is measured by the ASAP-D via self-esteem. As mentioned, self-esteem may be positively influenced by use of the Good Lives model for juveniles showing harmful sexual behavior, rather than by mainly focusing on recidivism risk and criminogenic treatment needs (Collie et al., 2007).

LIMITATIONS

Most prominently, this study was conducted on a small sample of non-randomly assigned juveniles. The latter restricts comparability, but increases the relevance of the results for clinical practice. A matching of level of treatment needs was used to create relative comparability. On some scales, however, even when the researched treatment modalities were specifically aimed at juveniles at risk for recidivism, few juveniles reported subclinical levels of treatment needs. This limits statistical power and hampers the generalization of study findings. Testing significant changes by means of RCI at the individual level, however, is considered a remedy to deal with lack of statistical power to detect significant change at the group level. The results of this study are to be considered indicative and exploratory. At the same time, our results are largely in line with previous research on the effectiveness of residential versus non-residential treatment of behaviorally disordered juveniles (De Swart et al., 2012; Strijbosch et al., 2015), increasing confidence in the current study findings.
Secondly, this study was limited because risk domains that MST-PSB specifically focusses on (e.g., family functioning and school functioning) were not assessed. The primary focus of this study was residential treatment, and, therefore, the choice was made to measure (improvement on) overlapping treatment goals of both the residential and non-residential comparison program MST-PSB. We acknowledge that the current study, as a result, provides a restricted view on the (intended) results of MST-PSB.

Thirdly, this study was carried out making use of self-report data, which may not provide a fully reliable representation of treatment needs. As reported, 25 percent of the juveniles scored high on social desirability, so some juveniles may have underreported their level of treatment needs. In the current study, the assessment of treatment needs was therefore complemented by an assessment of typology and diagnoses at admittance (a professional judgement on severity / treatment need), providing a broader description of the complexity of these cases.

CONCLUSION

Evidence-based residential and non-residential (contextual) treatment obtained equal positive changes in psychosocial treatment needs of juveniles with harmful sexual behavior. However, consistent with our expectations, the juveniles treated in the two contexts differed. As hypothesized, juveniles in residential care represent a more at recidivism risk group with higher (self-reported) treatment needs. Both treatment modalities might improve their results on several important (criminogenic) treatment needs. According to the authors, adaptation to a more holistic and less risk focused approach may prove fruitful.
General Discussion
INTRODUCTION

In general, media reports of sexual offending are met with social outcry and unrest. Especially when victims are (small) children, the image of a (sexually) distorted or devious perpetrator, even when he or she is a juvenile, is frequently used. In general, the conviction “once a sex offender, always a sex offender” still seems to exist among the general public, portraying juveniles as life course persistent sexual abusers. Even among professionals, the conviction that a biological predisposition (partially) caused sexual misconduct was common not that long ago (Vijselaar, Gijzeman, Bouwens, & Bertens, 2015).

In an attempt to do justice to the diversity of juveniles who have shown harmful sexual behavior, this dissertation is aimed at improving matched care for these juveniles; what ‘type’ of juvenile with harmful sexual behavior should be treated by (what ‘type’ of) intensive, mandated treatment, aimed at which individual treatment needs? Are there differences in approach or effects that could guide allocation of these juveniles to a better matched care? How can treatments improve their results by incorporating contemporary research findings into their practices? Making use of two Dutch samples of juveniles in intensive, mandated, sex offense specific treatment over a five year period, and synthesizing previous research by means of two meta-analyses, this dissertation integrates knowledge on the effects of treatment for juveniles with harmful sexual behavior, and explored the allocation to, aims, and effects of treatment for juveniles with harmful sexual behavior in the Netherlands.

In the Netherlands, a broad spectrum of treatment specifically aimed at juvenile sexual problem behavior is available. Juveniles with sexual problem behavior are generally treated separately from juveniles who have displayed predominantly aggressive or non-sexually transgressive problem behavior. However, indication criteria, adequately distinguishing between the available most intensive, mandated treatment types, are lacking. Furthermore, application of the dominant Risk-Need-Responsivity (RNR) model for offender rehabilitation (Andrews & Bonta, 2010) to juveniles with harmful sexual behavior constitutes a challenge. Juveniles with harmful sexual behavior form a notoriously heterogeneous group regarding their psychosocial treatment needs and (re)offending patterns. To date, no psychological typology has been found to represent a more at sexual recidivism risk subgroup. This makes it hard to identify juveniles in need of intensive mandated care at treatment admission. Adding to the difficulty of allocation are several persistent (erroneous) assumptions about juveniles with harmful sexual behavior (Worling, 2013). These assumptions play down the important individual differences in these juveniles regarding their level of treatment needs (i.e., their level of deviance, deficits or disorders), and the level of sexual (recidivism) risk these juveniles constitute, increasing the chance of providing a mismatched treatment intensity or type of care.
In this chapter, a summary of the research findings and their practical implications will be presented, followed by a discussion of the broader implications of the combined study results. The chapter concludes with a reflection on strengths and limitations of the studies conducted, the implications for future research and an overall conclusion.

