

# A Quasi-Experimental Study on the Effectiveness of Dutch Cell Dogs in Incarcerated Youth

International Journal of  
Offender Therapy and  
Comparative Criminology  
2021, Vol. 65(5) 644–666  
© The Author(s) 2021



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0306624X20983748  
journals.sagepub.com/home/ijo



Hanne M. Duindam<sup>1</sup> , Hanneke E. Creemers<sup>1</sup>,  
Machteld Hoeve<sup>1</sup>, and Jessica J. Asscher<sup>1,2</sup> 

## Abstract

This study examined the effectiveness of Dutch Cell Dogs (DCD), a prison-based dog training program that aims to improve socioemotional functioning of incarcerated youth by giving them the opportunity to train a shelter dog. Primary (aggression and institutional infractions) and secondary (wellbeing and therapeutic functioning) outcomes were assessed for the intervention ( $n = 61$ ) and comparison group ( $n = 77$ ) before the start of DCD, after 4 weeks, and after 8 weeks at posttest. Overall, DCD participation was not effective. Compared to the comparison group, institutional infractions decreased in DCD participants with an immigrant background and increased in DCD participants with a native Dutch background. In addition, DCD participation reduced the quality of the therapeutic alliance for younger participants and those in secure residential facilities. The current study demonstrated heterogeneity in DTP responsiveness. Future research with robust designs and sufficiently large samples is needed to further identify who benefits from DTPs.

## Keywords

prison-based dog program, incarceration, youth, aggression, wellbeing, correctional program

<sup>1</sup>University of Amsterdam, the Netherlands

<sup>2</sup>Utrecht University, the Netherlands

## Corresponding Author:

Hanne M. Duindam, Forensic Child and Youth Care Sciences, University of Amsterdam, Nieuwe Achtergracht 127, Amsterdam, 1018 WS, the Netherlands.

Email: h.m.duindam@uva.nl

## Introduction

Despite the wide variety in correctional practices around the world, there is growing consensus that rehabilitation of youth offenders is essential in reducing criminal behavior. Correctional programs for incarcerated youth are an important effort to off-set criminal life trajectories and several have been found to improve wellbeing and decrease reoffending rates (e.g., cognitive behavioral therapy, substance abuse treatment; Lipsey et al., 2010; Lipsey, 2009; Wilson, 2016). Programs that promote skill building, personal development, and self-insight have been found to be more effective than interventions oriented toward restriction and control (e.g., Lipsey et al., 2010; Lipsey, 2009). One type of intervention that aims to foster skill building and positive attitude development for incarcerated youth are prison-based dog training programs (DTPs). The goal of the current study was to examine the effectiveness of Dutch Cell Dogs—a community-service DTP—on primary (i.e., aggression, institutional infractions) and secondary (i.e., wellbeing and therapeutic functioning) outcomes in a sample of youth who were incarcerated for committing offences.

In correctional facilities, community-service DTPs such as Dutch Cell Dogs (DCD) are the most common form of prison-based dog programs (Cooke & Farrington, 2014). Even though there is variability among community-service DTPs in terms of program characteristics, such as duration (few weeks to ongoing) and access to the dog (24-7 vs. part-time), the overall goal of DTPs is for incarcerated individuals to train and care for shelter dogs to improve the dog's adoption chances. The aim is to create a "win-win" situation: asylum dogs get adopted into a new home thanks to the behavioral training they receive and the DTP participants' wellbeing and behavior are expected to improve due to them helping the dog through positive reinforcement (Leonardi et al., 2017). DCD is the only DTP offered in correctional facilities in the Netherlands. It is an 8 weeks long community-service DTP during which incarcerated people train, groom, and play with an asylum dog, to whom they have been matched, biweekly for 2 hours. Compared to some other DTPs, DCD is relatively short (8 weeks) and participants have access to "their" dog during training hours only.

Incorporating DTPs in correctional programming stems from the medical field, where Animal Assisted Therapy (AAT) has been implemented to tackle a wide variety of mental and medical problems (Furst, 2006; Nimer & Lundahl, 2007). DTPs are different from AAT in that DTPs do not have a therapeutic focus and pre-defined therapeutic goals, moreover, there are no "patients" in the program who need to be "helped." However, DTPs may improve participants' rehabilitation chances and increase overall wellbeing by harvesting the benefits of the human-animal bond, similar to AATs. Animals can have a calming effect by deactivating stress reactions, but they can also be activating by stimulating human's intrinsic motivation (Beetz, 2017). Accordingly, a wide array of psychosocial and physiological benefits of the human-animal bond has been identified, such as reduced stress and anxiety, enhanced social support, and improved communication (Beetz et al., 2012).

By forming a positive attachment to the dog during DTPs, offenders may become more trusting and open to relationships with others, including therapeutic professionals

and prosocial peers, which is important as poor attachment patterns are associated with delinquency (Hoeve et al., 2012; Jasperson, 2010; Leonardi et al., 2017). Dogs can be a source of social support particularly for offenders in closed settings with higher rates of insecure attachment patterns, as it may be more difficult for them to experience this support from contact with humans (Beetz, 2017). Interacting with a dog to whom one feels attached, has been linked to an increase in oxytocin, which in turn is related to displaying more prosocial behavior (Beetz et al., 2012). In addition, dogs provide offenders with instant behavioral feedback during DTPs due to their tendency to directly respond to environmental stimuli, which may enhance offenders' self-awareness and stimulate the development of more effective emotion regulation skills (Duindam et al., 2020). Moreover, DTPs can provide offenders with a sense of purpose during incarceration and contribute toward a more positive "anti-criminal" identity. By helping a shelter dog, incarcerated individuals may experience the benefits of prosocial behavior and conforming to societal norms (Hill, 2018). In general, achieving goals and targets during the training of the dogs may also boost self-esteem (Fournier et al., 2007).

Some empirical evidence exists in support of these hypothesized effects of DTPs for incarcerated individuals. For example, identified benefits of DTP participation are a reduction in anxiety and stress and an improvement in self-esteem, self-control, and social-emotional regulation skills (e.g., Cooke, 2014; Cooke & Farrington, 2016; Flynn et al., 2019; Leonardi et al., 2017). However, most of this support for DTPs comes from studies with significant methodological limitations (e.g., no control group, small sample sizes, cross-sectional design; Cooke, 2014; Flynn et al., 2019). A recent meta-analysis synthesized findings from (quasi)-experimental research on the effectiveness of prison-based dog programs up until now (Duindam et al., 2020), resulting in the inclusion of nine studies on DTPs. No effect was found for social-emotional functioning, however, prison-based dog programs were found to reduce recidivism although this was based on three retrospective studies only.

For incarcerated adolescents in particular, positive rehabilitating effects have been reported in studies with a qualitative design (e.g., Leonardi et al., 2017; Smith & Smith, 2019), suggesting DTPs may be promising for youth offenders. However, controlled studies focusing on DTPs for this group are scarce: the recent meta-analysis (Duindam et al., 2020) included only three studies that were conducted on incarcerated youth (e.g., Chianese, 2009; Grommon et al., 2018; Seivert et al., 2016) and these reported mixed findings about the DTP's effectiveness. Seivert et al. (2016) found that anxiety and empathy worsened slightly overtime in both the DTP and active control group; no changes were found in externalizing behavioral problems. Grommon et al. (2018) found no change in a variety of psychosocial factors (e.g., self-esteem, empathy, compassion, social competence) after participation in a DTP. Chianese (2009), on the other hand, concluded that DTPs are a promising intervention for incarcerated adolescent girls, as program participants reoffended at only half the rate compared to girls who had no dog contact. Differences in findings may be due to variation in program (e.g., limited vs. fulltime presence of the dogs), study design (e.g., RCT/quasi-experimental; prospective vs. retrospective design), and/or sample characteristics (e.g., male vs. female offenders).

In sum, previous research has been contradicting and more experimental studies are needed to get insight into the effectiveness of DTPs for youth offenders. The current study added to the limited body of research on the effectiveness of DTPs for incarcerated youth by conducting a prospective, quasi-experimental study in a relatively large sample of youth. To our knowledge, this is also the first DTP study to be conducted across multiple youth facilities, which improves generalizability of the findings. Moreover, moderators were included to examine program effectiveness. Studying the overall effectiveness of DTPs for incarcerated youth is important, however, just as important is it to also examine for whom the programs may work and under what conditions by conducting moderator analyses (e.g., Kraemer et al., 2002). This can help answer questions such as whether responsivity to DTPs differs among subgroups of youth offenders, thereby facilitating program assignment decisions in practice and giving direction to focus areas for future research on DTPs (Kraemer et al., 2002).

