Associations Between Neurocognitive Characteristics, Treatment Outcome, and Dropout Among Aggressive Forensic Psychiatric Outpatients

Danique Smeijers[1,2], Erik Bulten[3], Jan Buitelaar[2,4,5], and Robbert-Jan Verkes[1,2,3]

Abstract

Aggression Replacement Training (ART) is widely used to reduce aggression and is considered to be effective although there are also inconsistent results. Studies investigating the effectiveness of ART do not focus on neurocognitive characteristics. Focusing on these aspects would result in enhanced understanding of underlying mechanisms of ART. The current open uncontrolled treatment study assessed whether neurocognitive characteristics were associated with change in aggression during the social skills and anger control modules of ART among forensic psychiatric outpatients. Furthermore, differences between treatment dropouts and completers and change in these characteristics during ART were examined. A reduction of trait aggression, cognitive distortions, and social anxiety was observed. Neurocognitive characteristics were not associated with change in aggression, could not distinguish treatment completers from dropouts, and did not change after ART. It is suggested that new paradigms should be developed which take into account the social context in which these impairments appear.

Keywords

aggressive behavior, treatment, responsivity, neurocognitive, forensic psychiatric patients

[1]Department of Psychiatry, Radboud University Medical Centre Nijmegen, The Netherlands
[2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
[3]Forensic Psychiatric Centre Pompestichting, Nijmegen, The Netherlands
[4]Karakter Child and Adolescent Psychiatry University Center, Nijmegen, The Netherlands
[5]Department of Cognitive Neuroscience, Radboud University Medical Centre Nijmegen, The Netherlands

Corresponding Author:
Danique Smeijers, Pompestichting, Postbus 31435, Nijmegen 6503 CK, The Netherlands
Email: daniquesmeijers@gmail.com

Disproportionate aggressive behavior, defined as aggression disproportionate to the provocation (Siever, 2008), is one of the most important reasons for referral to forensic psychiatric institutions (Smith & Humphreys, 1997). Aggression Replacement Training (ART; Glick & Goldstein, 1987; Goldstein, Glick, & Gibbs, 1998) is frequently used in children and adolescents as well as in adults to reduce aggression (Brännström, Kaunitz, Andershed, South, & Smedslund, 2016). The ART consists of three modules: (a) *social skills training*, which focuses on responding in a prosocial way to difficult situations instead of using aggression; (b) *anger control training*, which learns to gain more control over aggressive thoughts and aggressive impulses; and (c) *moral reasoning training*, where patients learn to recognize certain cognitive distortions relating to aggression by themselves and think in less egocentric way by means of group discussions. The modules are given in three weekly sessions over 10 weeks.

In general, ART is considered to be effective in decreasing aggressive behavior among forensic psychiatric outpatients (FPOs), even though there are also inconsistent results (Brännström et al., 2016). Furthermore, studies indicate that a high disposition to act aggressively before treatment is associated with more change during treatment (Hornsveld, van, Dam-Baggen, Leenaars, & Jonkers, 2004; Hornsveld, Kraaimaat, Muris, Zwets, & Kanters, 2014; Hornsveld, Nijman, Hollin, & Kraaimaat, 2008; Hornsveld, 2005; Hornsveld, Nijman, & Kraaimaat, 2008; Smeijers, Bulten, Buitelaar, & Verkes, 2017), whereas presence of cognitive distortions is thought to be related to a less rapid reduction of aggression (Smeijers et al., 2017). In addition, patients who dropped out of ART are characterized by more weekly alcohol and cannabis use, higher levels of psychopathic traits, and proactive aggression (Hornsveld et al., 2014; Hornsveld, Nijman, & Kraaimaat, 2008; Smeijers et al., 2018).

Studies investigating the effectiveness of the ART, however, often focus on behavioral and/or personality characteristics. However, neurocognitive characteristics are also found to be associated with aggressive behavior. As suggested by the Social Information Processing (SIP) model, individuals with severe aggressive behavior have abnormal response repertoires in social situations due to deficient information processing (Crick & Dodge, 1996). In specific two steps of the SIP model are thought to be impaired: (a) *interpretation*: aggressive individuals tend to attribute malicious intent toward other's intention, and (b) *response decision*: aggressive individuals evaluate aggressive acts in such a way that it is likely to lead to the enactment of aggressive behavior (Crick & Dodge, 1996). Both of these patterns do not occur consciously such as the aforementioned cognitive distortions which are defined as inaccurate attitudes, thought, or beliefs regarding own or others' behavior. We also can refer to these automatic and implicit deficient patterns as cognitive biases. In specific, two forms of cognitive biases are frequently associated with aggressive behavior: attentional biases and hostile attribution/interpretation biases. An attentional bias refers to an a priori tendency to focus selectively on threatening information and is often measured by use of the Emotional Stroop Task (EST). An attentional bias toward aggressive- or violence-related words has been associated with higher levels of aggressive behavior in student populations as well as offenders and FPOs (e.g., Brugman et al., 2014; Chan,

Raine, & Lee, 2010; Domes, Mense, Vohs, & Habermeyer, 2013; Smith & Waterman, 2003, 2004). Even though a few studies examined this bias by use of the Stroop Task in clinical populations, it is not often used in clinical settings in association with treatment.

Hostile attribution/interpretation biases, however, refer to attributing hostile intent to others' actions and to interpret emotional facial expressions as hostile. Studies regarding these specific biases revealed that hostility biases are associated with higher levels of aggressive behavior in children as well as adults and in student populations as well as offenders and FPOs (e.g., Chen, Coccaro, & Jacobson, 2012; Dodge, 2006; Schonenberg & Jusyte, 2014; Smeijers et al., 2017). Both biases could be important causes as well as maintaining factors of aggressive behavior. In addition, possibly related to these biases is social approach and avoidance behavior. For instance, it is thought that angry facial expressions operate as a signal of threat by transferring the expresser's aggression (Blair, 2003). In general, this results in avoidance behavior as one wants to avoid potential danger. However, psychopaths appear to lack this automatic avoidance tendency, as measured by use of the Approach–Avoidance Task (AAT; von Borries et al., 2012). Furthermore, it is thought that anger is associated with approach motivation (Carver & Harmon-Jones, 2009). It is possible that one of the aforementioned biases, and aggressive behavior in itself, is related to this social approach and avoidance behavior.

Another neurocognitive characteristic often associated with aggressive and antisocial behavior is impulsivity which is described as the inability to withhold a response or thought, preference for immediate reward, acting without forethought, sensation-seeking, and a tendency to engage in risky behavior (for review, see Bari & Robbins, 2013). Computerized tasks often used to measure impulsivity and response inhibition are the Go/No Go paradigm, Stop Signal Task (SST), and the Continuous Performance Task (CPT). The latter is also used to measure sustained attention. Studies using these tasks have shown that higher levels of aggressive behavior is associated with impulsivity and impaired response inhibition, expressed in shorter reaction times or more errors, in student samples as well as in FPOs (e.g., Harmon-Jones, Barratt, & Wigg, 1997; Ogilvie, Stewart, Chan, & Shum, 2011; Pawliczek et al., 2013; Tonnaer, Cima, & Arntz, 2016; Vigil-Colet et al., 2004).