**SUMMARY OF RESULTS AND PRACTICAL IMPLICATIONS**

The first study addressed the question whether juveniles with harmful sexual behavior in intensive, mandated treatment in the Netherlands (i.e., Multisystemic Therapy - Problem Sexual Behavior; MST-PSB, Secure Youth Care for juveniles with Severe Sexual Problems; SYC-SSP, or Forensic Youth Care for juveniles with Severe Sexual Problems; FYC-SSP), were allocated according to RNR principles. Also, it described and compared background characteristics of these juveniles. It was found that juveniles with relatively less treatment needs (i.e., a lower sexual drive and impulsive/antisocial behavior and a higher problem insight, motivation, and community stability), and who were, consequently, at lower general recidivism risk, more often received community based MST-PSB. Estimated recidivism risk levels of juveniles in all three treatment modalities did not always support the need for risk reduction by the imposition of limitation of freedom of movement and maximum supervision provided. Based on the assessed sexual recidivism risk, 38% of the juveniles in FYC-SSP, 7% in SYC-SSP, and 24% in MST-PSB had received treatment that was too intensive, considered detrimental to the juvenile's treatment motivation and development. Taking the general overestimation of risk by actuarial risk assessments for juveniles into account, this finding is to be taken very seriously. Both referrers and treatment providers should more stringently adhere to admission criteria and only refer juveniles with more (specific) treatment needs, and consequently, a higher (sexual) recidivism risk. It was deemed important that a structured assessment of treatment needs and recidivism risk is administered and given more weight in the allocation to mandatory forms of treatment. Because of the restricted validity of actuarial assessments in predicting (sexual) recidivism by juveniles, the assessed treatment needs (i.e., etiology) were expected to provide the most valid base for treatment allocation purposes. Practical implications of the assessed background characteristics (e.g., relatively high age in combination with more prepubescent victims, more paraphilia, pervasive and personality disorders) and treatment needs (relatively high scores on sexual drive, intervention items, and community instability) of juveniles in FYC-SSP, appeared to be that stimulating moral development and learning to cope with an atypical sexual interest should be specific treatment targets. In contrast, practical implications of the background characteristics (e.g., more neglect, prior out of home placements, and disruptive disorders) and treatment needs of juveniles in SYC-SSP (relatively high scores on impulsive/antisocial behavior and community instability), implicated that SYC should focus less on specific
atypical sexual behavior and attitudes, and more on a broader set of (generalist) criminogenic needs. Improving personal adjustment (taking classes/qualifying for employment, anger management, and social skills training) and building (new) social networks could be key elements of treatment for this group of juveniles.

The second (quantitative review) study investigated the effect of treatment on recidivism by juveniles who have sexually offended. It also examined the potential moderating effect of type of recidivism, and several participant, treatment, and study characteristics by means of a multilevel meta-analysis. A small to moderate effect size was found ($d = 0.37, p < .001$). However, after controlling for publication bias, a significant treatment effect was no longer present ($d = 0.15, p = .176$). Type of recidivism did not moderate the effect of treatment, nor did participant or treatment characteristics. Regarding study characteristics, a shorter follow up time showed a trend for larger effect sizes. Secondly, effect size calculation based on proportions yielded larger effect sizes than calculation via mean frequency of offending.

This indicates that the proportion of recidivism was significantly lower in treatment groups compared to comparison groups, but the mean number of offenses for every reoffender in the two groups did not significantly differ. Regarding the effect of treatment on recidivism, it was noted that the treatments researched compared themselves to other (possibly as effective) treatments, reducing differences in effect size as measured by Cohens $d$. Furthermore, the treatment forms were aimed at a highly heterogeneous group of juveniles, possibly influencing effectiveness. Thirdly, sexual recidivism rates are generally low. Relatively low sexual recidivism rates, however, do not imply that juveniles who have sexually offended are not in need of treatment to reduce recidivism risk, if only because their general recidivism rates are higher. Mixed offenders, who display more antisocial tendencies, have been shown to be at higher risk for recidivism and, therefore, in greater need of (intensive) treatment. The study did show sex offense specific treatment to have an equal effect on other types of recidivism. Thus, juveniles who have sexually offended and who display specific as well as general criminogenic needs may be best off in mandated, intensive treatment. Not all juveniles who have sexually offended display these features equally, therefore, not all juveniles who have sexually offended may be in need of sex offense specific treatment as to reduce recidivism risk.

The third (quantitative review) study examined the effects of treatment for juveniles with harmful sexual behavior on psychosocial functioning, as well as the potential moderating effects of outcome, treatment, participant, and study characteristics by means of a multilevel meta-analysis. A moderate overall effect size was found of $d = 0.60 (p < .001)$, indicating groups receiving treatment achieved an estimated 33% relative improvement on psychosocial functioning. Type of outcome measure did moderate the effect of treatment, indicating that effects on atypical sexual arousal (as measured by penile plethysmography
in 48% percent of the cases) were smaller compared to effects on other outcomes (i.e., overall functioning, delinquency and aggression, impulse control, social and coping skills, emotion and self-image, cognitions and sexual knowledge, and family functioning). The treatment effect on empathy also tended to be smaller. Most prominently, studies of weak quality produced larger effect sizes. Unexpectedly, non-established treatments had more effect than did established treatments, which may be explained by the use of less rigorous study designs. Treatment groups with a higher percentage of juveniles with similar age victims or mixed type problem behavior also yielded larger effect sizes. Lastly, evaluation of treatment effects by professionals produced higher effect sizes, compared to other sources of information (e.g., adolescent self-report). It was concluded that treatment aimed at psychosocial functioning seems promising. Even if some psychosocial factors have not (yet) been established as criminogenic, they represent real life problems for juveniles. Recent developments in treatment methods describe a return to more holistic treatment frameworks, which, for juveniles with harmful sexual behavior, might prove especially relevant since (sexual) recidivism rates are low, but individual treatment needs may be high.