In line with the program's rationale, we examined to what extent DCD participation was associated with improvement in primary and secondary outcomes. Primary outcomes were identified as aggression and institutional infractions, which are risk factors of recidivism that may be positively impacted by DTP participation (Andrews et al., 2006; van Wormer et al., 2017). Secondary outcomes included those factors potentially associated with DTP participation that are likely to be protective factors of recidivism, such as wellbeing (i.e., self-esteem, self-control, stress, anxiety/depression, life satisfaction, withdrawn behavior, and empathy) and therapeutic functioning (i.e., treatment motivation, therapeutic alliance; Jolliffe & Farrington, 2004; Olver et al., 2011; Silver et al., 2008). Furthermore, we examined to what extent participants' age and cultural background influenced the effectiveness of the DTP under study. These moderators were chosen because previous research has demonstrated differential effects of prison-based DTPs depending on the age of participants (Duindam et al., 2020), and conflicting evidence with regards to the influence of cultural background on the effectiveness of correctional programs (Shearer et al., 2001; Usher & Stewart, 2014). Finally, we added facility type (i.e., correctional facility and secure residential youth care center) as a moderator to examine its influence on the effectiveness of DTPs. In line with the assumed benefits of DTPs, we hypothesized that DCD participation would be associated with an improvement in primary and secondary outcomes.

## Methods

### Design

A pre-posttest quasi-experimental design with two conditions (intervention and comparison) was used. Participants in the intervention and comparison group received treatment-as-usual (TAU) as offered at the respective facilities. In addition, participants in the intervention group took part in Dutch Cell Dogs (TAU+DCD), which we will refer to as DCD. Due to the relatively low number of DCD applicants, it was not possible to randomly assign participants to study condition. For this study, data were

collected at three different time points: before the start of the DCD training (T1), half-way through the training (T2), and at the end of the training (T3).

### **Participants**

This study took place in six mandatory treatment facilities in the Netherlands, including four correctional facilities and two secure residential youth care centers. The aim of these mandatory treatment settings in the Netherlands is to provide professional care for youth with severe behavioral problems (Roest et al., 2016). Even though these youth are similar in terms of behavioral difficulties (e.g., Roest et al., 2016), placement in correctional facilities is mandated by a judge based on juvenile penal law for 12- to 23-years-olds who have committed an offense. Placement in secure residential care, on the other hand, is based on civil law for 12- to 18-years-olds who are a risk to themselves and/or others (Eltink et al., 2018).

The target sample size for this study was determined a priori, 128 participants were needed to detect a medium effect size (.25), given a power of .80 and an alpha of .05 (Schenk et al., 2018). The intervention group consisted of ( $n=61$ ) participants who were selected for the DCD program between 2017 and 2019. Of the approached youth, only one refused participation in the study. The comparison group (TAU;  $n=77$ ) consisted of youth residing in the same facilities in the same period as the DCD participants, however, they followed the regular daily activities.

In Table 1 the characteristics of the participants are presented. As shown, a majority of the participants were male (83.3%). The average age of participants was 18.4 years ( $SD=2.5$ ; range = 12.9–25.5 years-old) and the majority of participants had a first or second-generation immigrant background (59.4%). A smaller group of participants had a native Dutch cultural background (40.6%). Additionally, 73.9% of participants were incarcerated offenders residing in a correctional facility; they were between 14.9 and 25.5 years-old. The rest (26.1%) were juveniles between the ages of 12.9 and 17.5 years-old, who stayed in a secure residential youth care center. In general, there were no significant differences between DCD and TAU in terms of demographics (see Table 1) and pre-intervention scores (see Table 2).

Despite extensive efforts to follow up with participants, 17 (12.3%) and 28 (20.3%) participants at the second (T2) and third (T3) wave respectively, did not continue to participate in the research for a variety of reasons (e.g., not motivated, transferred to another—unknown—facility). When comparing participants who dropped out versus participants who completed all assessments, there were no differences in terms of background variables and pre-intervention scores.

### **Procedure**

Prior to the start of data collection, the Ethical Committee of the Faculty of Social and Behavioral Sciences of the University of Amsterdam approved study procedures (No. 2015-CDE-6363). Participants were recruited for the intervention group during the initial meeting between DCD staff and participants, where the research project was

**Table 1.** Background Characteristics of Research Participants.

	DCD (n=61)		TAU (n=77)		F
	M (range)	SD	M (range)	SD	
Mean age (years)	18.4 (12.9–23.2)	2.5	18.5 (13.5–25.5)	2.5	0.043
Average incarceration (years) <sup>a</sup>	0.6 (0.4–2.1)	0.5	0.8 (0.0–10.11)	1.5	0.890
	%	N	%	n	$\chi^2$
					0.252
Response rate T1	100.0	61	100.0	77	
T2	93.4	57	83.1	64	
T3	80.3	49	79.2	61	
Type of offense <sup>b</sup>					
(Attempted) homicide	8.2	5	2.6	2	—
Violent behavior	24.6	15	23.4	18	
Theft or fraud	1.6	1	3.9	3	
Sexual offences	11.5	7	5.2	4	
Other and unknown	34.4	21	33.8	26	
Residential youth care	19.7	12	31.2	24	
Offense category					1.275
Single	12.5	6	9.4	5	
Mix	45.8	22	37.7	20	
Unknown	41.7	20	52.8	28	
Cultural background					0.375
Native Dutch	37.7	23	42.9	33	
1st or 2nd generation immigrant	62.3	38	57.1	44	
Educational background					7.309
None or primary education	18.0	11	6.5	5	
Secondary education	37.7	23	53.2	41	
Tertiary education	24.6	15	28.6	22	
Other or unknown	19.7	12	11.7	9	
Psychiatric diagnosis					0.353
Yes	50.8	31	55.3	42	
No	13.1	8	10.5	8	
Unknown	31.6	22	34.2	26	
Type of facility					2.333
Youth correctional	19.7	49	68.8	53	
Secure residential youth care	80.3	12	31.2	24	

Note. DCD=Dutch Cell Dogs; TAU=treatment-as-usual; n=sample size; M=mean; SD=standard deviation; offense category—single=individual is currently serving sentence based on a single offense; offense category—mix=individual is currently serving sentence based on multiple offenses.

<sup>a</sup>This information was only available for a subset of our sample (n=93). <sup>b</sup>Categorization based on most severe crime, p<.05, “—” Chi square test could not be performed because categories were insufficiently filled.

**Table 2.** Comparison of DCD and TAU on Pre-Intervention (T1) Scores.

	DCD (n=61)		TAU (n=77)		F <sup>b</sup>
	M (range)	SD	M (range)	SD	
Primary outcomes					
Aggression	0.41 (0.00–1.18)	0.32	0.34 (0.00–1.35)	0.34	1.743
(T1) <sup>a</sup> Institutional infractions	4.43 (0.00–21.00)	5.51	2.93 (0.00–10.00)	2.82	2.452
Secondary outcomes					
Wellbeing					
Self-esteem	2.17 (0.20–3.00)	0.60	2.00 (0.00–3.00)	0.65	2.658
Self-control	3.36 (2.00–4.50)	0.61	3.32 (2.00–4.67)	0.60	0.183
Stress	1.64 (0.40–4.00)	0.80	1.62 (0.00–3.30)	0.71	0.011
Anxiety/depression	0.37 (0.00–1.92)	0.47	0.37 (0.00–1.38)	0.34	0.004
Life satisfaction	3.54 (1.00–7.00)	1.36	3.71 (1.00–7.00)	1.36	0.479
Withdrawn behavior	0.40 (0.00–1.33)	0.35	0.45 (0.00–1.44)	0.35	0.497
Empathy	3.35 (2.40–4.65)	0.48	3.27 (1.30–4.44)	0.58	0.642
Therapeutic functioning					
Therapeutic alliance	2.62 (1.16–3.84)	0.66	2.7 (1.05–3.89)	0.64	0.480
Treatment motivation	2.03 (1.09–3.00)	0.44	2.04 (1.09–2.91)	0.40	0.036

Note. DCD=Dutch Cell Dogs; TAU=treatment-as-usual; M=mean; SD=standard deviation.