To date, no previous studies investigated neurocognitive characteristics in relation to aggression treatment. This, however, is of great importance as it contributes to our understanding of the working mechanisms of aggression treatment. Furthermore, it may provide more understanding of treatment dropout and treatment responsivity. This is in line with the Risk–Need–Responsivity (RNR) model for offender rehabilitation which postulates that offenders at higher risk for reoffending will benefit most from more intensive treatment, that only those factors associated with reductions in recidivism should be targeted during treatment, and that interventions should be matched to offender characteristics (Andrews et al., 1990; Polaschek, 2012). This can result in better interference related to these deficits to develop personalized treatment. In the long run, this might result in enhanced treatment adherence and eventually a more successful reduction of aggressive behavior.

The current open uncontrolled treatment study was in continuation of Smeijers et al. (2017) and was now aimed (a) at examining whether neurocognitive characteristics at baseline were associated with the change in aggression during ART among FPOs with severe aggressive behavior; (b) to explore whether treatment dropouts differed from treatment completers in these underlying neurocognitive characteristics; and (c) to examine which characteristics, other than aggression, changed during treatment.

Method

Participants

In the period from January 1, 2012, to June 15, 2015, 963 FPOs were referred to "Kairos," the outpatient unit of Forensic Psychiatric Centre the Pompestichting in Nijmegen, The Netherlands, because of aggression regulation problems. Inclusion to the study required to meet each of the following criteria: (a) a diagnosis of an antisocial, borderline, and/or narcissistic personality disorder and/or the intermittent explosive disorder (IED), and (b) a total score of 5 points or higher on the Social Dysfunction and Aggression Scale (SDAS; Wistedt et al., 1990). In addition, FPOs were excluded if there was a current severe addiction, current major depression, or lifetime bipolar disorder or psychosis. They were excluded because in such a condition, a proper treatment of these disorders will be a priority, and furthermore, these conditions will seriously limit the responsivity of patients to ART. In this study, 213 male and female FPOs were eligible and willing to participate. An overview of reasons for exclusion is provided in Table 1.

Of the 213 FPOs, 44 were excluded because of no show at the screening appointment ($n = 22$), no show at the baseline assessment start of treatment ($n = 6$), not willing to participate anymore ($n = 4$), no current aggressive behavior ($n = 10$; only aggression in the past), or current major depression ($n = 2$). Eventually, 169 FPOs participated in the present study, of which 125 performed the halfway measurement and 115 performed the end of treatment measurement as well. Of the 169 participating FPOs, 58 dropped out of the current study because of no show during intervention ($n = 47$), not willing to participate anymore ($n = 2$), reference to other type of treatment (for addiction or for autism; $n = 7$), and death ($n = 2$). Unfortunately, no follow-up information regarding this drop-out group is available. Furthermore, admission to Kairos occurs on either obligatory (e.g., when sentenced by a judge) or voluntary basis (based on reference by general practitioner). In the current study, 124 patients were referred voluntarily and 45 obligatory.

Demographic information is provided in Table 2. All participants signed a consent form after receiving information about the study and obtained a monetary compensation (€30). The current study was approved by the regional Ethics Committee, Commissie Mensgebonden Onderzoek (CMO) Regio Arnhem-Nijmegen, The Netherlands. The initial design included a waiting list control group. However, this design did not get ethical approval of the Ethics Committee. Due to the vulnerability of this population and their

Table 1. Reason for Exclusion.

	n
Total	750
Reason	
Negative decision by therapist due to severity of psychopathology	12
Dropout after intake/not suitable for treatment	102
Refused to participate	324
Exclusion criteria	
Current major depression	52
Lifetime psychosis	25
Current severe alcohol/drug dependency	56
Insufficient understanding of Dutch language	12
No current aggressive behavior (only past)	167

Table 2. Demographic Information.

	M/n
Age	M = 35.79 (SD = 10.94)
Male	n = 159
Female	n = 10
IQ[a]	M = 87.83 (SD = 11.91)
Alcohol use, unit/week	n = 112; M = 7.99 (SD = 15.34)
Cannabis use, joint/week	n = 53; M = 4.31 (SD = 13.54)
Diagnosis	
Antisocial personality disorder	n = 68
Borderline personality disorder	n = 35
Narcissistic personality disorder	n = 6
Intermittent explosive disorder	n = 145
ADHD	n = 46
History of depressive disorder	n = 110

Note. ADHD = attention deficit hyperactivity disorder.
[a]As measured by using the Dutch Adult Reading Test (Schmand, Bakker, Saan, & Louman, 1991).

critical need of psychological care, the committee reasoned that these patients should receive immediate care and considered a waiting list control group as not ethical. The current study, therefore, was an uncontrolled open trial study. This is one of the difficulties of conducting research in clinical settings and that it was not possible to deliver the ART components as they were originally intended. Due to the nature of the study, no causal conclusions can be drawn. Nevertheless, cross-sectional studies are thought to be more generalizable to the general population of the population of interest; more suitable to identify risk factors; and, therefore, of significant clinical relevance (Besen & Gan, 2014).

Materials

Questionnaires. The SDAS (Wistedt et al., 1990) is an observer scale that measures the severity of actual aggressive behavior. It consists of nine items measuring outward aggression and two items measuring inward aggression. Items have to be scored on a 4-point Likert-type scale, with 0 = *not present* and 4 = *severely to extremely present as extremes*. The SDAS has adequate observer reliability (Wistedt et al., 1990). In the current study, due to lack of observers, the SDAS was used as self-report. Participants had to rate their aggressive behavior over a period of 3 months. The SDAS as self-report demonstrated acceptable test–retest stability: intraclass correlation coefficient (ICC) baseline – halfway measurement = .686, $p <$ .01; halfway – end of treatment measurement = .763, $p < $ .01; and baseline – end of treatment measurement = .651, $p < $ .01, and acceptable internal consistency: Cronbach's α baseline = .76, halfway measurement = .82, and end of treatment measurement = .82 (Smeijers et al., 2017).

The Aggression Questionnaire (AQ; Buss & Perry, 1992) is a self-report questionnaire to assess an overall trait level of aggression. It consists of 29 items which are divided into four subscales: physical aggression, verbal aggression, anger, and hostility. The items are scored on a 5-point Likert-type scale (1 = *extremely unlike me* to 5 = *extremely like me*). The Dutch translation has adequate psychometric properties (Morren & Meesters, 2002). In the present study, the internal consistency has also proven to be good (Cronbach's α = .91 for baseline measurement and .93 for end of treatment measurement). The AQ was administered at baseline and at end-of-treatment measurement.