The fourth study investigated the effects of intensive, mandated residential and non-residential treatment on psychosocial functioning among adolescents with harmful sexual behavior. It compared their psychosocial treatment needs at admittance and after having received cognitive-behavioral based residential treatment or multi-systemic contextual treatment. Results replicated the findings of the first study in that juveniles in residential treatment represent a group at higher risk for recidivism (more repeat offending, more sex-plus offending), and with higher (self-reported) treatment needs (i.e., higher orientation towards sex, more emotional loneliness, more impulsivity) than juveniles in non-residential treatment. In general, juveniles in intensive, mandated treatment reported low self-esteem relatively often (72%), specifically in SYC-SSP, emotional loneliness (65%) and impulsivity (59%) were also often reported. Residential cognitive behavioral based treatment and contextual multi-systemic treatment obtained equal positive changes on psychosocial treatment needs. However, many juveniles reporting a (sub)clinical treatment need at admittance also reported no significant change after treatment. Some juveniles even showed significant deterioration on aspects of their psychosocial functioning after treatment. The study results specifically indicated room for improvement of treatment effect with regard to self-esteem, locus of control, empathy, aggression, hypersexuality, and orientation onto sex, yielding 80 up until 92% of no change or even deterioration after treatment. The mandatory administration and risk oriented focus of the treatments studied were hypothesized to not provide the most fertile ground on which to improve results. Focusing less prominently on the sexual problem behavior and risks (possibly worsening self-esteem) and more
prominently on life-fulfillment (e.g., competence, relatedness, and autonomy) may proof useful to explicitly counteract universal detrimental aspects of having displayed harmful sexual behavior that juvenile's face in general (i.e., stigma and social deprivation).

DISCUSSION

The (level of) individual treatment needs of juveniles with harmful sexual behavior differ significantly. They do not form a homogeneous group regarding their offending patterns and the etiology of their problem behavior (Barbaree & Marshall, 2006; Fanniff & Kimonis, 2014; Hendriks, 2006; Ryan, Leversee & Lane, 2010; Seto & Lalumiere, 2010). A distinction is often made through use of the ‘specialist’ and ‘generalist’ view on juveniles with harmful sexual behavior (Van den Berg, 2015). The first -specialist- view focuses on specific explanations for the development and continuation of sexually transgressive behavior in juveniles. The second -generalist- view conceptualizes sexual offending as a subtype of general (juvenile) conduct problems, emphasizing similarities between juveniles who have sexually offended and those who display other forms of offensive behavior. Both views receive support in recent research (Fanniff & Kimonis, 2014; Seto & Lalumiere, 2010), as sexually transgressive juveniles (especially those with child victims) indeed display specific intra- and interpersonal problems, but, at the same time, display an equal age of onset of problem behavior, and equal levels of antisocial thoughts, behavior patterns, and callous-unemotional traits (i.e., especially those juveniles with peer victims or mixed offending behavior, see also Drew, 2013; Hendriks, 2006; Leroux, Pullman, Montayne, & Seto, 2016).

To date, in accordance with the dominant RNR model for offender rehabilitation (Bonta & Andrews, 2007), empirically established recidivism risk factors and factors derived from theories on etiology (regarding sexual problem behavior specifically as well as regarding conduct problems in general), determine at what factors (deficits) treatment for these juveniles is aimed. Research on sexual recidivism by juveniles is hampered by low prevalence rates of official sexual reoffending (Caldwell, 2016), only small sample sizes are available for statistical analyses. Still, several sexual recidivism risk factors have been established (Carpentier & Proulx, 2011; Christiansen & Vincent, 2013; Worling & Långström, 2003). Several of these factors are, however, considered static; they represent past occurrences that cannot be altered (e.g., number of previous victims, type of offending behavior, age of onset). Those that are considered dynamic (to be influenced by treatment) are also designated as criminogenic treatment needs (e.g., antisocial attitudes, sexual deviation, social isolation, family functioning, school functioning). Intensive treatment programs tailored to juveniles with harmful sexual behavior indeed generally are aimed at remediating cognitive distortions (sexual as well as antisocial), improving social and relational skills,
incorporate relapse prevention techniques (Bromberg & O’Donohue, 2014; Ryan, Leversee, & Lane, 2010; Veneziano & Veneziano, 2002). Importantly, establishing a correlation between a (historical) characteristic and recidivism does not yet establish its predictive validity, that is, whether change in a criminogenic treatment need is indeed related to a change in criminal offense recidivism (Baglivio, Wolff, Jackowski, & Greenwald, 2015).

The main treatment method in use is multimodal cognitive behavioral based treatment (CBT), as this treatment type has rendered the best results in reducing recidivism in general (Dopp, Borduin, & Brown, 2015; Hanson, Bourgon, Helmus, & Hodgson, 2009; Lösel & Schmucker, 2005; Schmucker & Lösel, 2015; Walker, McGovern, Poey, & Otis, 2004). The prominence of reducing sexual recidivism as a treatment aim, notably the first ‘R’ in the RNR model, does, however, remain problematic. The scarcity of sexual recidivism (with a contemporary weighted mean of 2.75%; Caldwell, 2016), makes risk assessment, and allocation based on recidivism risk levels, difficult. Our meta-analyses indicated treatments to be (more) effective in reducing recidivism, as (mostly) compared to treatment as usual, which represent treatments that are regularly administered but have not (yet) been found to be effective. This result, however, proved affected by publication bias, indicating smaller, possibly non-significant, differences in daily practice (Ter Beek et al., 2017b). Usually, established treatments are found to be more effective than ‘treatment as usual’. Differences found in effects between treatment as usual and established treatments, however, are not always as expected for all (types of) clients, underscoring the importance of matched care and responsivity (Weisz et al., 2013).

As a result of the dominant risk focus, juveniles with harmful sexual behavior receive mandated care, specifically aimed at antisocial cognitions and possible atypical sexual preferences. They undergo restrictions in freedom during adolescence, a key developmental phase, when risk assessment does not always support the need for these restrictions (Ter Beek et al., 2017a). A focus on (sexual) deviance and deficits might even become an iatrogenic effect of treatment; a possible negative influence on self-image (Cecile & Born, 2009; Deci & Ryan, 2000; Gatti, Tremblay, & Vitaro, 2009). The stigmatization juveniles face by being labeled a sex offender is high. In some states in North America lifelong public registration is a consequence of an adjudication. In the Netherlands registration, although not public, also exists, hampering the acquirement of work, which is a potent protective factor (Spruit, Van der Put, Gubbels, & Bindels, 2017; Van den Berg, 2015).