<sup>a</sup>These data were only available for a subset of the sample (n=80). <sup>b</sup>No significant group differences were found at the  $p < 0.05$  level.

briefly introduced. Participants for the comparison group were recruited by the researchers and/or designated DCD contact persons (e.g., prison officer, psychologist) at the respective facilities. Furthermore, flyers and posters were distributed across the facilities to familiarize inmates with the research project. Before study participation, written informed consent was obtained from all participants.

Upon completion of each assessment (T1, T2, T3), participants received a small stipend (e.g., €5) or gift (e.g., shower gel, candy). At T1, reading assistance was offered to participants; at subsequent measurements assistance was only offered if needed. The assessments took place in a private, quiet room at the respective facility and each assessment lasted for approximately 60 to 90 minutes. Participants were asked to fill out questionnaires on a laptop. For a more elaborate description of the research procedures, we refer to our study protocol (Schenk et al., 2018) and trial registration (Netherlands National Trial Register TC=6894).

### Study Condition

*Intervention group (Dutch Cell Dogs).* DCD is a prison-based DTP where incarcerated people train, take care of, and play with an asylum dog biweekly during 2 hour long sessions for 8 weeks. In each DCD training, there is a maximum of six dog-participant

dyads and two DCD staff leading the sessions. The two DCD staff members are certified canine instructors with extensive group training experience. Dog-participant pairs, and DCD staff, remained the same throughout the trainings. A big emphasis is placed on responsibility and dedication of participants during the program; the dogs are reliant on participants to be trained in order to have a better chance for adoption. Therefore, truancy (two session or more) results in program termination.

Incarcerated individuals who would like to participate in DCD have to write a motivation letter. Inclusion criteria for the program are: a level of physical/mental fitness that allows for training participation as well as a stay in the facility for the entire program duration. Staff of participating facilities select DCD participants. In addition to DCD, participants continue to receive TAU at their respective facilities, similar to the comparison group. DCD staff match each participant to their own asylum dog based on behavioral observations during an intake meeting and potential goals. For example, a “busy” dog may be paired with a participant with “similarly” (hyper)active behavior in order to foster self-insight. The DCD training consists of four phases. During the first *intake phase*, several days prior to the first training session, the incarcerated juveniles express motivation to join the program and are matched to a personal shelter dog. During the second phase, week 1 to 7, the *dog training in theory and practice*, DCD staff teach training content and supervise participants’ practice. Examples of training topics are: teaching the dog desired behavior and how to respond to basic commands, taking care of the dog, and relaxation by playing. Incarcerated juveniles learn to observe, understand, and anticipate body language, emotions and behaviors of the dog. During the third phase (week 8), *Graduation day*, incarcerated juveniles and their dogs demonstrate what they have learned to those interested (e.g., staff of correctional facilities, family members, shelter personnel) and receive a certificate, a large picture of their shelter dog, and a DCD T-shirt for successful completion of the program. During the fourth phase, *evaluation*, which takes place about 1 month after termination of the program, incarcerated juveniles meet with DCD staff to evaluate the program experience. Participants are invited to provide feedback and various topics are discussed (i.e., current living situation of the dog, anecdotes of training experience, the experience of saying goodbye to the dog). For a more detailed description of training phases and content, we refer to our study protocol (Schenk et al., 2018).

*Comparison group (TAU).* Participants in the comparison group had to meet the same inclusion criteria as the DCD participants. However, participants in the comparison group did not want to participate in DCD or could not participate due to various reasons (e.g., no time, scheduling conflict). Similar to the DCD group, participants in the comparison group received TAU (e.g., Multisystemic Therapy, Cognitive Behavioral Therapy, Occupational Therapy).

## Outcome Measures

*Primary outcomes.* Primary outcomes included aggression and the number of institutional infractions participants were involved in prior to and during DCD. *Aggression*



was assessed using the Dutch version of the aggression subscale of the Youth Self Report form (YSR; Achenbach & Rescorla, 2001), which consists of 17 items (e.g., *I argue a lot*), to be answered on a 3-point Likert-type scale, ranging from 0 (*not true*) to 2 (*often true*). The mean of item scores was taken to form a total score. In the current study, Cronbach alpha's indicated good reliability ( $\alpha$  T1=0.900;  $\alpha$  T2=0.879;  $\alpha$  T3=0.913).

The number of *institutional infractions* in the 2 months prior to the start of DCD and during the (8 weeks) DCD program were (upon permission from participants) subtracted from case files by the researcher or research contact person at the respective facilities. Institutional infractions include instances such as fighting and drug use and are logged by prison officers and mental health professionals at the correctional and secure residential facilities respectively. To register infractions, respondents had to select how many infractions occurred in the described time period, answer options included 0, 1, 2, . . ., up to 20 or *more than 20* (for this latter category, infractions were set to 21 for the analyses). Table 2 includes the means and standard deviations of the amount of infractions per study group before the start of DCD (T1).

**Secondary outcomes.** All secondary outcomes (i.e., wellbeing and therapeutic functioning) were assessed using Dutch versions of validated scales. Higher total scores indicated a larger level of the measured construct. Negatively worded items were reverse-coded before calculating the total score and the mean of item scores was taken to form a total score.

*Wellbeing* was assessed by examining participants' levels of self-esteem, self-control, stress, anxiety/depression, withdrawn behavior, overall life satisfaction, and empathy. *Self-esteem* was measured using the Dutch translation of the Rosenberg's Self Esteem Scale (RSES; Rosenberg, 1979). The RSES consists of 10 items (including five negatively-worded items) scored on a 4-point Likert-type scale, ranging from 0 (*strongly disagree*) to 3 (*strongly agree*); an example item is: *On the whole, I am satisfied with myself*. The RSES demonstrated good internal consistency in the current study ( $\alpha$  T1=0.890;  $\alpha$  T2=0.842;  $\alpha$  T3=0.857).

*Self-control* was measured by using the Brief Self Control Scale (BSCS) consisting of 13 items, including nine negatively-worded items (Tangney et al., 2004). Scale items (e.g., *I often act without thinking through all the alternatives*) were scored on a 5-point Likert-type scale, ranging from 1 (*not at all*) to 5 (*very much*). For the present study, the BSCS demonstrated insufficient internal consistency ( $\alpha$  T1=0.607;  $\alpha$  T2=0.513;  $\alpha$  T3=0.569). However, to improve internal consistency, item 6 was removed at all time points resulting in better Cronbach's alpha's (i.e.,  $\alpha$  T1=0.731;  $\alpha$  T2=0.668;  $\alpha$  T3=0.701).

*Stress* was assessed by using 10 items of the Perceived Stress Scale, including four reverse-coded items (PSS; Cohen et al., 1983). Items (e.g., *In the last month, how often have you felt nervous and stressed*) are answered on a 5-point Likert-type scale ranging from 0 (*never*) to 4 (*very often*). The internal consistency of PSS in the present study was good ( $\alpha$  T1=0.815;  $\alpha$  T2=0.817;  $\alpha$  T3=0.795).

*Anxiety and depression* were measured by using the anxious/depressed subscale of the YSR (13 items; Achenbach & Rescorla, 2001, 2003). Items (e.g., *I feel lonely*) were answered on a 3-point Likert-type scale, 0 (*not true*) to 2 (*often true*). Internal consistency of this subscale was good ( $\alpha$  T1=0.900;  $\alpha$  T2=0.879;  $\alpha$  T3=0.913).

*Withdrawn behavior* was assessed by using the Withdrawn behavior subscale of the Adult Self Report (ASR; Achenbach & Rescorla, 2003) form (e.g., *I would rather be alone than with others*), which consists of nine items to be answered on a 3-point Likert-type scale, ranging from 0 (*not true*) to 2 (*often true*). Internal consistency of this subscale was good ( $\alpha$  T1=0.788;  $\alpha$  T2=0.838;  $\alpha$  T3=0.845).

Overall *life satisfaction* was measured by using the Satisfaction With Life Scale (SWLS; Diener et al., 1985). Participants are asked to what extent they agree with five statements about their life (e.g., *In most ways my life is close to ideal*) on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Overall, internal consistency of the SWLS was good ( $\alpha$  T1=0.796;  $\alpha$  T2=0.854;  $\alpha$  T3=0.907).