The How I Think (HIT) questionnaire (Barriga & Gibbs, 1996) is a 54-item self-report questionnaire to assess self-serving cognitive distortions at baseline and end-of-treatment measurement. The items are divided into four cognitive distortion subscales (self-centered, blaming others, minimizing/labeling, and assuming the worst) and four behavioral referent categories (physical aggression, opposition-defiance, lying, and stealing). Items have to be answered on a 6-point Likert-type scale. In the current study, the internal consistency has also proven to be good (Cronbach's α = .93 for baseline measurement and .94 for end-of-treatment measurement). The Dutch translation has proven reliability and validity (Nas, Brugman, & Koops, 2008).

The Inventory of Interpersonal Situations (IIS; Dam-Baggen & Kraaimaat, 2000) is a self-report measure that assesses two aspects of social anxiety: discomfort experiences in interpersonal situations and frequency of assertive social responses in those situations. The IIS contains two scales with 35 items each. The items are scored on a 5-point Likert-type scale (1 = *not at all* to 5 = *a lot*). The Dutch version of the IIS has demonstrated adequate validity and reliability (Dam-Baggen & Kraaimaat, 2000). In the current study, only the subscale regarding discomfort experiences/social anxiety was of interest. In the current study, the internal consistency of this subscale has also proven to be good (Cronbach's α = .96 for baseline measurement and .96 for end-of-treatment measurement). The IIS was administered at baseline and end-of-treatment measurement.

Paradigms. The Point Subtraction Aggression Paradigm (PSAP) measures aggression in response to provocation. The goal of the task was to earn as much money as possible. Participants were instructed to play an online game against a confederate. Two response options were available: (a) pressing the "A" button 40 times, the participants earns US$2, and (b) pressing the "B" button 10 times, US$2 will be deducted from the participants' fictitious opponent. The fictitious opponent also subtracts points from the participant on a predetermined basis. The number of times the point subtraction button is pressed is the dependent measure of aggression. The PSAP has proven psychometric properties (Cherek, Tcheremissine, Lane, & Nelson, 2006). The PSAP was administered at baseline and end-of-treatment measurement.

The Hostile Interpretation Bias Task (HIBT; Smeijers, Rinck, Bulten, van den Heuvel, & Verkes, 2017) was used to assess a hostile interpretation bias (HIB) at baseline and end-of-treatment measurement. Photographs of faces with emotional affect (angry, fear, disgust, and happy) of four male and four female models were selected from the Radboud Faces Database (Langner et al., 2010). Each affective picture was morphed (using WinMorph 3.01) 5 times with the neutral image of the same individual, creating 20%, 40%, 60%, 80%, and 100% emotion intensity, respectively. The neutral expression was in all models displayed with mouth closed, whereas the emotional pictures where displayed with mouth open. This difference in mouth opening resulted in pictures showing ambiguous expressions.

The task consisted of a practice block and two experimental blocks. The practice block consisted of 16 trials (8 models × 2 emotions). Only pictures with happy and angry affect and of 100% intensity were used to familiarize participants with the task. Each experimental block consisted of 168 trials (8 models × 4 emotions × 5 intensity levels + 8 neutral images). The order of the pictures was pseudo-randomized and equal in both blocks. Participants were instructed to indicate whether the picture looked hostile or not. In case they thought they saw a hostile picture, they were asked to press the Z-key, otherwise the M-key (on a QWERTY keyboard). They had to respond as quickly as possible. The picture, size 8.5 cm × 10.5 cm, was presented for 4 s, in the center of the computer screen, against a black background. The pictures remained on the screen until a response was given or until 4 s had passed. After a pretrial pause of 1 s, a new picture was displayed immediately. Labels were displayed in the left (Yes, hostile) and right (No, not hostile) bottom corner of the screen in white Arial font, size 30. Responses given by pressing the Z-key, indicating that the participant saw a hostile picture, were defined as "hostile" responses. If a response was not given within 4 s, the words "Too late" appeared on the screen in red. A HIB was defined as the percentage of "hostile" responses to the emotional pictures. The hostile responses were dummy coded (0 = no, not hostile, 1 = yes, hostile), and the mean was calculated which then immediately revealed the percentage of the pictures that were interpreted as hostile. Trials without a response (due to late responding) were not taken into account.

The SST (Logan, 1994; Logan, Schachar, & Tannock, 1997) was used to measure motor impulsivity and response inhibition. This task requires participants to make quick key responses to visually presented Go signals and to inhibit any response when an auditory stop signal is suddenly presented. The Go signals were arrows pointing left

or right which were presented for 750 ms in the center of the computer screen after the presentation of a fixation cross for 400 ms. The variable time interval between the end of a Go signal and the start of the fixation cross was 1,000 to 1,200 ms. Participants were instructed to respond as quickly and accurately as possible on all trials by pressing the corresponding arrow on the keyboard but to withhold any response on "stop trials." The stop signal consisted of an auditory cue, that is, a high-pitched sound of 1000 Hz. The SST used in the current study consisted of four blocks which all consisted of 60 trials. Blocks 1 and 4 consisted of 48 Go trials and 12 stop trials. Blocks 2 and 3 consisted of 12 Go trials and 48 stop trials. Only trials with reaction times >150 and <1,500 ms were of interest. No responses on Go trials were defined as omission errors, responses to No Go trials were defined as commission errors. Mean omission and commission errors were used as dependent variables.

The CPT was used to measure sustained attention and response inhibition. Letters were sequentially presented on the computer screen and participants had to press the spacebar as soon as the letter-combination "AX" appeared (Go trial). The CPT consisted of 20 practice trials and 400 experimental trials, of which 15% consisted of Go trials. Each letter was presented for 200 ms. The interstimulus interval was 900 ms. Only trials with reaction times >150 and <1,500 ms were of interest. No responses on Go trials were defined as omission errors, and responses to No Go trials were defined as commission errors. Mean omission and commission errors were used as dependent variables. The CPT was found to be significantly correlated with other psychometric measures of inattention, ratings of inattention, impulsivity, and hyperactivity (Klee & Garfinkel, 1983).

The EST was used to measure attentional bias. Two versions of the EST were used: (a) aggression-related words versus neutral words and (b) anxiety-related words versus neutral words. All words were presented in color (red, yellow, green, and blue) in the center of the computer screen and were presented for 100 ms with an interstimulus interval of 1,500 ms. Both versions consisted of 132 trials, of which 50% consisted of neutral words. Only trials with reaction times >150 and <1,500 ms were of interest. Reaction times to emotionally laden words are longer than to emotionally neutral words, as the emotional content appears to capture attention. Slowed responses are due to the emotional relevance of the word for the individual (interference (bias) effect). The EST bias score was used as dependent variable. Bias scores were calculated by subtracting the mean reaction time of the neutral words of the mean reaction time of the anxiety- or aggression-related words.