Juveniles with harmful sexual behavior do, generally, self-report several (albeit non-criminogenic) treatment needs. For example, low self-esteem, emotional loneliness, and external locus of control (helplessness or defeatism) seem relatively often reported problems by these juveniles (Ter Beek et al., 2017d). Self-report, however, especially by juveniles with
harmful sexual behavior, generally is met with a certain amount of skepticism (Worling, 2013). The supposedly deceitful nature of delinquent juveniles, combined with an expected relatively low insight into one’s own level of functioning (e.g., because of developmental problems or lower intellectual functioning) is reason for many practitioners to rely more on professional assessment of (criminogenic) treatment needs (Breuk, Clauser, Stams, Slot, & Doreleijers, 2007). Juveniles with harmful sexual behavior, however, generally tend to be less antisocial, and, therefore, arguably less deceitful than their non-sexually transgressive peers (Fanniff & Kimonis, 2014; Seto & Lalumière, 2010). Furthermore, risk assessment tools for juveniles with harmful sexual behavior were found not to be very accurate. They often overrate risks and, as a consequence, level of treatment need (Hempel, 2013). Additionally, our meta-analytic study on the effects of treatment on psychosocial functioning indicated professional judgment to produce divergent results on the effects of treatment as compared to self-report and parent-report, indicating professional judgment not to be free from bias itself (Ter Beek et al., 2017c; Van Vugt et al., 2012).

Several research practitioners have recently advocated a more positive and holistic (less risk focused) approach to juvenile sexual problem behavior (Dopp, Borduin, & Brown, 2015; Ward, Yates, & Willis, 2012; Worling, 2013). Although the importance of reserving intensive treatments for juveniles most at recidivism risk was substantiated, Lipsey (2009), through a meta-analytic review, described three factors to be of pivotal importance for the efficacy of juvenile offender treatment: 1) use of a ‘therapeutic intervention philosophy’, 2) serving high risk offenders, and 3) quality of implementation (i.e., treatment fidelity). The first factor mentioned, a ‘therapeutic intervention philosophy’, was defined as “an approach that attempts to engage the juvenile in a supportive, constructive process of change” (pp. 128), underscoring the importance of intrinsic motivation, and of the last ‘R’ in the RNR model (i.e., responsivity; tailoring treatment to individual characteristics such as cultural background, gender, age, cognitive abilities, and individual treatment motivation).

The Self Determination Theory (SDT) on intrinsic motivation by Ryan and Deci (2017) describes the mechanisms of human development, postulating that “humans are active, growth-oriented organisms who are naturally inclined towards integration of their psychic elements into a unified sense of self and integration of themselves into larger social structures” (Deci & Ryan, 2000, pp. 229). Through extensive research, they established three basic needs (or ‘drives’, or ‘motivations’) in every human being; the need for competence (feeling productive and useful), relatedness (feeling a part of or connected to others), and autonomy (feeling in charge of one’s own life or being able to choose freely between alternatives). They describe how ‘thwarting’ of these needs (while growing up), creates alternative, defensive or self-protective processes, which, importantly, have functional utility. These processes include egocentrism and antisocial activities, such as harmful sexual behavior.
The Good Lives Model applies this paradigm to the treatment of juveniles with harmful sexual behavior. It proposes to primarily focus treatment on ‘approach goals’ (i.e., well-being, or the adequate prosocial fulfillment of basic needs), instead of on ‘avoidance goals’ (i.e., relapse prevention, remediating deficits; Ward, Yates, & Willis, 2012). It emphasizes the assessment of specific aspects of importance to the well-being of an individual juvenile. To achieve this, the juvenile self-reports (narratively) on what to him constitutes a ‘good live’; what activities or situations provide joy and fulfillment? Thereby prominently assessing individual protective factors. The professional then assesses what basic needs are most important in the juvenile’s life. The harmful sexual behavior is conceptualized as an inappropriate strategy to obtain these basic needs, or, well-being, the latter being inherently ‘normal’ or non-deviant. Secondly, in accordance with the RNR model, recidivism risk factors are assessed (Fortune, Ward, & Print, in preparation). This process constitutes a function analysis of (harmful) behavior and individual motivation, and importantly adds to a more holistic comprehension of the juvenile; understanding his choice of goal directed behavior in a more comprehensive manner, rather than focusing on risk factors alone, the latter running the risk of stimulating repression rather than growth (De Valk et al., 2016). In their response to recent criticism on the RNR model facilitating an excessive focus on risks, Andrews, Bonta and Wormith (2011) have endorsed the importance of targeting the well-being of higher risk cases as an element of ethical, professional, humane, and decent practice. They postulate that within the RNR model this practice is well described, positioned and endorsed through its last R; responsivity.