*Empathy* was assessed using the Basic Empathy Scale (BES; Jolliffe & Farrington, 2006; Van Langen et al., 2009). The BES consists of 20 items (e.g., *I can usually work out when people are scared*), including eight negatively-worded items (e.g., *My friends' emotions don't affect me much*), to be answered on a 5-point Likert-type scale, ranging from 1 (*completely disagree*) to 5 (*completely agree*). Overall, internal consistency of the BES was good ( $\alpha$  T1=0.835;  $\alpha$  T2=0.864;  $\alpha$  T3=0.881).

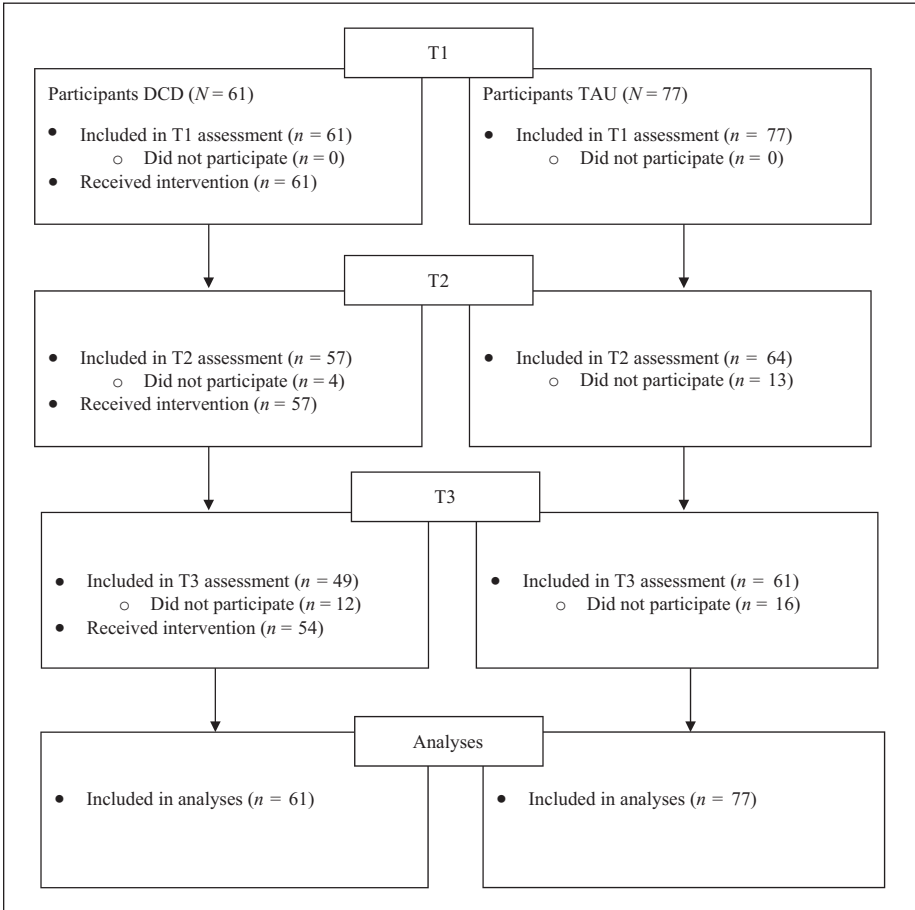
*Therapeutic functioning* was measured by participants' therapeutic motivation and their perception of the therapeutic alliance with their therapist, mentor, or prison officer. *Therapeutic motivation* was assessed with the Adolescent Treatment Motivation Questionnaire (ATMQ; Van Der Helm et al., 2012), which consist of 11 items about therapeutic motivation (e.g., *My treatment helps me*) to be answered on a 3-point Likert-type scale ranging from 1 (*not true*) to 3 (*true*). Internal consistency of the ATMQ was good ( $\alpha$  T1=0.802;  $\alpha$  T2=0.843;  $\alpha$  T3=0.879).

The *therapeutic alliance* was measured by using the Psychological Availability and Reliance on Adult scale (PARA; Schuengel & Zegers, 2003), which consists of 19 items, including three negatively-worded items, that have to be rated on a 4-point Likert-type scale, ranging from 1 (*disagree*) to 4 (*agree*). Not all participants had a therapist so they were instructed to answer the items about the relationship with their mentor, dedicated ward prison officer or therapist (e.g., *For support and advice you go to your [mentor, therapist, prison officer]*). The internal consistency of the PARA was good ( $\alpha$  T1=0.900;  $\alpha$  T2=0.918;  $\alpha$  T3=0.918).

**Moderators.** Participants' *age* in years at T1 and *cultural background* (native Dutch vs. first or second generation immigrant) were included in the analyses as moderators. In addition, *facility type* (correctional vs. secure residential) was added as moderator.

## Data Analysis

In accordance with intention-to-treat analysis, we included all participants in the analyses to reduce the influence of treatment motivation as a possible confounding effect



**Figure 1.** Flow diagram of participants Dutch Cell Dogs study.

(Montori & Guyatt, 2001). This means that regardless of whether or not participants completed DCD, they were included in the analyses. The majority of the participants in the DCD group completed the training program (88.5%), whereas a minority of seven participants did not complete the training (11.5%; see Figure 1). Missing data were imputed using multiple imputation, by employing the expectation-maximization algorithm (Graham, 2009). Data were imputed separately for the intervention and comparison group, in order to assess program effects in an unbiased manner (Yamaguchi et al., 2020). Self-report questionnaires data were imputed for the 17 participants at T2 (12.3%) and the 28 participants at T3 (20.3%), who did not continue to participate in the research. Institutional infractions were only available for a subset (n=80) of the sample. For other participants this information could not be obtained from case files or, to a smaller extent, no permission for file analysis was provided. For

institutional infractions data were imputed for participants who had missing data at one of the two time-points ( $n=3$ ).

The effectiveness of DCD was assessed by conducting two repeated measures univariate analysis of variance (ANOVA) tests, one for the primary outcome of aggression and one for institutional infractions. Repeated measures multivariate analysis of variance (MANOVA) tests were conducted for the secondary outcome variables. Intervention condition (DCD vs. TAU) was entered as a factor. Effect sizes (Cohen's  $d$ ) were calculated for each outcome variable by using the effect size calculator of the Campbell collaboration, using standardized means and errors.

In addition, moderator analyses were conducted to determine the influence of age, cultural background, and facility type on the overall effectiveness of DCD. The moderators were added as factors to the repeated measures MANOVA's and an interaction between condition  $\times$  moderator  $\times$  time was added. If a moderator effect was significant, post hoc analyses were conducted by running univariate and multivariate analyses for various levels of the moderator.

## Results

Assumptions were tested and met before conducting the analyses, with the exception of the outcomes aggression, institutional infractions, depression/anxiety, and withdrawn behavior. For these outcome measures the normality assumption was violated. Therefore, analyses were repeated using Robust Mean Modeling (RMM), which is a robust method for handling skewed distributions (Fan & Hancock, 2012). The overall results regarding the effects of DCD on these outcomes remained the same (available on request). Results of the repeated measures ANOVA's and repeated measures MANOVA's for the primary and secondary outcomes and Cohen's  $d$  effect sizes are reported in Table 3.

**Primary outcomes.** The analyses examining whether DCD was effective in changing the primary outcomes, revealed no differences between DCD and TAU in change over time for aggression,  $F(2, 272)=0.146$ ,  $p=.864$ , or institutional infractions,  $F(1, 78)=1.352$ ,  $p=.248$ . There were no time effects for primary outcomes.

**Secondary outcomes.** There were no differences between DCD and TAU in change over time for wellbeing (i.e., self-esteem, self-control, stress, anxiety/depression, life satisfaction, withdrawn behavior, and empathy),  $F(14, 123)=1.714$ ,  $p=.061$ , and therapeutic functioning (i.e., therapeutic alliance and treatment motivation),  $F(4, 133)=1.659$ ,  $p=.163$ . Regardless of group condition, three secondary outcomes changed over time: there was an increase in participants' self-control,  $F(14, 123)=5.628$ ,  $p=.006$  and life satisfaction,  $F(14, 123)=4.961$ ,  $p=.008$ , and a reduction in their stress levels,  $F(14, 123)=3.632$ ,  $p=.028$ .

**Moderator analyses.** Significant moderator effects, including mean, standard deviations, and Cohen's  $d$  effect sizes, are displayed in Table 4.