The AAT (Rinck & Becker, 2007) was used to measure approach and avoidance tendencies regarding emotional faces at baseline and end-of-treatment measurement. Black–white photographs (sized 8.4-13.5 cm) of facial expressions of eight actors (four men and four women) were selected from Ekman. Participants were instructed to evaluate the facial expressions (i.e., happy or angry) and to respond as fast and accurate as possible to the stimuli by either pulling or pushing a joystick. When the joystick is pulled, the stimulus grows in size, when it is pushed, it shrinks. Participants receive alternately an affect-congruent or an affect-incongruent instruction. The affect-congruent instruction indicates pulling the joystick for happy faces and pushing the

joystick for angry faces. In the affect-incongruent condition, the required response to the facial expression is reversed.

The AAT consisted of six blocks (push happy–pull neutral, pull happy–push neutral, push neutral–pull angry, pull neutral–push angry, pull happy–push angry, and push happy–pull angry) of, in total, 288 trials. Each block was preceded by a practice block of 16 trials. During this practice block, participants received feedback in case they made an error: The picture did not disappear. The trials were semirandomized: No more than three of the same stimulus response combination was presented successively. Each trial was self-paced: Participants had to press the fire button while the joystick was in the resting (upward) position. Only trials with reaction times >150 and <1,500 ms were of interest. AAT bias scores were calculated for angry, happy, and neutral faces. The individual mean reaction time for pull movements were subtracted from the individual mean reaction time for push movements. Negative scores indicate stronger avoidance, whereas positive scores reflect stronger approach tendencies.

Intervention

All FPOs were referred to the ART. The ART as offered by "Kairos" consisted of solely two of the original modules: (a) *social skills training* and (b) *anger control training*. Moreover, a slightly adapted form of ART was offered for perpetrators of intimate partner violence which is called "Stop Domestic Violence" (SDV). The SDV consisted of identical modules as the ART; in addition, the partners of the FPOs were involved during the intervention (SDV; $n = 58$). Both the ART and SDV consisted of two, 90-min, weekly sessions during 12 weeks and occurred either in groups ($n = 116$) or individually ($n = 46$). Allocation to SDV was determined by a clinician who judged whether the aggression displayed by the FPO only occurred in domestic settings or whether it could be considered as generalized aggressive behavior. Also, allocation to group or individual treatment was determined by the clinician and was dependent of a patients learning style. The first 10 weeks consisted of the social skills and anger control training. Week 11 consisted of a session to integrate all that was learned in the previous weeks. Finally, Week 12 consisted of an evaluation session. Indication for ART was determined by a multidisciplinary team. ART therapists were all formerly trained in applying the ART and, in addition, made use of a detailed intervention manual and participated in intervision.

Procedure

Clinicians at "Kairos" asked FPOs who were referred to ART (group or individual) whether they agreed to be contacted about the study. When they agreed, FPOs were contacted by the researcher. All FPOs received treatment as indicated, whether they participated in the study or not.

After receiving information about the nature of the study, a consent form was signed. The FPOs were screened by trained clinicians with the Research Criteria set for Intermittent Explosive Disorder (IED-IR; Coccaro, Kavoussi, Berman, & Lish,

1998), the Structured Clinical Interview for *DSM-IV* Axis II personality disorders (SCID-II; Weertman, Arntz, & Kerkhofs, 2000), and the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998; Van Vliet & De Beurs, 2007) regarding the aforementioned inclusion and exclusion criteria.

Once FPOs were suitable for participation they proceeded with the baseline measurement which was completed before the start of their treatment. Halfway, after 6 weeks, the level of aggressive behavior was determined by use of the SDAS. An end-of-treatment measurement took place after 12 weeks. The FPOs were instructed not to use alcohol or drugs 24 hr prior to any of the assessments.

Statistical Analysis

First, it is was examined whether FPOs who received ART as usual or ART for domestic violence perpetrators and FPOs who received group or individual ART could be considered as one sample in subsequent analyses. Therefore, a MANOVA was conducted to examine whether there were differences on the SDAS, AQ, HIT, IIS, HIBT, CPT, SST, EST, AAT, and PSAP between aforementioned groups. Bonferroni correction was used to control for multiple comparisons.

Subsequently, in accordance with Smeijers et al. (2017), a linear mixed model (SPSS, Version 24) was used to examine whether implicit baseline characteristics moderated the treatment response. One advantage of this analysis is that it is possible to include individuals with incomplete data (Heck, Thomas, & Tabata, 2013). As the aggression assessment halfway and end of treatment were not completed for all participants, this method was favored. The basic model was a repeated-measures design with aggressive behavior as measured with the SDAS as dependent variable and Time of measurement (baseline, halfway, and end of treatment) as within-subjects factor. Repeated covariance type was set at diagonal. With respect to Time, the slope was set as a fixed effect and the intercept as a random effect. These random effects were defined to assess variation in the dependent variable because variation among individuals, regarding change in aggression over time, was assumed (Bolker et al., 2009; West, Welch, & Galecki, 2014). The covariance type for the random effects was set at unstructured. As our previous study already revealed that aggressive behavior reduced over the course of treatment, the main effects of HIBT, CPT, EST aggression, EST anxiety, SST, and PSAP and two-way interactions with Time were directly added to the basic model.

Subsequently, to investigate the underlying differences between FPOs who completed treatment versus FPOs who dropped out, a MANOVA regarding baseline measurements was conducted. Bonferroni correction was used to control for multiple comparisons. Finally, paired *t* tests were conducted to examine whether end-of-treatment measurements changed as compared with baseline.

Results

Table 3 displays the means on the SDAS, HIBT, CPT, EST aggression, EST anxiety, SST, and PSAP. Means are presented for the total sample as well as separately for the

Table 3. Descriptives (*M, SD*) of the Total Sample, Completers Versus the Dropouts, Baseline, and End of Treatment.