Aligning treatment goals with personal goals of the juvenile, does directly support ‘agency’ (e.g., self-direction or autonomy) and treatment motivation (Deci & Ryan, 2000; Koestner, Lekes, Powers, & Chicoine, 2002). Adolescents generally are particularly in need of agency or self-direction. Instead of undergoing mandatory treatment (De Valk et al., 2017), aligning treatment goals with personal goals offers the opportunity to actively build towards (a shared) perspective and simultaneously reduce (risk of future) harmful behaviors. “..., needs (i.e., competence, autonomy and relatedness) are the linking pin between the affordances and demands of the social world on one hand and either people’s natural tendencies toward growth and well-being or their accommodative tendencies toward self-protection with the accompanying psychological costs on the other hand” (Deci & Ryan, 2000, pp. 262). Thus, when juveniles with harmful sexual behavior self-report issues, such as low self-esteem, emotional loneliness, and an external locus of control, they seem in need of treatment that assists them in regaining or building self-confidence (competence), self-direction (autonomy), and relatedness (Deci & Ryan, 2000; Ryan & Deci, 2017). The effects of sex offender specific treatment on overall psychosocial functioning seem promising (Ter Beek et al., 2017c), indicating treatment to be able to assist juveniles in achieving personal life goals in a more prosocial manner. The most promising treatment effects are achieved on overall
psychosocial functioning, rule breaking and aggression, impulse control, social and coping skills, emotion and self-image, cognitions and sexual knowledge, and family functioning, aspects generally considered to be important protective factors for antisocial behavior (Lösel & Farrington, 2012).

Finally, some juveniles with harmful sexual behavior will not self-report or even portray psychosocial problems. This fact is generally met with certain skepticism, because harmful sexual behavior by juveniles is considered deviant and, as a consequence, juveniles displaying this behavior must also be somewhat deviant (Worling, 2013). However, alternative (non-intrapsychic) explanations for harmful sexual behavior are available. Peer group affiliation, including being easily influenced (common in juveniles), the influence of alcohol or drugs and situational characteristics, such as family dysfunction or a lack of (parental) monitoring in a difficult developmental phase, might contribute to harmful sexual behavior by juveniles (Dopp, Borduin, & Brown, 2015; Ryan, Levee, & Lane, 2010; Van Outsem, 2009). Guidance or contextual treatment might still be important for these juveniles, but intensive individual treatment does not seem an adequate fit. Accepting the fact that not all juveniles with harmful sexual behavior are ‘inherently deviant’, is deemed imperative to a more matched care delivery and the achievement of better treatment results.

STRENGTHS AND LIMITATIONS

Some aspects of the studies in this dissertation merit reflection. Firstly, in the studies conducted, restricted sample sizes were used. Most studies on juveniles with harmful sexual behavior encounter a small number of juveniles adjudicated or treated specifically for this type of problem behavior (Fannif & Kimonis, 2014). Levels of unreported sexual problem behavior (dark numbers) are estimated to be relatively high, so these samples might constitute the ‘tip of the iceberg’ (White, 2011; Wittebrood, 2006). Official recidivism reports involve getting caught, and getting caught again (Yun & Lee, 2013). This means that the juveniles currently studied represent a subgroup of juveniles with harmful sexual behavior, namely those that get caught. This fact restricts the generalizability of the results to this specific group of sexually transgressive juveniles. And although sufficient studies were found to perform two multilevel meta-analyses, the total amount of participants remained relatively low for this type of statistical analyses. For both meta-analyses a three-level mixed effects model (Assink & Wibbelink, 2016) was used to maximize statistical power and preserve as much information as possible, accounting for both within and between sources of variance. The two Dutch samples sizes are also considered small, they, however, did contain all juveniles in intensive specialized treatment in the Netherlands over a two year period (2010-2012), and all juveniles in specialized mandated Dutch secure youth
care over a three year total period (2012-2015). Testing of significant changes by means of Reliable Change Index at an individual level, was used as a remedy to deal with lack of statistical power to detect significant change at group level.

Secondly, the inclusion of older studies in the meta-analyses, a prerequisite for systematic reviews, and the inclusion of mostly North American studies, may limit generalizability (Bijleveld, 2015), although studies examining risk factors for juvenile (general) criminal offense recidivism in the United States and Europe showed remarkable similar results (Van der Put et al., 2011, 2012). Interestingly, Leijten, Melendez-Torres, Knerr, and Gardner (2016) showed that parenting interventions for reducing child disruptive behavior that were developed in the United States could be transported to other continents without loss of effectiveness, which proved not true for the effects of multi-systemic treatment on juvenile criminal offense recidivism (Van der Stouwe, Asscher, Stams, Dekovic, & Van der Laan, 2014). Studies are conducted within a certain time frame and context, which especially influences studies on sexual problem behavior; what is considered atypical in some parts of the world may not be considered so in other parts of the world. Also, time alters perceptions on normalcy of sexual behavior (in adolescence). Results, therefore, should be cautiously applied to other (especially non-Western) parts of the world.

Finally, only few participant characteristics could be included in the moderator analyses in the meta-analytic studies, because not many studies comprehensively reported on sample characteristics. The heterogeneity of juveniles with harmful sexual behavior demands the thorough reporting of sample characteristics to enable assessment of external validity of study results, and to conduct moderator analyses to test intervention effects in subgroups of juvenile sex offenders (Bijleveld, 2015). Because this was not the case in many studies, possible moderating effects of participant characteristics might have been missed.

**DIRECTIONS FOR FUTURE RESEARCH**

Further research into the field of (relapse into) sexual transgressions by juveniles is needed in order to obtain a fuller understanding of its etiology and of which juveniles persist and desist. Aiming mandated intensive treatment at juveniles not in need of intensive treatment is a stigmatizing waste of resources, sooner causing harm than benefit. Research on protective factors, and the attainment of well-being specifically, is still in its infancy, while research on risk factors is ongoing, but still incomplete. Knowledge on how well-being of these juveniles may be enhanced and supported short and long term, and whether well-being indeed prevents (sexual) reoffending, seems imperative to successfully incorporate the concept of well-being into mandated treatment.
Further research on the last R of the RNR model (i.e., how to improve responsivity and its relation to established treatment integrity) is also advised. Treatment integrity is an important predictor of treatment success, as is responsivity. The question how to ‘stick with a program’, but at the same time create sufficient flexibility in treating juveniles with a complex and highly individualized treatment need is a topic worthy of further investigation.