**Table 3.** Effects of DCD ( $n = 61$ ) Versus TAU ( $n = 77$ ) on Primary and Secondary Outcomes.

	DCD										TAU				F	Cohen's $d$ (CI)	Time
	T1		T2		T3		T1		T2		T3		F	Group *Time			
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)					
<b>Primary outcomes</b>																	
Aggression	0.41 (0.32)	0.37 (0.32)	0.38 (0.38)	0.34 (0.33)	0.31 (0.33)	0.33 (0.34)	0.31 (0.33)	0.31 (0.33)	0.31 (0.33)	0.31 (0.33)	0.33 (0.34)	0.31 (0.33)	0.31 (0.33)	0.146	0.066 (-0.271 to 0.402)	1.195	
Infractions	—	4.43 (5.51)	3.60 (4.32)	—	2.93 (2.82)	3.14 (3.66)	—	2.93 (2.82)	2.93 (2.82)	2.93 (2.82)	3.14 (3.66)	2.93 (2.82)	2.93 (2.82)	1.352	0.261 (-0.181 to 0.702)	0.487	
Well-being														1.714†	0.320 (-0.018 to 0.658)	2.181*	
Self-esteem	2.17 (0.60)	2.16 (0.56)	2.03 (0.57)	2.00 (0.65)	2.14 (0.48)	2.12 (0.53)	2.00 (0.65)	2.00 (0.65)	2.14 (0.48)	2.14 (0.48)	2.12 (0.53)	2.00 (0.65)	2.00 (0.65)	0.146	0.096 (-0.240 to 0.432)	1.398	
Self-control	3.36 (0.61)	3.41 (0.63)	3.54 (0.65)	3.32 (0.60)	3.34 (0.53)	3.42 (0.59)	3.32 (0.60)	3.32 (0.60)	3.34 (0.53)	3.34 (0.53)	3.42 (0.59)	3.32 (0.60)	3.32 (0.60)	0.096	0.032 (-0.304 to 0.368)	5.628**	
Stress	1.64 (0.80)	1.60 (0.80)	1.48 (0.69)	1.62 (0.71)	1.56 (0.74)	1.46 (0.72)	1.62 (0.71)	1.62 (0.71)	1.56 (0.74)	1.56 (0.74)	1.46 (0.72)	1.62 (0.71)	1.62 (0.71)	0.032	0.138 (-0.198 to 0.475)	3.632*	
Anxiety/depression	0.37 (0.47)	0.29 (0.39)	0.31 (0.43)	0.37 (0.34)	0.34 (0.33)	0.33 (0.36)	0.37 (0.34)	0.37 (0.34)	0.34 (0.33)	0.34 (0.33)	0.33 (0.36)	0.37 (0.34)	0.37 (0.34)	0.138	0.245 (-0.092 to 0.582)	3.026†	
Life satisfaction	3.54 (1.36)	3.89 (1.56)	3.76 (1.81)	3.71 (1.36)	3.82 (1.46)	4.08 (1.51)	3.71 (1.36)	3.71 (1.36)	3.82 (1.46)	3.82 (1.46)	4.08 (1.51)	3.71 (1.36)	3.71 (1.36)	0.245	0.098 (-0.238 to 0.434)	4.961**	
Withdrawn behavior	0.40 (0.35)	0.36 (0.36)	0.41 (0.37)	0.45 (0.35)	0.42 (0.37)	0.42 (0.37)	0.45 (0.35)	0.45 (0.35)	0.42 (0.37)	0.42 (0.37)	0.42 (0.37)	0.45 (0.35)	0.42 (0.37)	0.098	0.325 (-0.013 to 0.663)	0.630	
Empathy	3.35 (0.48)	3.33 (0.57)	3.23 (0.63)	3.27 (0.58)	3.32 (0.58)	3.34 (0.63)	3.27 (0.58)	3.27 (0.58)	3.32 (0.58)	3.32 (0.58)	3.34 (0.63)	3.27 (0.58)	3.27 (0.58)	1.659	0.110 (-0.226 to 0.446)	0.685	
<b>Therapeutic functioning</b>																	
Therapeutic alliance	2.62 (0.66)	2.57 (0.99)	2.46 (0.73)	2.70 (0.65)	2.64 (0.69)	2.62 (0.69)	2.70 (0.65)	2.70 (0.65)	2.64 (0.69)	2.64 (0.69)	2.62 (0.69)	2.70 (0.65)	2.70 (0.65)	0.110	0.218 (-0.119 to 0.555)	1.346	
Treatment motivation	2.03 (0.44)	2.09 (0.55)	2.04 (0.60)	2.04 (0.40)	1.98 (0.43)	1.99 (0.46)	2.04 (0.40)	2.04 (0.40)	1.98 (0.43)	1.98 (0.43)	1.99 (0.46)	2.04 (0.40)	2.04 (0.40)	0.110	0.218 (-0.119 to 0.555)	1.346	

Note. DCD = Durch Cell Dogs; TAU = treatment-as-usual; M = mean; SD = standard deviation; CI = confidence interval; F = test statistic based on multivariate and univariate analyses. † $p < .10$ . \* $p < .05$ . \*\* $p < .01$ .

**Table 4.** Significant Moderator Effects.

	DCD				TAU				F	Cohen's <i>d</i> (CI)		
	T1	T2	T3	T1	T2	T3	T1	T2			T3	
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)			M (SD)	
Age												
Primary outcomes												
Institutional infractions												
Juvenile ( <i>n</i> = 48)	—	4.38 (6.29)	4.57 (5.28)	—	2.18 (2.22)	2.67 (3.40)	—	2.18 (2.22)	2.67 (3.40)	—	4.673*	0.262 (-0.311 to 0.834)
Adult ( <i>n</i> = 32)	—	4.50 (4.49)	2.31 (2.15)	—	4.19 (3.33)	3.93 (4.04)	—	4.19 (3.33)	3.93 (4.04)	—	0.060	0.511 (-0.193 to 1.215)
Therapeutic functioning												
Therapeutic alliance												
Juvenile ( <i>n</i> = 66)	2.76 (0.65)	2.49 (0.81)	2.44 (0.75)	2.62 (0.67)	2.61 (0.73)	2.66 (0.64)	2.62 (0.67)	2.61 (0.73)	2.66 (0.64)	2.62 (0.67)	3.103*	0.427 (-0.043 to 0.895)
Adult ( <i>n</i> = 72)	2.50 (0.64)	2.64 (1.14)	2.47 (0.73)	2.78 (0.63)	2.68 (0.61)	2.60 (0.74)	2.78 (0.63)	2.68 (0.61)	2.60 (0.74)	2.72	4.441*	0.227 (-0.238 to 0.692)
Therapeutic motivation												
Cultural background												
Primary outcomes												
Institutional infractions												
Native Dutch ( <i>n</i> = 32)	—	2.50 (3.50)	5.17 (4.75)	—	3.10 (3.00)	2.40 (3.44)	—	3.10 (3.00)	2.40 (3.44)	—	16.474***	-1.188 (-1.961 to 0.416)
1st/2nd gen ( <i>n</i> = 48)	—	5.46 (6.10)	2.84 (3.99)	—	2.78 (2.71)	3.78 (3.79)	—	2.78 (2.71)	3.78 (3.79)	—	8.887**	0.861 (0.269 to 1.453)
Facility type												
Therapeutic functioning												
Therapeutic alliance												
COR ( <i>n</i> = 102)	2.49 (0.63)	2.53 (1.04)	2.42 (0.74)	2.68 (0.66)	2.56 (0.64)	2.56 (0.72)	2.68 (0.66)	2.56 (0.64)	2.56 (0.72)	2.68 (0.66)	3.679**	0.159 (-0.230 to 0.548)
SRC ( <i>n</i> = 36)	3.15 (0.48)	2.73 (0.74)	2.60 (0.71)	2.73 (0.64)	2.81 (0.72)	2.77 (0.61)	2.73 (0.64)	2.81 (0.72)	2.77 (0.61)	2.73	3.737*	0.316 (-0.380 to 1.013)
Treatment motivation												

Note. COR = correctional facility; SRC = secure residential care facility; Native Dutch = participants with a native Dutch background; 1st/2nd gen = participants with 1st/2nd generation immigrant background; juvenile < 18 years-old, Adult > 18 years-old; F = test statistic based on multivariate and univariate (post hoc) analyses.

