Chapter 5

Does A Cognitive Skills Training Program for Prisoners Affect Neurocognitive Functioning and Heart Rate Activity?

Abstract

The current study investigates changes in neurobiological characteristics after cognitive skills training for prisoners. It was hypothesized that prisoners who completed a cognitive skills training program would show improved neurocognitive functioning and a normalization of heart rate activity. In addition, it was expected that neurobiological changes were related to behavioral improvement. A total of 190 male adult prisoners were included in the study. Of this sample, 121 prisoners participated in a cognitive skills training program and 69 prisoners served as a wait-list control group. Several neurocognitive skills and heart rate activity measures were assessed at pre and post-test assessment. In addition, trainers, prison officers, and prisoners were requested to evaluate behavioral changes over time. Results did not confirm the hypotheses. The absence of both neurobiological and behavioral improvement is discussed in light of the employed measures, the content and duration of the current intervention program, and the prison setting.
Introduction

Numerous brain-imaging studies have revealed impairments in prefrontal brain functioning associated with antisocial and violent behavior (for reviews see Koenigs, 2012; Wahlund & Kristiansson, 2009; Yang & Raine, 2009). The prefrontal brain area is known to give rise to complex 'neurocognitive' functions; that is, a collection of functions necessary for self-regulation and the regulation of socially appropriate behavior. Impairments in these neurocognitive functions have received considerable research attention in relation to different operationalizations of antisocial behavior (for reviews see Brower & Price, 2001; Meijers, Harte, Jonker, & Meynen, 2015; Morgan & Lilienfeld, 2000; Ogilvie, Stewart, Chan, & Shum, 2011). Overall, there is empirical evidence for a link between prefrontal brain impairments increasing the risk of engaging in antisocial behavior through decreasing, for example, empathic capacities, self-reflection, behavioral inhibition and emotion regulation (Raine, 2008).

During the last decade, several researchers have suggested that neuroscientific findings might provide us with more insight into how to best intervene with criminal behavior (Beauchaine, Neuhaus, Brenner, & Gatzke-Kopp, 2008; Glenn & Raine, 2014; Ross & Hoaken, 2010; Ross & Hilborn, 2008; Van Goozen & Fairchild, 2008; Vaske, Galyean, & Cullen, 2011). This is an important matter, since current intervention programs aimed to reduce antisocial behavior show varying degrees of success. For example, cognitive behavioral therapy (CBT), one of the most well-accepted treatment options for offenders, has a success rate varying from less than 10% up to almost 50% reduction of criminal recidivism across different studies (Landenberger & Lipsey, 2005; Lipsey & Cullen, 2007; Lipsey, Landenberger, & Wilson, 2007; McDougall, Perry, Clarbour, Bowles, & Worthy, 2009). A certain amount of that variability reflects statistical noise and unsystematic differences in study procedures and methods. However, according to Lipsey and Cullen (2007) much of it is related to substantive characteristics of the treatment and the offender samples to which the interventions are applied. Overall, the varying results suggest that criminal recidivism rates are not substantially reduced for a fairly large group of offenders.

Recent empirical literature has suggested promising effects of neuroscience in understanding why some intervention programs are effective in reducing recidivism rates and why some offenders benefit more from therapy than others. For example, Mullin and Simpson (2007) reported that poor neurocognitive functioning before a cognitive skills training program for prisoners was predictive of better treatment outcomes. However, in contrast to this finding, Fishbein et al. (2009) found that prisoners with poor neurocognitive functioning benefitted less from correctional intervention compared to those with normal
to high neurocognitive functioning. Results from our own research group (Cornet et al., 2015b) indicate that a specific neurocognitive task, the D-II Cancellation Task, has the potential to predict prisoners’ treatment dropout above and beyond several traditional background and behavioral characteristics, including self-reported treatment motivation. Although these studies indicate that there is a relationship between level of neurocognitive functioning and the effectiveness of correctional intervention, the underlying mechanisms remain poorly understood (Cornet, De Kogel, Nijman, Raine, & Van der Laan, 2014).

Interest in the relationship between neuroscience and the effectiveness of CBT is by far not limited to the correctional field. Much more research has been conducted on this issue in the field of clinical psychiatry and psychology. For example, Kandel (1998), a well-known psychiatrist and neuroscientist, has hypothesized that long-lasting changes in behavior due to psychotherapies (including CBT) are supported by changes in gene expressions. These changes in gene expression in turn modify the strength of synaptic connections in the brain and structural changes that alter the anatomical pattern of interconnections between nerve cells. In other words, psychotherapy does not only change the mind, but also the brain (Kandel, 1998). Empirical studies have confirmed Kandel’s hypotheses, showing strong evidence for ‘normalization’ of neural patterns after CBT in patients with various psychiatric disorders (including panic disorder and obsessive compulsive disorder; for reviews on this subject see Karlsson, 2011; Linden, 2006; Porto et al., 2009). With regard to correctional CBT programs, a theoretical article by Vaske et al. (2011) suggests that CBT may be effective in reducing problem behavior, including crime, because the intervention affects specific areas of the brain that are related to antisocial behavior (such as the prefrontal region). Overall, both empirical and theoretical studies have shown that there is a clinically relevant relationship between CBT and brain functioning in the way that the effectiveness of CBT is likely rooted in how this intervention program affects one’s brain functioning.