Lastly, the adequate assessment of dynamic treatment needs that are of specific importance to juveniles with harmful sexual behavior, also remains an important topic of interest. Validity of psychodiagnostic measurements and treatment need assessments for this highly heterogeneous and relatively small group of juveniles should remain a focus point of research; what aspects can be influenced through treatment and are therefore to be made measurable? The operationalization of the concept of well-being and protective factors seem of prominent interest, as to be able to assess treatment progress on these topics in the future.

**CONCLUSION**

The central question of this dissertation was “how to provide matched care for juveniles with harmful sexual behavior”. In sum, juveniles who self-report high levels of treatment needs (specific or generic) should be allocated to intensive forms of treatment tailored to these specific needs. Intensive, mandated treatment should generally focus more explicitly on the improvement of well-being (i.e., the fulfillment of the three basic human needs; competence, relatedness, and autonomy). Through this, the juvenile’s goal directed behavior (i.e., the sexual harmful behavior) is able to change into more prosocial behavior. This foremost constitutes adapting a more holistic view on harmful sexual behavior by juveniles by current mandated treatments.

When juveniles do not self-report a high level of treatment need, an assessment of callous-unemotional traits and their transgressive behavior in general may clarify which juveniles are in need of mandated intensive treatment focused on conduct problems in general, and which juveniles indeed are not in need of mandated, intensive individual treatment. The last group may require less intensive and less restrictive guidance, (family) counseling or psycho-education, depending on the (contextual) factors that facilitated the harmful behavior.

In conclusion, self-reported treatment need levels, combined with recidivism risk assessment, may best guide allocation to and contents of treatment. Improving well-being of the individual juvenile, thereby also preventing relapse into harmful behavior, is deemed the most important feature of matched and effective treatment in general.
References
References marked with an asterisk (*) are studies included in meta-analyses

A


Chapter 7


B

References


D


Chapter 7

E


F


G


H


Hart-Kerkhoffs, L. (2010). *Juvenile sex offenders, mental health and reoffending*. Amsterdam, the Netherlands: VU University Medical Center Amsterdam.


K


L


Chapter 7


150
References


Chapter 7


N


Chapter 7

Q


R


S


Chapter 7


V


Chapter 7


Worling, J. R. (2013). What were we thinking? Five erroneous assumptions that have fueled specialized interventions for adolescents who have sexually offended. *International Journal of Behavioral Consultation and Therapy, 8*, 80-88.


References

Y


Z
Juveniles with harmful sexual behavior constitute a heterogeneous group regarding treatment needs and reoffending patterns. In general, intensive mandated treatment is aimed at reducing recidivism risk and, therefore, aimed at criminogenic treatment needs. Sexual recidivism in juveniles, however, is scarce and mostly adolescence limited. No psychological typology to date has been proven to represent a higher at sexual recidivism risk subgroup. Risk assessment instruments generally overestimate recidivism risk, so allocation to mandated intensive treatment aimed at reducing sexual recidivism risk is a difficult task. Up to 38% percent of juveniles with harmful sexual behavior in the Netherlands prove to be allocated to too intensive and restrictive treatments, considered detrimental to development and treatment motivation. Treatment effect on reduction of recidivism was established as small to moderate ($d = 0.37$), but influenced by publication bias, indicating small, possibly non-significant, results in treatment practice. The effect of treatment on psychosocial functioning was established as moderate ($d = 0.60$), with several moderating influences of outcome, participant, treatment, and study characteristics. Importantly, a primary focus on (sexual) deviance or risk reduction in juveniles with harmful sexual behavior may interfere with treatment effects on relatively often self-reported psychosocial treatment needs (i.e., low self-esteem, emotional loneliness, and external locus of control; helplessness or defeatism). A shift to a more holistic paradigm for treating juveniles with harmful sexual behavior is advocated. Aiming treatment at approach goals instead of avoidance goals, thereby aligning treatment goals prominently with the attainment of well-being, is expected to improve treatment results and reduce relapse into transgressions.
Appendices

Nederlandse Samenvatting
Dankwoord
Curriculum Vitae
Publicaties
NEDERLANDSE SAMENVATTING (DUTCH SUMMARY)

Intensieve, gedwongen vormen van behandeling zijn in het algemeen gericht op het verminderen van het risico op herhaling van grensoverschrijdend gedrag, en dus gericht op criminogene behandelbehoeften (kenmerken aangetoond een relatie hebbende met dat grensoverschrijdende gedrag). Herhaald seksueel grensoverschrijdend gedrag door jeugdigen is echter schaars en beperkt zich meestal tot de adolescentie. Jeugdigen die ernstig seksueel grensoverschrijdend gedrag vertonen vormen daarnaast geen homogene groep wat betreft type en mate van behandelbehoeft en (de kans op) herhaling van grensoverschrijdend gedrag; zij verschillen fors van elkaar. Tot op heden identificeert geen enkele psychologische typering een subgroep binnen deze groep jeugdigen met een hoger risico op herhaling van seksueel probleemgedrag. Wetenschappelijk onderzoek naar risicofactoren voor herhaald seksueel grensoverschrijdend gedrag (criminogene behandelbehoeften) wordt daarnaast belemmerd door kleine steekproefgroottes van jeugdigen die daadwerkelijk in herhaling vallen (en hiervoor opgepakt / veroordeeld worden). Grotere steekproefgroottes zijn nodig om verschillen tussen jeugdigen die wel terugvallen in dit gedrag, en degenen die dit niet doen betrouwbaar vast te stellen. De in gebruik zijnde specifieke risicobeoordelingsinstrumenten (zoals de J-SOAP-D in Nederland) overschatten het risico op herhaald seksueel grensoverschrijdend gedrag over het algemeen. De toewijzing aan gedwongen intensieve specialistische behandervormen, expliciet gericht op jeugdigen met een hoog recidive risico, is daarom gecompliceerd. Tot 38% van de jeugdigen met ernstig seksueel grensoverschrijdend gedrag in Nederland blijkt toegewezen aan te intensieve en te restrictieve behandervormen. Dit wordt beschouwd als schadelijk voor de algemene ontwikkeling en de behandelmotivatie van de jeugdige en is daarmee zeer onwenselijk.