\**p* < .05, \*\**p* < .01, \*\*\**p* < .001.

**Age.** Age moderated the effectiveness of DCD on institutional infractions,  $F(2, 77)=4.673$ ,  $p=.012$  and therapeutic functioning,  $F(4, 131)=3.103$ ,  $p=.018$ . For post hoc analyses, median split was used to create two age categories: juveniles (<18-years-old, 47.8% of the sample) and adults ( $\geq$ 18-years-old, 52.2% of the sample). These analyses showed that DCD participation was not significantly associated with changes in infraction rates in both the juvenile and adult group. In terms of therapeutic functioning, DCD was—relative to TAU—associated with a significant reduction in the quality of the therapeutic alliance in juveniles,  $F(2, 140)=3.248$ ,  $p=.042$ , not in adults.

**Cultural background.** Cultural background moderated the effectiveness of DCD on institutional infractions,  $F(1, 76)=16.474$ ,  $p<.001$ . Post hoc analyses revealed that, compared to TAU, there was a reduction in institutional infractions following DCD for participants with a first or second generation immigrant background,  $F(1, 46)=8.887$ ,  $p=.005$ . In contrast, institutional infractions of DCD participants with a native Dutch background increased over time,  $F(1, 30)=10.588$ ,  $p=.003$ .

**Facility type.** Facility type (i.e., correctional vs. secure residential) moderated the effectiveness of DCD on therapeutic functioning,  $F(4, 131)=3.679$ ,  $p=.007$ . Post hoc analyses demonstrated that the quality of the therapeutic alliance reduced over time for DCD participants in secure residential care, compared to TAU,  $F(2, 68)=4.599$ ,  $p=.013$ . This was not the case for those in correctional facilities.

## Discussion

The present study examined the short term effectiveness of DCD for incarcerated youth in terms of primary (i.e., institutional infractions, aggression) and secondary (i.e., wellbeing and therapeutic functioning) outcomes. Contrary to expectations, no effect of DCD was found for improving primary or secondary outcomes. However, DCD was effective in reducing institutional infractions for participants with a first or second generation immigrant background, whereas infractions increased for DCD participants with a native Dutch background. For juvenile (18–) participants and participants in secure residential care facilities, DCD was associated with a decrease in the quality of the treatment alliance.

The lack of overall improvement in primary and secondary outcomes as a result of DCD participation is consistent with the results of two RCT's on DTPs for youth offenders that did not find positive program effects in externalizing behavioral problems (Seivert et al., 2016) and in psychosocial functioning (i.e., self-esteem, empathy, optimism, pessimism, compassion, and social competence; Grommon et al., 2018; Seivert et al., 2016). This is in contrast with previous research in adults, that did demonstrate positive overall effects for DTPs in primary (i.e., infractions, externalizing behavior problems; e.g., Cooke & Farrington, 2016; van Wormer et al., 2017) and secondary outcomes, such as social-emotional (e.g., Cooke, 2014) and therapeutic (e.g., Fournier et al., 2007) functioning. Furthermore, smaller effect sizes for younger

participants have been found in a recent meta-analysis on prison-based dog programs (Duindam et al., 2020). These findings suggest that the responsiveness of youth offenders to DTPs may differ from that of adults. This may be because generally the impact of dynamic risk factors on primary outcomes (i.e., recidivism) is stronger with increasing age (Spruit et al., 2017). In late adolescence, only a weak association between dynamic risk factors and recidivism has been found, which may explain why the chance of success of interventions that aim to target these dynamic risk factors (such as DTPs) is lower in adolescence (Spruit et al., 2017).

The importance of the role of age is also supported by results of the moderator analyses. Although post hoc analyses failed to reach significance (potentially due to a lack of power), mean infraction scores suggested an improvement for adults (18+) and not for juvenile (18-) participants following DCD. In addition, for the juvenile (18-) participants, the quality of the therapeutic alliance decreased over time following DCD. This may suggest that DTP participation requires that a certain developmental stage is reached in order to harvest the benefits of the human-animal bond. In general, young adolescents have been found to be more self-centered in relation to their pets due to their stage of moral development (Gage & Christensen, 1990). Juveniles value pets because of the pleasure they bring and personal needs they satisfy (Gage & Christensen, 1990; Kohlberg & Kramer, 1969). In this stage of development, training and caring for a shelter dog may be challenging because program outcomes for participants may not immediately be positive (and require patience). This may mean that, to allow for DTPs to have an effect in this younger subgroup, program delivery needs adjustment in line with the developmental maturity stage of these juveniles (e.g., more compliments, guidance, and feedback of DTP staff may be important). Tailoring DTPs to the developmental level of youth is an interesting area for future research.

Alternatively, the absence of significant effects of DTPs in youth may be associated with program characteristics. In the current and previous research on DTPs for youth, program duration was generally shorter and the dogs did not stay with participants fulltime, while this was the case in research focusing on DTPs for adults. This may account for some of the contradicting findings between age groups.

Another consideration is that most of the quasi-experimental studies on DTPs that found positive program effects have been conducted in correctional facilities in the United States (US), where prison programming is not accessible to all inmates (Dick, 2018). As access to rehabilitation, education, and treatment programs is a legal right of detainees in Western Europe (Dick, 2018), there unavoidably is variation in DTP study conditions across countries. This may account for some of the contradicting findings as program effects would presumably be larger for incarcerated persons in the US for whom the DTP may be the only form of prison program they participate in, especially when compared to a control group that is not enrolled in any form of program. In the current study, all participants had access to various prison programs in addition to TAU, conceivably leaving relatively less room for improvement for the DCD group.

A relatively large proportion of the 18- juvenile participants resided in secure residential care (55%); we also found a reduction in the quality of the therapeutic



relationship for DCD participants in secure residential care (compared to correctional facilities). These findings suggest that in addition to age, facility type may have also played a role in the effectiveness of DCD. Informal observations by the DCD trainers suggest that, particularly in the secure residential care facilities, participation in DCD may have affected the perspective of the juveniles on their relationship with their group leaders and social workers. It may even have resulted in increased awareness of dysfunctional relationships, which may have caused adverse effects.

This touches upon a broader challenge concerning the implementation of programs in secured settings: it is important to consider how positive growth sparked by the program can be maintained in the deprivational and repressing climates of secured facilities (Van Der Helm et al., 2014). For DTPs—and correctional programs in general—it is important to consider positive valuable experiences can be maintained in the difficult prison environment once the sessions have finished. DTPs are recognized for their humane and respectable approach, creating a safe environment for emotional expression (e.g., Aufderheide & Renck Jalongo, 2019). In many ways, this approach is at odds with the “restricted” nature of prisons, often characterized by power imbalance, social isolation, victimization, and negative peer influence (e.g., Lambie & Randell, 2013). To foster growth in these settings, it therefore seems important for DTPs to consider how they can help participants integrate their program experience into daily prison life by, for example, working more closely with staff at the respective facilities.

Finally, we found that institutional infractions reduced only for participants with a first or second generation immigrant background. Even though we found no other AAT studies to have directly examined cultural background as a moderator; previous research about its influence on the effectiveness of correctional program in general has been conflicting (e.g., Usher & Stewart, 2014). One potential explanation for why DCD may have helped those with an immigrant background in particular, is that these youth may particularly improve their connection to society by engaging in prosocial behavior in line with societal norms; one of the hypothesized benefits of DTPs. A lack of social integration, which is linked to a higher likelihood to ignore rules and norms and to engage in criminal behavior according to the Social Control Theory (Hirschi, 1969), has been identified as one of the factors contributing to the engagement in criminal behavior of immigrant minorities in particular (e.g., Junger-Tas, 2001). Improving the connection to society may be particularly important for immigrant minorities who have more to gain in this regard due to their socially marginalized position (Junger-Tas, 2001). Experiencing these social benefits, may boost a positive, anti-criminal identity that helps them refrain from prison misconduct (Hill, 2018), however, more research is needed to test this hypothesis. Another potential explanation for this significant moderator finding, is that it was driven by the unexpected increase in incidents for DCD participants with a native Dutch background, compared to TAU. As we do not have a plausible explanation for this surprising finding more future research is needed.