	Total sample: Baseline measurement (n = 169)	Completers: Baseline measurement (n = 111)	Dropouts: Baseline measurement (n = 58)	Completers: End of treatment (n = 111)
SDAS baseline	M = 15.14 (SD = 6.49)	M = 15.14 (SD = 6.19)	M = 15.14 (SD = 7.09)	—
SDAS halfway	—	M = 10.79 (SD = 6.67)	—	—
SDAS end of treatment	—	—	—	M = 8.48 (SD = 6.35)
AQ[a]	M = 93.02 (SD = 18.69)	M = 91.99 (SD = 18.41)	M = 95.00 (SD = 19.21)	M = 85.13 (SD = 20.53)
HIT[a]	M = 2.46 (SD = 0.76)	M = 2.41 (SD = 0.73)	M = 2.56 (SD = 0.81)	M = 1.83 (SD = 0.68)
IIS[a]	M = 77.01 (SD = 28.00)	M = 78.42 (SD = 28.39)	M = 74.29 (SD = 27.26)	M = 71.62 (SD = 26.54)
CPT omission	M = 5.37 (SD = 7.69)	M = 5.37 (SD = 8.29)	M = 5.38 (SD = 6.51)	—
CPT commission	M = 8.52 (SD = 9.35)	M = 7.83 (SD = 9.24)	M = 9.64 (SD = 9.52)	—
SST omission	M = 2.08 (SD = 5.95)	M = 2.28 (SD = 6.98)	M = 1.66 (SD = 2.89)	—
SST commission	M = 6.80 (SD = 13.89)	M = 7.58 (SD = 16.49)	M = 5.21 (SD = 5.34)	—
PSAP	M = 232.08 (SD = 418.68)	M = 187.91 (SD = 284.03)	M = 321.24 (SD = 599.21)	M = 241.08 (SD = 371.65)
EST aggression	M = 124.62 (SD = 11,109.00)	M = 1,307.02 (SD = 10,714.54)	M = −402.47 (SD = 11,615.12)	—
EST anxiety	M = −428.79 (SD = 4,172.83)	M = −558.22 (SD = 4,229.55)	M = −15.471 (SD = 4,077.78)	—
AAT angry	M = −8.07 (SD = 97.64)	M = −2.67 (SD = 101.37)	M = −19.06 (SD = 89.47)	M = −7.95 (SD = 106.33)
AAT happy	M = 17.36 (SD = 111.02)	M = 17.24 (SD = 106.21)	M = 17.61 (SD = 121.19)	M = 18.61 (SD = 107.21)
AAT neutral	M = −9.63 (SD = 102.86)	M = 29.31 (SD = 111.78)	M = 11.37 (SD = 135.22)	M = 10.86 (SD = 95.07)
HIB angry	M = 56.03 (SD = 19.03)	M = 54.56 (SD = 19.78)	M = 58.94 (SD = 17.28)	M = 57.26 (SD = 17.52)
HIB happy	M = 11.54 (SD = 16.43)	M = 10.03 (SD = 14.81)	M = 14.54 (SD = 19.06)	M = 9.34 (SD = 13.02)
HIB disgust	M = 49.57 (SD = 22.28)	M = 47.34 (SD = 22.32)	M = 33.93 (SD = 28.02)	M = 50.09 (SD = 22.50)
HIB fear	M = 27.74 (SD = 25.92)	M = 24.62 (SD = 24.35)	M = 53.98 (SD = 21.78)	M = 24.32 (SD = 23.50)

Note. The mean total score on the Social Dysfunction and Aggression Scale (SDAS), the Aggression Questionnaire (AQ), the How I Think (HIT) questionnaire, and the Inventory of Interpersonal Situations (IIS); the mean total omission and commission errors of the Stop Signal Task (SST) and Continuous Performance Task (CPT); the mean total amount of "B" responses on the Point Subtraction Aggression Paradigm (PSAP); and the mean bias score on the aggression and anxiety Emotional Stroop Task (EST) and Approach–Avoidance Task (AAT) are reported. Hostile interpretation bias (HIB) is defined as the percentage of "hostile responses" per emotion.
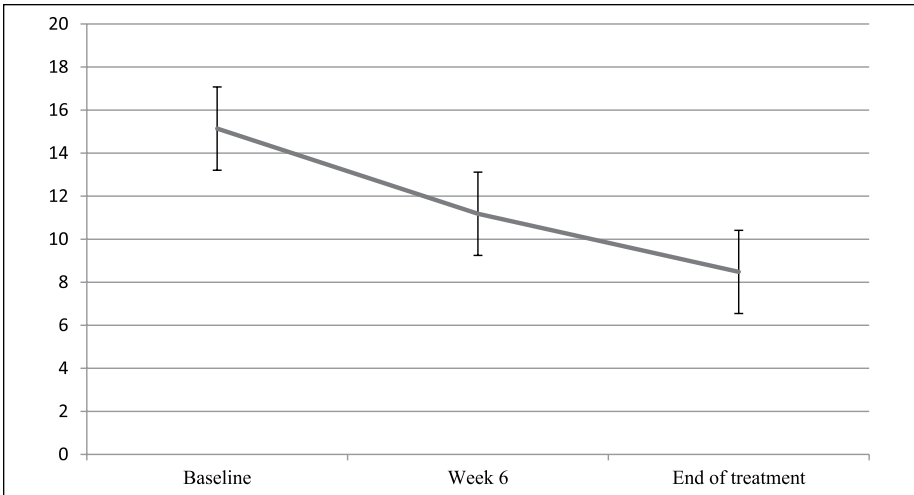[a]Baseline and end-of-treatment measurements significantly differ from each other at least p < .004.

Figure 1. Graphical representation of the change in aggressive behavior according to the self-report of FPOs (SDAS).
Note. Baseline: *n* = 169, *SD* = 6.49; Week 6: *n* = 125, *SD* = 6.69; end of treatment: *n* = 115, *SD* = 6.25. FPOs = forensic psychiatric outpatients; SDAS = Social Dysfunction and Aggression Scale.

treatment dropouts and completers and baseline and end-of-treatment measurements. In supplementary material also, the means on all measures are presented for all subgroups: male, female, individual treatment, group treatment, standard ART, and ART for domestic violence.

Differences Among FPOs

The MANOVA, which was conducted to investigate whether FPOs who received regular ART versus ART for domestic violence perpetrators and FPOs who received group versus individual treatment could be considered as one sample, revealed no significant multivariate effects of ART versus ART for domestic violence and group versus individual treatment emerged, Wilks's Lambda = .784, $F(20, 43) = 0.592$, $p = .897$, $\eta^2 = .216$; Wilks's Lambda = .575, $F(20, 43) = 1.586$, $p = .102$, $\eta^2 = .425$, respectively. Therefore, in subsequent analyses, the FPOs were considered as one sample.

Effect on Aggressive Behavior

The basic model (Time, SDAS) showed a significant main effect of Time, indicating that aggressive behavior significantly reduced during treatment (Smeijers et al., 2017). A graphic representation of the change in aggressive behavior (SDAS) is displayed in Figure 1. As an effect size is not provided by linear mixed models, Cohen's *d* was calculated by dividing the mean difference between pre–post SDAS scores by the pooled standard deviation: $15.19 - 8.48 / \left(\sqrt{\left(\left(6.14^2 + 6.25^2 \right) / 2 \right)} \right) = 1.08$.

This basic model was extended by adding main effects of HIBT, CPT, SST, EST aggression, EST anxiety, AAT, and PSAP. A significant main effect of CPT commission errors and AAT_Neutral was found, suggesting that impulsivity/attention and a stronger avoidance of neutral faces were associated with differences in aggressive behavior. Therefore, in the final model, only the main effect of CPT commission errors and AAT_Neutral and a two-way interaction between CPT and Time and AAT and Time were included to examine whether this characteristics might explain variability in aggression reduction during treatment. The analysis of the final model revealed a significant main effect of Time and CPT. No significant main effect of AAT_Neutral as well as no interaction of Time × CPT and Time × AAT (see Table 4) emerged. The results suggest that impulsivity/difficulties to inhibit behavior and avoidance tendencies of neutral faces were not related to the treatment response.