Op seksueel grensoverschrijdend gedrag gerichte behandervormen blijken in het algemeen een klein tot middelgroot ($d = 0.37$) incrementeel effect op het verkleinen van herhaald grensoverschrijdend gedrag (recidive) te behalen. Dit resultaat blijkt echter beïnvloed door een overmatige publicatie van positieve studieresultaten (‘publication bias’), wat wijst op het in de praktijk behalen van kleine, mogelijk niet significant, verschillen in effect op het verminderen van herhaald (seksueel) grensoverschrijdend gedrag. Geen type behandeling lijkt specifiek meer effect te sorteren dan een andere.

Het effect van specialistische behandervormen op psychosociaal functioneren werd vastgesteld als middelgroot ($d = 0.60$), met modererende invloeden van type uitkomstmaat, type jeugdige, type behandeling en studiekenmerken. Op atypische seksuele opwinding en empathie werd minder behandeleffect vastgesteld dan op andere vormen van psychosociaal functioneren (zoals algemeen functioneren, regel brekend gedrag & agressie,
impulsiviteit, sociale vaardigheden & coping, emoties & zelfbeeld, seksuele cognities en kennis en gezinsfunctioneren). Bij jeugdigen met seksueel grensoverschrijdend gedrag naar leeftijdgenoten ('peers') werden betere behandelresultaten aangetoond dan bij jeugdigen met seksueel grensoverschrijdend gedrag naar kinderen (het slachtoffer is dan jonger dan twaalf én minimaal 5 jaar jonger dan de dader). Ook rapporteerden oudere studies (van een lagere studiekwaliteit) betere behandelresultaten. Lage studiekwaliteit bleek in bijzondere mate bij te dragen aan verschil in effect. Dit laatste pleit voor het uitvoeren van meer robuust onderzoek naar het verbeteren van psychosociaal functioneren, op grotere schaal en met een uniforme operationalisering van het begrip welbevinden.

Een studie naar de effecten van specialistische, intensieve behandelvormen in Nederland (MST-PSB en ESP in JeugdzorgPlus) op psychosociaal functioneren, zoals gemeten met de ASAP-D, toonde aan dat deze behandelvormen vergelijkbare resultaten behalen. Op een aantal behandelbehoeften zoals door de jeugdigen gerapporteerd, verdient het resultaat echter verbetering. Een primaire focus op behandelen van (seksuele) afwijking en risicoreductie bij jeugdigen met ernstig seksueel grensoverschrijdend gedrag, beïnvloedt het effect van behandeling op relatief vaak voorkomende psychosociale behandelingsbehoeften (laag zelfbeeld, emotionele eenzaamheid en een externe locus of control; hulpeloosheid of defaitisme) mogelijk negatief. Een verschuiving naar een meer holistisch paradigma in de behandeling van jeugdigen met ernstig seksueel grensoverschrijdend gedrag wordt daarom voorgestaan. Door de behandeling explicieter te richten op toenaderingsdoelen (intrinsiek motiverende doelen), en minder prominent op vermijdingsdoelen (door middel van terugvalpreventie), wordt de behandeling meer afgestemd op het verhogen van het individuele welzijn. Hiermee wordt verwacht dat behandelresultaten kunnen worden verbeterd én terugval in (seksueel) grensoverschrijdend gedrag wordt voorkomen.
DANKWOORD (ACKNOWLEDGMENTS)

Staan op de schouders van reuzen, zo voelt het om een proefschrift samen te stellen. Kennis verzamelen over probleemgedrag is zonder alle anderen die zich ook hiermee bezig (hebben ge)houden, een onmogelijkheid. Bij deze mijn eerste grote dankbetuiging aan al deze mensen.

Maar ik heb ook specifieke mensen te danken voor hun bijdrage aan dit proefschrift, en specifiek aan mijn persoonlijke ontwikkeling en welbevinden. Dat laatste is, zoals iedereen die dit proces eens heeft meegemaakt weet, ook een belangrijk punt van aandacht.


Daarnaast ben ik natuurlijk veel dank verschuldigd aan collega’s en studenten die mee hebben gewerkt aan dataverzameling en dataverwerking: Ina Somsen (!), Chris Rijnders, Matthijs Sturm, Corinne Peeters, Olga van Bellen, Jorrin van Dellen, Chantal Goertz en Hieke Deen, wil ik graag bij naam noemen en bij deze bedanken voor hun inzet. Joan van Horn, dank voor je bijdrage aan hoofdstuk 5 en je interesse en steun. Anouk Spruit, UvA collega en coauteur, dankjewel voor het zijn van een vraagbaak. Statistiek was voor mij toch wel erg lang geleden. Fijn dus om altijd even te kunnen afstemmen, inclusief over wat er ’s avonds allemaal aan gekkigheid voorbij komt op TV. Harmke Leloux, Horizon collega, fijn om af en toe ook iemand anders bezig te zien met dezelfde thema’s en te kunnen sparren / delen.
Appendices