Strengths of the current study were the inclusion of a relatively large youth sample across several (correctional and residential) facilities, reporting of effect size estimations, and an elaborate description of the studied sample and DTP, which allows for

comparison across studies (May et al., 2016). Another important strength was the inclusion of moderators. Our study has demonstrated the importance of examining the effectiveness of a DTP in subgroups, by including moderators, as this has revealed variation in responsiveness to DCD. This variation would have remained undiscovered when only looking at overall program effectiveness by examining mean scores. Design adjustments that illuminate heterogeneity in program responses are essential for guiding clinical applications of interventions and future research (e.g., Kraemer et al., 2002).

Some limitations of the current study should also be considered. First, the Cohen's *d* effect sizes of our findings all had wide confidence intervals, indicating low precision in effect estimates. This means effects ranged from small (negligible) to large. Future research with larger samples can help provide more accurate estimates of effect size to determine (clinical) relevance of DTPs for youth participants. Second, we employed a quasi-experimental design as opposed to a RCT, which is seen as the golden standard as program allocation (as opposed to random assignment) may introduce bias (Farrington, 2003). At the same time, we found no pre-existing differences between the DCD and TAU, suggesting their levels of functioning (in the measured constructs) were comparable before the start of the program. Another limitation is that we could not assess the extent to which DCD was carried out as intended as there is no standardized instrument developed to examine program integrity. It is important for future research on DTPs to include treatment fidelity measures so that its influence on program effectiveness can be assessed (Duindam et al., 2020). Finally, our findings may not be generalizable to programs with a higher "DTP dosage" where dogs are more integrated into the daily life of participants. DCD is a relatively short program with limited dog access compared to some other DTPs that are continuous with full-time dog access (e.g., dog resides with participants).

Several recommendations for future research can be formulated based on our study, such as examining the long term effects of participating in a DTP: it may be that certain changes are only detectable at a later stage, as social-emotional change generally takes time (Grommon et al., 2018). Therefore, follow-up assessments are important. In addition, including staff perception in future research should also be considered. Most studies on DTPs are on how the youth see themselves; staff may provide another perspective on the effectiveness of DTPs (e.g., Antonio et al., 2017). Moreover, future research should also broaden the scope of types of outcomes measured. Our study—and most previous research—has generally focused on psychological and behavioral constructs (e.g., Duindam et al., 2020), while findings from qualitative studies suggest that DTPs may also positively impact rehabilitation by lessening the criminogenic effect of the prison environment, by, for example, improving detainee-staff interactions or aiding in the development of vocational skills (e.g., Minton et al., 2015). Including such environmental, vocational, and process-oriented measures in future experimental research may provide a broader perspective on the workings of DTPs. The relatively low attrition rate of DCD participants (11.5%), compared to correctional programming attrition rates more generally (19.9%; Olver et al., 2011), may also reflect positive program experiences not fully captured by the included constructs in this study.

In general, there is a need for more robust research on DTPs, which can be said for interventions based on the benefits of the human-animal interaction (HAI) more generally, as most studies in the field have had methodological limitations (e.g., no control group; no post-intervention assessments; e.g., May et al., 2016). To our knowledge, this is the first study that confirms youth' differential responding to a DTP, depending on age, cultural background, and facility type, thereby aiming to identify for whom DCD may (not) work. Future research should continue this search by examining to what extent these and other sample (e.g., gender, offense type, callous-unemotional traits), program (e.g., duration, dog access), and setting (e.g., group climate, integration of DTP at facility) characteristics may moderate program effectiveness. At the same time, it is important for future research to study how DTPs may work (Furst, 2019). Most likely, the same working mechanisms are not applicable to each individual (e.g., Beetz, 2017); therefore, conducting more experimental research to determine for whom DTPs may work, and why, will help the HAI field in general move forward in solidifying its potential to help a population in high-need.

In conclusion, the current study demonstrated that overall, DCD participation was not associated with an improvement in primary and secondary outcomes on top of care as usual. In addition, there was heterogeneity in response to DCD, demonstrating the need for future research to further address potential variations in effectiveness. This research adds to the growing body of more rigorous research on DTPs: a special type of correctional program that has the potential to help incarcerated youth, dogs, and the community at large.


### Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iDs

Hanne M. Duindam  <https://orcid.org/0000-0002-5934-820X>

Jessica J. Asscher  <https://orcid.org/0000-0002-8169-4301>

### References

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms & profile*. University of Vermont, Research Center for Children, Youth, & Families.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, 52(1), 7–27. <https://doi.org/10.1177/001128705281756>

- Antonio, M. E., Davis, R. G., & Shutt, S. R. (2017). Dog training programs in pennsylvania's department of corrections. *Society and Animals*, 25(5), 475–489. <https://doi.org/10.1163/15685306-12341457>
- Aufderheide, C., & Renck Jalongo, M. (2019). Celebrating the achievements of prison dog programs. In M. Renck Jalongo (Ed.), *Prison dog programs: Renewal and rehabilitation in correctional facilities* (pp. 275–303). Springer.
- Beetz, A. (2017). Theories and possible processes of action in animal assisted interventions. *Applied Developmental Science*, 21(2), 139–149. <https://doi.org/10.1080/10888691.2016.1262263>
- Beetz, A., Uvnäs-Moberg, K., Julius, H., & Kotrschal, K. (2012). Psychosocial and psychophysiological effects of human-animal interactions: The possible role of oxytocin. *Frontiers in Psychology*, 3, 1–15. <https://doi.org/10.3389/fpsyg.2012.00234>.
- Chianese, N. (2009). *Girls, jails, and puppy dog tails: An evaluation of the new leash on life program* (Master's thesis). Department of Sociology, California State University.
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385–396.
- Cooke, B. J. (2014). *Bad to the bone?: The effects of dog training programs on factors related to desistance* (Doctoral dissertation). Institute of Criminology, University of Cambridge.
- Cooke, B. J., & Farrington, D. P. (2014). Perceived effects of dog-training programmes in correctional settings. *Journal of Forensic Practice*, 16(3), 171–183. <https://doi.org/10.1108/JFP-08-2013-0041>
- Cooke, B. J., & Farrington, D. P. (2016). The effectiveness of dog-training programs in prison: A systematic review and meta-analysis of the literature. *Prison Journal*, 96(6), 854–876. <https://doi.org/10.1177/0032885516671919>
- Dick, A. (2018). The immature state of our union: Lack of legal entitlement to prison programming in the United States as compared to European countries. *Journal of International and Comparative Law*, 35(2), 287–324. <https://doi.org/10.15439/2019F121>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71–75. <https://doi.org/10.1097/HTR.0000000000000004>
- Duindam, H. M., Asscher, J. J., Hoeve, M., Stams, G. J. J. M., & Creemers, H. E. (2020). Are we barking up the right tree? A meta-analysis on the effectiveness of prison-based dog programs. *Criminal Justice and Behavior*, 47(6), 749–767.
- Eltink, E. M. A., Ten Hoeve, J., De Jongh, T., Van der Helm, G. H. P., Wissink, I. B., & Stams, G. J. J. M. (2018). Stability and change of adolescents' aggressive behavior in residential youth care. *Child and Youth Care Forum*, 47(2), 199–217. <https://doi.org/10.1007/s10566-017-9425-y>
- Fan, W., & Hancock, G. R. (2012). Robust means modeling: An alternative for hypothesis testing of independent means under variance heterogeneity and nonnormality. *Journal of Educational and Behavioral Statistics*, 37(1), 137–156. <https://doi.org/10.3102/1076998610396897>
- Farrington, D. P. (2003). Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Science*, 587(1), 49–68. <https://doi.org/10.1177/0002716202250789>
- Flynn, E., Combs, K. M., Gandenberger, J., Tedeschi, P., & Morris, K. N. (2019). Measuring the psychological impacts of prison-based dog training programs and in-prison outcomes for inmates. *Prison Journal*, 100(2), 224–239. <https://doi.org/10.1177/0032885519894657>
- Fournier, A., Geller, E., & Fortney, E. (2007). Human-animal interaction in a prison setting: Impact on criminal behavior, treatment progress, and social skills. *Behavior and Social Issues*, 16(1), 89–105.