Dropout

Means and *SD*s for baseline measures of treatment dropouts and completers are presented in Table 3. No significant multivariate effect of group was found, Wilks's Lambda = .876, $F(14, 92) = 0.934$, $p = .526$, $\eta^2 = .124$. The results suggested that treatment completers could not be differentiated from treatment dropouts regarding neurocognitive characteristics.

Baseline Versus End of Treatment

Means and *SD*s for baseline and end-of-treatment measures are presented in Table 3. The results from the paired *t* tests showed that only baseline scores of the AQ, HIT, and IIS decreased at end-of-treatment measurement, $t(113) = 5.001$, $p < .001$; $t(113) = 3.911$, $p < .001$; and $t(113) = 4.057$, $p < .001$, respectively.

Discussion

The present study examined whether neurocognitive characteristics were associated with change in aggression during the social skills and anger control module of the ART among FPOs displaying severe aggressive behavior. Results suggested that response inhibition was associated with change in aggression during treatment; more difficulties to inhibit responses did result in higher levels of aggressive behavior during the course of treatment. However, no interaction effects were found suggesting that response inhibition could not explain variability in aggression reduction during treatment. Our previous study found that treatment dropouts showed higher levels of proactive aggression, psychopathic traits, and more weekly substance use (Smeijers et al., 2017). This knowledge can be used to make an informed decision whether ART is suitable for a specific individual patient and to develop personalized treatment. As suggested by the RNR model for offender rehabilitation, interventions should be matched to offender characteristics (Andrews et al., 1990; Polaschek, 2012). However, in the current study, treatment completers and treatment dropouts could not be further

Table 4. Results of Linear Mixed Model.

Model	Parameter	Estimate	95% CI	t	df	p
Basic model	Intercept	14.91	[13.81, 16.02]	26.82	114.388	<.001
	Time (SDAS)	−3.28	[−3.86, −2.69]	−11.21	114.199	<.001
Model including main effects of baseline characteristics	Intercept	11.39	[7.98, 14.79]	6.65	79.769	<.001
	Time (SDAS)	−3.46	[−4.17, −2.76]	−9.74	87.110	<.001
	EST aggression	0.00	[0.00, 0.00]	1.65	73.672	.103
	EST anxiety	0.00	[−0.00, 0.00]	0.73	73.586	.467
	AAT_Angry	0.00	[−0.00, 0.01]	0.14	72.987	.889
	AAT_Happy	−0.01	[−0.01, 0.00]	−1.26	72.563	.210
	AAT_Neutral	−0.01	[−0.02, 0.00]	−2.50	72.637	.015
	PSAP	0.00	[−0.00, 0.01]	1.38	73.931	.171
	HIBT_Angry	0.04	[−0.04, 0.12]	0.89	72.830	.376
	HIBT_Happy	0.00	[−0.08, 0.09]	0.07	72.565	.944
	HIBT_Fear	−0.05	[−0.12, 0.01]	−1.48	73.500	.141
	HIBT_Disgust	0.02	[−0.07, 0.12]	0.53	73.108	.592
	CPT_Omission	0.05	[−0.10, 0.21]	0.65	72.989	.516
	CPT_Commission	0.24	[0.10, 0.39]	3.38	72.752	<.001
	SST_Omission	−0.05	[−0.40, 0.28]	−0.33	72.857	.741
	SST_Commission	−0.00	[−0.17, 0.17]	−0.02	72.881	.981
Final model	Intercept	13.85	[12.48, 15.23]	20.03	103.148	<.001
	Time (SDAS)	−3.18	[−3.93, −2.42]	−8.35	103.081	<.001
	AAT_Neutral	−0.00	[−0.01, 0.00]	−0.32	102.840	.747
	CPT_Commission	0.21	[0.08, 0.34]	3.26	102.615	<.001
	Time × AAT_Neutral	−0.00	[−0.00, 0.00]	−1.34	102.618	.184
	Time × CPT	−0.00	[−0.07, 0.06]	−0.20	102.586	.839

Note. All results were comparable when excluding women (*n* = 10) from analyses. CI = confidence interval; SDAS = Social Dysfunction and Aggression Scale; EST = Emotional Stroop Task; AAT = Approach–Avoidance Task; PSAP = Point Subtraction Aggression Paradigm; HIBT = Hostile Interpretation Bias Task; CPT = Continuous Performance Task; SST = Stop Signal Task.

distinguished based on underlying neurocognitive characteristics. Finally, we investigated whether end-of-treatment measurements changed as compared with baseline. Besides neurocognitive characteristics, also self-report measures were included. These analyses revealed that the disposition to act aggressively, cognitive distortions, and social anxiety all decreased after the intervention. This suggests that both modules of the ART affect not only aggressive behavior but also other relevant characteristics. No differences were found regarding the neurocognitive measures.

When treating aggressive behavior, one aims not only to increase impulse control but also to affect maintaining factors to reduce the chance of recidivism. Indeed, the current results showed that trait aggression, cognitive distortions, and social anxiety declined after treatment. The level of the neurocognitive deficits, however, remained stable and did not improve after treatment. This suggests that the ART modules did not

affect these cognitive mechanisms. One may consider that ART is primarily a training of behavioral skills and does not specifically target these underlying cognitive mechanisms. It is of importance to investigate whether other forms of interventions, such as cognitive bias modification, are suitable to modify these characteristics, such as hostile interpretation or attentional biases, to treat aggressive behavior successfully and reduce the change of recidivism.

The current negative results might have occurred due to a couple of reasons. First of all, the neurocognitive characteristics might be seen as general responsivity factors. It may be speculated that clinicians are able to respond to these deficits and adjust their interventions accordingly. These deficits might then not be associated with treatment outcome or dropout but determine the implementation of the intervention. Subsequently, neurocognitive characteristics may not be associated with aggression during treatment at all. Yet, in previous research, these paradigms are often used to distinguish patient groups from healthy controls. These studies showed that individuals high in aggressive or antisocial traits display several neurocognitive deficits, such as a HIB toward emotional facial expressions, attentional bias toward aggressive-related words, and aggression on a laboratory measure (e.g., Chan et al., 2010; New et al., 2009; Smeijers et al., 2017). The level of the neurocognitive characteristics measured at baseline, in the current study, was comparable with these previous studies. Then, it may be the case that the paradigms used to measure these characteristics are not sensitive enough to distinguish within a subgroup of FPOs. It will be an important avenue for prospective research to elucidate these impairments in further detail and to develop paradigms which are more sensitive to measure subtle individual differences.