Mijn werkgever, Horizon Jeugdzorg en Onderwijs, en in het bijzonder Krijnie Schotel, wil ik danken voor het stellen van de vraag, en bieden van de mogelijkheid om ook daadwerkelijk onderzoek te doen. Ik heb de bijzondere vrijheid gekregen om in mijn eigen forensische interessegebied te werken. Daarnaast kost (promotie)onderzoek wel héél veel tijd, de praktijk is allang vijf stappen verder als er eindelijk weer eens resultaten zijn. Maar goed, ze zijn er! Ik hoop, en verwacht, dat mijn werk praktisch van nut zal zijn. Aan de andere deelnemende instellingen die zo goed zijn geweest om mij in hun keuken toe te laten, de Viersprong en RJJI Den Hey-Acker / DJI-JJI; de groep jeugdigen die ik onderzocht is erg klein, en vergelijkend (beschrijvend) onderzoek is erg spannend. Veel dank voor het vertrouwen, zonder jullie deelname was het niet gelukt. Tenslotte, veel dank aan alle anonieme jongeren die de (veel te lange) testen hebben ingevuld en hun gegevens beschikbaar hebben gesteld. Onderzoek doen duurt lang, tests invullen ook…. Maar het was niet voor niets, ik hoop van harte dat jullie het met de conclusies eens kunnen zijn en we samen (toewijzing aan) behandeling kunnen verbeteren.

Dan mijn persoonlijk welbevinden. Pap en Mam, zonder ook maar iets te begrijpen van waarom ik in godsnaam zoveel op dat kamerj dump te typen en waarom ik nóg meer moest ‘studeren’, jullie zijn er altijd. Zonder jullie was ik, maar ook dit, niet mogelijk. Een stabiele basis is heel veel waard, zo getuigt ook dit proefschrift. Dank voor alle ondersteuning die ik mag genieten, al mijn hele leven lang. Agnes, ‘my person’, wat moet iemand zoals ik zonder een beetje echt begrip? (en af en toe een schop onder mijn….). Let’s get fabulously old together! Mijn andere paranif, Fatima, SPH-maatje, wat een lange weg zijn we beiden inmiddels gegaan. Thanks for reminding me aan mijn jonge(re) enthousiaste zelf. Rosanna, dank voor je luisterend oor, alle kopjes thee en je creativiteit op de omslag van dit proefschrift, een kers op mijn taart…..

Olav, blijkbaar; opposites attract. Stabiele factor, (gelukkig) niet snel onder de indruk van mijn temperament, no-nonsense en meester in het ‘heanig an’. We verwonderen ons regelmatig over elkaar. Hoe mooi kan het leven zijn, enne: Ja…. Dan mijn spiegeltjes: Sterre, al lang samen, maar niet vanaf het begin, en toch ‘heb’ je iets van me. Ook van jou leer ik, over mezelf, over jou en over zorgen voor (een puber ©). Laten we hopen dat ik nog lang bij je in de leer mag zijn. Tenslotte, last but not least, Lumen, wat een wonderlijk cadeautje ben je. Ik kijk naar jou en zie mezelf, ik had vooraf niet bedacht hoe leerzaam dat kan zijn. Ik zie er naar uit om te ervaren hoe jij volwassen wordt, welke keuzes je maakt, en jouw avonturen mee te beleven. Kom je nog vaak even met me knuffelen? Daar word ik zo blij van.
“There are, it has been said, two types of people in the world. There are those who, when presented with a glass that is exactly half full, say: this glass is half full. And then there are those who say: this glass is half empty. The world belongs, however, to those who can look at the glass and say: What’s up with this glass? Excuse me? Excuse me? This is my glass? I don’t think so. My glass was full! And it was a bigger glass! Who’s been pinching my beer? And at the other end of the bar the world is full of the other type of person, who has a broken glass, or a glass that has been carelessly knocked over (usually by one of the people calling for a larger glass) or who had no glass at all, because he was at the back of the crowd and had failed to catch the barman’s eye.”

— Terry Pratchett, The Truth

Final thanks to Sir Terry for having it all figured out already.
CURRICULUM VITAE

Ellis ter Beek was born in Enschede on September 29, 1975. She studied Social Pedagogical Care and worked in forensic youth care (FYC) as a group counselor from 1997 to 2000. She then studied Remedial Sciences; Family and Behavior studies. During this, she continued to work in FYC as a psycho-diagnostic judicial researcher. After obtaining her master’s degree in 2003, she developed training programs and counseling methods for juveniles and personnel in mandated youth care. She also obtained her ‘Orthopedagoog-Generalist’ (Remedial Generalist) registration (2009). She became HRD-manager and, subsequently, head of treatment of an FYC institution and a Secure Youth Care (SYC) institution, and general manager of another SYC institution. In 2015, she made a transition to senior researcher, while performing Ph.D. research on treatment of juveniles with severe sexual problem behavior. Ellis has clinical expertise on juveniles with a several complex issues. Her personal area of interest is comorbidity of psychiatric disorders, borderline intellectual functioning and severe (forensic) behavioral problems in adolescents, focusing explicitly on the juvenile in interaction with his/her context.

PUBLICATIES² (LIST OF PUBLICATIONS)


---

² in volgorde van verschijnen / in order of appearance
UITNODIGING

Voor het bijwonen van de openbare verdediging van het proefschrift

TO TREAT or
NOT TO TREAT?

Harmful Sexual Behavior in Adolescence:
Needs before Risk

Door Ellis ter Beek

Woensdag 14 februari 2018 om 12.00 uur in
de Agnietenkapel, Oudezijds Voorburgwal
229-231, 1012 EZ te Amsterdam

Receptie na afloop ter plaatse

Paranimfen
Agnes de Lima-Heyns
Fatima Ben Allouch

Voor vragen:
DrEllisPromoveert@gmail.com

IELS TER BEEK

TO TREAT or
NOT TO TREAT?

Harmful Sexual Behavior in Adolescence:
Needs before Risk

ELLIS TER BEEK

15206-terbeek-cover.indd   1
19/12/2017   12:24