- Furst, G. (2006). Prison-based animal programs. *The Prison Journal*, 86(4), 407–430. <https://doi.org/10.1177/0032885506293242>
- Furst, G. (2019). The empirical evidence supporting dog training programs for incarcerated people. In M. R. Jalongo (Ed.), *Prison dog programs: Renewal and rehabilitation in correctional facilities* (pp. 17–36). Springer.
- Gage, M. G., & Christensen, D. H. (1990). Early adolescents' values about their pets. *Journal of Psychology: Interdisciplinary and Applied*, 124(4), 417–425. <https://doi.org/10.1080/00223980.1990.10543237>
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60(1), 549–576. <https://doi.org/10.1146/annurev.psych.58.110405.085530>
- Grommon, E., Carson, D. C., & Kenney, L. (2018). An experimental trial of a dog-training program in a juvenile detention center. *Journal of Experimental Criminology*, 41(1), 62–80. <https://doi.org/10.1007/s11292-018-9346-2>
- Hill, L. B. (2018). Becoming the person your dog thinks you are: An assessment of Florida prison-based dog training programs on postrelease recidivism. *Corrections, Policy, Practice, and Research*, 5(3), 149–169. <https://doi.org/10.1080/23774657.2018.1433564>
- Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.
- Hoeve, M., Stams, G. J. J. M., Van Der Put, C. E., Dubas, J. S., Van Der Laan, P. H., & Gerris, J. R. M. (2012). A meta-analysis of attachment to parents and delinquency. *Journal of Abnormal Child Psychology*, 40(5), 771–785. <https://doi.org/10.1007/s10802-011-9608-1>
- Jaspersen, R. A. (2010). *Therapeutic interventions and animal assisted therapy with incarcerated females* (Doctoral dissertation). College of Social Work, University of Utah.
- Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior*, 9(5), 441–476. <https://doi.org/10.1016/j.avb.2003.03.001>
- Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of Adolescence*, 29(4), 589–611. <https://doi.org/10.1016/j.adolescence.2005.08.010>
- Junger-Tas, J. (2001). Ethnic minorities, social integration and crime. *European Journal of Criminal Policy and Research*, 9, 5–29.
- Kohlberg, L., & Kramer, A. (1969). Continuities and discontinuities in childhood and adult moral development. *Human Development*, 12(2), 93–120.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59(10), 877–883. <https://doi.org/10.1001/archpsyc.59.10.877>
- Lambie, I., & Randell, I. (2013). The impact of incarceration on juvenile offenders. *Clinical Psychology Review*, 33(3), 448–459. <https://doi.org/10.1016/j.cpr.2013.01.007>
- Leonardi, R. J., Buchanan-Smith, H. M., McIvor, G., & Vick, S. J. (2017). “You think you’re helping them, but they’re helping you too”: Experiences of Scottish male young offenders participating in a dog training program. *International Journal of Environmental Research and Public Health*, 14(8), 945–972. <https://doi.org/10.3390/ijerph14080945>
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders*, 4(2), 124–147. <https://doi.org/10.1080/15564880802612573>
- Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., & Carver, D. (2010). *Improving the effectiveness of juvenile justice programs*. [https://rhyclearinghouse.acf.hhs.gov/sites/default/files/docs/19740-Improving\\_the\\_Effectiveness\\_of.pdf](https://rhyclearinghouse.acf.hhs.gov/sites/default/files/docs/19740-Improving_the_Effectiveness_of.pdf)
- May, D. K., Seivert, N. P., Cano, A., Casey, R. J., Johnson, A., & May, D. (2016). Animal-assisted therapy for youth: A systematic methodological critique. *Human-Animal Interaction Bulletin*, 4(1), 1–18.

- Minton, C. A., Perez, P. R., & Miller, K. (2015). Voices from behind prison walls: The impact of training service dogs on women in prison. *Society and Animals, 23*(5), 484–501. <https://doi.org/10.1163/15685306-12341379>
- Montori, V. M., & Guyatt, G. H. (2001). Intention-to-treat principle. *Canadian Medical Association Journal, 165*(10), 1339–1341.
- Nimer, J., & Lundahl, B. (2007). Animal-assisted therapy: A meta-analysis. *Anthrozoös, 20*(3), 225–238. <https://doi.org/10.2752/089279307X224773>
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology, 79*(1), 6–21. <https://doi.org/10.1037/a0022200>
- Roest, J. J., Van der Helm, G. H. P., & Stams, G. J. J. M. (2016). The relation between therapeutic alliance and treatment motivation in residential youth care: A cross-lagged panel analysis. *Child and Adolescent Social Work Journal, 33*(5), 455–468. <https://doi.org/10.1007/s10560-016-0438-4>
- Rosenberg, M. (1979). *Conceiving the self*. Basic Books.
- Schenk, G., Duindam, H. M., Creemers, H. E., Hoeve, M., Stams, G. J. J. M., & Asscher, J. J. (2018). The effectiveness of dutch cell dogs in correctional facilities in the Netherlands: A study protocol of a quasi-experimental trial. *BMC Psychiatry, 18*(1). <https://doi.org/10.1186/s12888-018-1797-5>
- Schuengel, C., & Zegers, M. (2003). *Psychological Availability and Reliance on Adult Manual*. Amsterdam.
- Seivert, N. P., Cano, A., Casey, R. J., Johnson, A., & May, D. K. (2016). Animal assisted therapy for incarcerated youth: A randomized controlled trial. *Applied Developmental Science, 22*(2), 139–153. <https://doi.org/10.1080/10888691.2016.1234935>
- Shearer, R. A., Myers, L. B., & Ogan, G. D. (2001). Treatment resistance and ethnicity among female offenders in substance abuse treatment programs. *Prison Journal, 81*(1), 55–72. <https://doi.org/10.1177/0032885501081001005>
- Silver, E., Felson, R. B., & Vaneseltine, M. (2008). The relationship between mental health problems and violence among criminal offenders. *Criminal Justice and Behavior, 35*(4), 405–426. <https://doi.org/10.1177/0093854807312851>
- Smith, H. P., & Smith, H. (2019). A qualitative assessment of a dog program for youth offenders in an adult prison. *Public Health Nursing, 36*(4), 507–513. <https://doi.org/10.1111/phn.12622>
- Spruit, A., van der Put, C., Gubbels, J., & Bindels, A. (2017). Age differences in the severity, impact and relative importance of dynamic risk factors for recidivism. *Journal of Criminal Justice, 50*(December 2016), 69–77. <https://doi.org/10.1016/j.jcrimjus.2017.04.006>
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 173–212. <https://doi.org/10.4324/9781315175775>
- Usher, A. M., & Stewart, L. A. (2014). Effectiveness of correctional programs with ethnically diverse offenders: A meta-analytic study. *International Journal of Offender Therapy and Comparative Criminology, 58*(2), 209–230. <https://doi.org/10.1177/0306624X12469507>
- Van Der Helm, G. H. P., Wissink, I. B., De Jongh, T., & Stams, G. J. J. M. (2012). Measuring treatment motivation in secure juvenile facilities. *International Journal of Offender Therapy and Comparative Criminology, 57*(8), 996–1008. <https://doi.org/10.1177/0306624X12443798>
- Van Der Helm, P., Beunk, L., Stams, G. J., & van der Laan, P. (2014). The relationship between detention length, living group climate, coping, and treatment motivation among juvenile delinquents in a youth correctional facility. *The Prison Journal, 94*(2), 260–275. <https://doi.org/10.1177/0032885514524884>

- Van Langen, M., Stams, G., & Van Vugt, E. (2009). *Dutch validation of the basic empathy scale*. Unpublished manuscript.
- van Wormer, J., Kigerl, A., & Hamilton, Z. (2017). Digging deeper: Exploring the value of prison-based dog handler programs. *Prison Journal*, 97(4), 520–538. <https://doi.org/10.1177/0032885517712481>
- Wilson, D. B. (2016). Correctional programs. In D. Weisburd, D. P. Farrington, & C. Gill (Eds.), *What works in crime prevention and rehabilitation: Lessons from systematic reviews* (pp. 193–218). Springer. <https://doi.org/10.1007/978-1-4939-3477-5>
- Wilson, S. J., Lipsey, M. W., & Soydan, H. (2003). Are mainstream programs for juvenile delinquency less effective with minority youth than majority youth? A meta-analysis of outcomes research. *Research on Social Work Practice*, 13(1), 3–26. <https://doi.org/10.1177/1049731502238754>
- Yamaguchi, Y., Ueno, M., Maruo, K., & Goshō, M. (2020). Multiple imputation for longitudinal data in the presence of heteroscedasticity between treatment groups. *Journal of Biopharmaceutical Statistics*, 30(1), 178–196. <https://doi.org/10.1080/10543406.2019.1632878>