Furthermore, neurocognitive measurement paradigms have several advantages such as that they are less easy to control deliberately; have, for the participant, less transparent purposes; rely less on self-report; and are able to measure automatic attitudes and behavioral dispositions (Schmidt, Banse, & Imhoff, 2015). However, to assess the real deficits, it might be conceivable that these paradigms should be made more real life like by using, for instance, virtual reality. The advantage of virtual reality is that underlying mechanisms can be investigated in controlled experimental designs but with enhanced ecological validity (for review, see Parsons, 2015). These advantages have also been found for use in forensic psychiatry (Benbouriche, Nolet, Trottier, & Renaud, 2014). Besides measurements, there are indications that these techniques may also be deployed in treatment (e.g., Rosenberg, Baughman, & Bailenson, 2013; Seitz, Poyrazli, Harrisson, Flickinger, & Turkson, 2014). Another next step will be to develop norm scores to make it possible to measure individual patients by use of these paradigms. This will enable forensic clinical practice to make use of these paradigms, together with explicit measures, to determine deficits and select interventions to improve these impairments alongside general aggression treatment. In the long run, this will lead to a more successful reduction of the recurrence of aggressive behavior.

The present study and results have a number of limitations which merit further comment. First, the ART examined in the current study only consisted of two modules instead of three. Therefore, the current results are not generalizable to the use of the entire ART. As there exist a lot of variations in the application of the ART (Brännström

et al., 2016), this is not the first study examining a different setup as compared with the original ART. These variations result in difficulties in comparing treatment effectiveness studies. In future research, the focus should maybe be on the modules separately to be able not only to compare different studies and treatment setups but also to understand the working mechanisms of and the added value of each module. Second, as this was an open uncontrolled treatment study and lacked a comparison or control group, the results have to be interpreted with care. Based on the current results, no conclusions can be drawn about the effectiveness of the ART. Third, as a few women participated, the current findings may, therefore, not be generalizable to a female population with aggression regulation problems. Future research is needed to elucidate the effectiveness over a proper control condition and also more research is needed on the working mechanisms of the ART in forensic psychiatric clinical practice.

Notwithstanding the limitations and nature of the current study, the present results contribute to our knowledge of the social skills and anger control modules of the ART by showing that the ART produced not only a decline in aggressive behavior but also in the disposition to act aggressively, social anxiety, and cognitive distortions in FPOs. In addition, the reported negative results give new directions for future research; the current neurocognitive paradigms are suitable to distinguish patients from healthy controls; however, further research is warranted to develop paradigms, which are more sensible to measure subtle individual differences within FPOs regarding these neurocognitive characteristics. In addition, the use of techniques like virtual reality seems to be promising due to their enhanced ecological validity. As individuals high in aggressive behavior, as compared with healthy controls, exhibit neurocognitive deficits, it is of extreme importance to elucidate in further detail whether characteristics at a neurocognitive level are associated with treatment outcome to enhance the understanding of the working mechanisms of the ART in forensic psychiatric clinical practice.

Acknowledgment

Declaration of Conflicting Interests

Funding

Supplemental Material

Supplementary material for this article is available online.

References

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369-404.

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of response control. *Progress in Neurobiology*, *108*, 44-79.

Barriga, A. Q., & Gibbs, J. C. (1996). Measuring cognitive distortion in antisocial youth: Development and preliminary validation of the "How I Think" questionnaire. *Aggressive Behavior*, *22*, 333-343, Laval, France.

Benbouriche, M., Nolet, K., Trottier, D., & Renaud, P. (2014, April). *Virtual reality applications in forensic psychiatry*. Paper presented at the Proceedings of the 2014 Virtual Reality International Conference, Laval, France.

Besen, J., & Gan, S. D. (2014). A critical evaluation of clinical research study designs. *The Journal of Investigative Dermatology*, *134*(3), e18.

Blair, R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*, 561-572.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*, 127-135.

Brännström, L., Kaunitz, C., Andershed, A.-K., South, S., & Smedslund, G. (2016). Aggression Replacement Training (ART) for reducing antisocial behavior in adolescents and adults: A systematic review. *Aggression and Violent Behavior*, *27*, 30-41.

Brugman, S., Lobbestael, J., Arntz, A., Cima, M., Schuhmann, T., Dambacher, F., & Sack, A. T. (2014). Identifying cognitive predictors of reactive and proactive aggression. *Aggressive Behavior*, *41*, 51-64.

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, *63*, 452-459. doi:10.1037/0022-3514.63.3.452

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, *135*, 183-204.

Chan, S.-C., Raine, A., & Lee, T. M. (2010). Attentional bias towards negative affect stimuli and reactive aggression in male batterers. *Psychiatry Research*, *176*, 246-249.

Chen, P., Coccaro, E. F., & Jacobson, K. C. (2012). Hostile attributional bias, negative emotional responding, and aggression in adults: Moderating effects of gender and impulsivity. *Aggressive Behavior*, *38*, 47-63. doi:10.1002/ab.21407

Cherek, D. R., Tcheremissine, O. V., Lane, S. D., & Nelson, R. (2006). Psychopharmacology of human aggression: Laboratory and clinical studies. In R. J. Nelson (Ed.), *Biology of aggression* (pp. 424-446). New York, NY: Oxford University Press.

Coccaro, E. F., Kavoussi, R. J., Berman, M. E., & Lish, J. D. (1998). Intermittent Explosive Disorder–Revised: Development, reliability, and validity of research criteria. *Comprehensive Psychiatry*, *39*, 368-376. doi:10.1016/S0010-440X(98)90050-5

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development*, *67*, 993-1002. doi:10.1111/j.1467-8624.1996. tb01778.x

Dam-Baggen, C. M. J., & Kraaimaat, F. (2000). *Inventarisatielijst Omgaan met Anderen (IOA). Handleiding (2e geheel herziene druk)* [Inventory of Interpersonal Situations (IOA). Manual (2nd completely revised edition)]. Lisse, The Netherlands: Swets Test Services.

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology*, *18*, 791-814. doi:10.1017/S0954579406060391

Domes, G., Mense, J., Vohs, K., & Habermeyer, E. (2013). Offenders with antisocial personality disorder show attentional bias for violence-related stimuli. *Psychiatry Research*, *209*, 78-84. doi:10.1016/j.psychres.2012.11.005

Glick, B., & Goldstein, A. P. (1987). Aggression Replacement Training. *Journal of Counseling & Development*, *65*, 356-362.

Goldstein, A. P., Glick, B., & Gibbs, J. C. (1998). *Aggression Replacement Training: A comprehensive intervention for aggressive youth*. Illinois: Research Press.

Harmon-Jones, E., Barratt, E. S., & Wigg, C. (1997). Impulsiveness, aggression, reading, and the P300 of the event-related potential. *Personality and Individual Differences*, *22*, 439-445.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2013). *Multilevel and longitudinal modeling with IBM SPSS*. New York, NY: Routledge.

Hornsveld, R. H. J. (2005). Evaluation of aggression control therapy for violent forensic psychiatric patients. *Psychology, Crime & Law*, *11*, 403-410. doi:10.1080/10683160500255596

Hornsveld, R. H. J., Kraaimaat, F. W., Muris, P., Zwets, A. J., & Kanters, T. (2014). Aggression Replacement Training for violent young men in a forensic psychiatric outpatient clinic. *Journal of Interpersonal Violence*, *30*, 3174-3191.

Hornsveld, R. H. J., Nijman, H. L., Hollin, C. R., & Kraaimaat, F. W. (2008). Aggression control therapy for violent forensic psychiatric patients: Method and clinical practice. *International Journal of Offender Therapy and Comparative Criminology*, *52*, 222-233. doi:10.1177/03 06624X07303876

Hornsveld, R. H. J., Nijman, H. L. I., & Kraaimaat, F. W. (2008). Aggression control therapy for violent forensic psychiatric patients: First results. *Psychology, Crime & Law*, *14*, 1-18. doi:10.1080/10683160701340569

Hornsveld, R. H. J., van Dam-Baggen, R., Leenaars, E., & Jonkers, P. (2004). Aggression Control Therapy for forensic psychiatric patients: Development and preliminary results. *Tijdschrift voor Psychotherapie*, *30*(1), 13-22.

Klee, S. H., & Garfinkel, B. D. (1983). The computerized continuous performance task: A new measure of inattention. *Journal of abnormal child psychology*, *11*(4), 487-495.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377-1388. doi:10.1080/02699930903485076

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego, CA: Academic Press.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science*, *8*, 60-64.

Morren, M., & Meesters, C. (2002). Validation of the Dutch version of the Aggression Questionnaire in adolescent male offenders. *Aggressive Behavior*, *28*, 87-96. doi:10.1002/ab.90010

Nas, C. N., Brugman, D., & Koops, W. (2008). Measuring self-serving cognitive distortions with the "How I Think" questionnaire. *European Journal of Psychological Assessment*, *24*, 181-189.

New, A. S., Hazlett, E. A., Newmark, R. E., Zhang, J., Triebwasser, J., Meyerson, D., . . . Buchsbaum, M. S. (2009). Laboratory induced aggression: A positron emission tomography study of aggressive individuals with borderline personality disorder. *Biological Psychiatry*, *66*, 1107-1114.

Ogilvie, J. M., Stewart, A. L., Chan, R. C., & Shum, D. H. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. *Criminology*, *49*, 1063-1107.

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience*, *9*, Article 660.

Pawliczek, C. M., Derntl, B., Kellermann, T., Kohn, N., Gur, R. C., & Habel, U. (2013). Inhibitory control and trait aggression: Neural and behavioral insights using the emotional stop signal task. *NeuroImage*, *79*, 264-274.

Polaschek, D. L. (2012). An appraisal of the risk–need–responsivity (RNR) model of offender rehabilitation and its application in correctional treatment. *Legal and Criminological Psychology*, *17*, 1-17.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 105-120.

Rosenberg, R. S., Baughman, S. L., & Bailenson, J. N. (2013). Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PLoS ONE*, *8*(1), e55003.

Schmand, B. A., Bakker, D., Saan, R. J., & Louman, J. (1991). De Nederlandse Leestest voor Volwassenen: een maat voor het premorbide intelligentieniveau [The Dutch Adult Reading Test: A measure of premorbid intelligence]. *Tijdschrift voor Gerontologie en Geriatrie*, *22*, 15-19.

Schmidt, A. F., Banse, R., & Imhoff, R. (2015). Indirect measures in forensic contexts. In F. J. R. van de Vijver, & T. Ortner (Eds.), *Behavior-based assessment in personality, social, and applied psychology*, 173-194. Göttingen, Germany: Hogrefe.

Schonenberg, M., & Jusyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *European Archives of Psychiatry and Clinical Neuroscience*, *264*, 61-69. doi:10.1007/s00406-013-0440-1

Seitz, C. A., Poyrazli, S., Harrisson, M. A., Flickinger, T., & Turkson, M. (2014). Virtual reality exposure therapy for military veterans with posttraumatic stress disorder: A systematic review. *The New School Psychology Bulletin*, *11*(1), 15-29.

Sheehan, D., Lecrubier, Y., Sheehan, K., Amorim, P., Janavs, J., Weiler, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*, 22-33.

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry*, *165*, 429-442. doi:10.1176/appi.ajp.2008.07111774

Smeijers, D., Bulten, B. H., Buitelaar, J. K., & Verkes, R. J. (2017). Treatment Responsivity of Aggressive Forensic Psychiatric Outpatients *International Journal of Offender Therapy and Comparative Criminology*. doi:10.1177/0306624X17747052

Smeijers, D., Rinck, M., Bulten, B. H., van den Heuvel, T., & Verkes, R. J. (2017). Generalized hostile interpretation bias regarding facial expressions: Characteristic of pathological aggressive behavior. *Aggressive Behavior*, *43*, 386-397.

Smith, A. D., & Humphreys, M. (1997). Characteristics of in-patients transferred to a locked ward in a Scottish psychiatric hospital. *Health Bulletin*, *55*(2), 77-82.

Smith, P., & Waterman, M. (2003). Processing bias for aggression words in forensic and nonforensic samples. *Cognition & Emotion*, *17*, 681-701.

Smith, P., & Waterman, M. (2004). Role of experience in processing bias for aggressive words in forensic and non-forensic populations. *Aggressive Behavior*, *30*, 105-122.

Tonnaer, F., Cima, M., & Arntz, A. (2016). Executive (dys)functioning and impulsivity as possible vulnerability factors for aggression in forensic patients. *The Journal of Nervous and Mental Disease*, *204*, 280-286.

Van Vliet, I., & De Beurs, E. (2007). Het Mini Internationaal Neuropsychiatrisch Interview (MINI). Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV en ICD-10-stoornissen [The Mini International Neuropsychiatric Interview (MINI). A short structured diagnostic psychiatric interview for DSM-IV and ICD-10 disorders]. *Tijdschrift voor Psychiatrie*, *49*, 393-397.

Vigil-Colet, A., & Codorniu-Raga, M. J. (2004). Aggression and inhibition deficits, the role of functional and dysfunctional impulsivity. *Personality and Individual Differences*, *37*, 1431-1440.

von Borries, A. K. L., Volman, I., de Bruijn, E. R. A., Bulten, B. H., Verkes, R. J., & Roelofs, K. (2012). Psychopaths lack the automatic avoidance of social threat: Relation to instrumental aggression. *Psychiatry Research*, *200*, 761-766.

Weertman, A., Arntz, A., & Kerkhofs, M. (2000). *SCID-II: Gestructureerd klinisch interview voor DSM-IV as II persoonlijkheidsstoornissen* [structured clinical interview for DSM-IV axis II personality disorders: **SCID-II**]. Lisse, The Netherlands: Swets & Zeitlinger.

West, B. T., Welch, K. B., & Galecki, A. T. (2014). Linear mixed models: A practical guide using statistical software. Boca Raton, FL: CRC Press.

Wistedt, B., Rasmussen, A., Pedersen, L., Malm, U., Träskman-Bendz, L., Wakelin, J., & Bech, P. (1990). The development of an observer-scale for measuring social dysfunction and aggression. *Pharmacopsychiatry*, *23*, 249-252. doi:10.1055/s-2007-1014514