Nowadays, several correctional intervention programs (e.g., cognitive skills training and anger management training) claim to address neurocognitive deficits, such as poor problem solving and impulsivity. However, it is unclear as to whether these capacities are actually being targeted or improved (Ross & Hoaken, 2010). To our knowledge, only three studies have been conducted examining brain functioning after intervention programs aimed to reduce antisocial behavior (Lewis et al., 2008; Ross, 2012; Woltering, Granic, Lamm, & Lewis, 2011). Both Woltering et al. (2011) and Lewis et al. (2008) found that among children with problem behavior who showed improved behavior after a training program, called “Stop Now and Plan”, there was a reduction in ventral prefrontal brain activity as measured with electroencephalography (EEG). According to the authors, a reduction in
the ventral brain activity indicates increased efficiency of self-regulatory mechanisms. In contrast, the dissertation by Ross (2012) showed no significant improvement in neurocognitive functioning among adult prisoners after different intervention programs.

Overall, research on neurocognitive changes after correctional intervention is extremely limited and results are mixed. On the other hand, there seems to be relatively strong empirical support for changes in psychophysiological factors in response to intervention programs aimed to reduce antisocial behavior. A systematic literature review conducted by our research group revealed that values of specific physiological processes (e.g., basal cortisol levels) changed towards ‘less abnormal’ after intervention in individuals with different types of antisocial behavior (Cornet et al. 2015a). There is poor understanding of the underlying mechanisms that causes these physiological changes, but one idea is that (electro)physiological measures are more sensitive to detect treatment induced changes than behavioral measures (Bruce, McDermott, Fisher, & Fox, 2009). This suggests that future examination of treatment effectiveness might benefit from a multilevel assessment procedure including neurobiological methods.

The aim of the current study is to further explore the effect of a cognitive skills training program for prisoners on their neurocognitive and physiological (i.e., heart rate activity) functioning. Results might lead to suggestions on how to improve pre –and post-intervention assessment and how to best tailor current intervention programs towards offenders’ neurocognitive deficits/needs to increase the success rate of intervention and reduce the chances of reoffending. First, the current study examines changes in neurocognitive functioning and heart rate activity in male adult prisoners after cognitive skills training compared to a waitlist control group. Second, the relationship between behavioral changes after intervention and neurobiological changes is investigated in an attempt to better understand the working mechanism of the intervention. We hypothesize the following: 1) prisoners who complete a cognitive skills training program show improved neurocognitive functioning, and 2) a change in heart rate activity towards ‘less abnormal’. Finally, 3) it is expected that treatment induced neurobiological changes are related to behavioral improvement.

**Method**

**Participants**

Participants consisted of male adult prisoners selected by Probation Service officers for cognitive skills training. In total, 121 prisoners started intervention (‘intervention group’) and 69 prisoners served as a waitlist control comparison group (‘control group’). The control
group participants were selected for the cognitive skills training program, but were not able to start the intervention due to various practical reasons (e.g., no trainers available or insufficient number of candidates to start intervention). For both groups, the only reason for exclusion from participation in the current study was an unstable psychological or physical condition as reported by the prisoners themselves. The study was approved by the Medical Ethics Committee of the VU University Medical Center Amsterdam (Ref.no: NL36062.029.11 CCMO), while informed consent to participate in this study was sought from the prisoners. Participation was voluntary and was compensated with €25.

Cognitive Skills Training
Treatment content as well as the treatment selection procedure are described elsewhere (Cornet et al. 2015b). In short, the cognitive skills training (called ‘CoVa’) central to this study is an adapted and translated version of the English ‘Enhanced Thinking Skills’ (ETS) program (Clark, 2000). In the Netherlands, a screening instrument based on the Risk-Need-Responsivity model called the ‘RISc’ (Recidive Inschattings Schalen, a Dutch translation of the Offender Assessment System (OASys; Howard, Clark, & Garnham, 2003) is used on a national scale to allocate prisoners to intervention programs based on their criminal needs and cognitive deficits (Vinke, Vogelvang, Erftermeijer, Veltkamp, & Bruggeman, 2003).

Design and Procedure
Figure 1 displays the recruitment process of participants in both groups. Lists of selected participants for treatment were requested from several prisons in the Netherlands between 2011 and 2014. Recruitment process details of the intervention group are described elsewhere (Cornet et al. 2015b). Control group participants were put forward by prison contact persons. Inclusion criteria for control group participants were: 1) being selected by Probation Service officers for CoVa participation, 2) no training participation within the upcoming two months, and 3) no history of CoVa participation. Once potential participants showed interest in the current study, an introductory meeting was scheduled either between the contact person and the prisoner or between one of the researchers and the prisoner. In total, 77 prisoners volunteered to take part in the study. Of these, 69 prisoners attended the pre-test measurement. Reasons for attrition before pre-test phase were ‘no-show’ or practical reasons (e.g., other activities were scheduled). Of the 69 participants, 16 prisoners did not complete the post-test measurement due to various reasons, such as provisional release from prison, transfer, or unforeseen CoVa participation. Overall, 53 prisoners in the control group completed both the pre and post-test measurement.

1 Contact persons were prison officers who were involved in the organizational aspects of the cognitive skills training (e.g., arranging trainers, workspace, etc.).
Researchers and Master students with a background in psychology, who were independent of the correctional facility, assessed participants at different time intervals. At pre-test assessment, participants completed several neurocognitive tasks and underwent heart rate measurement. This assessment was repeated after intervention completion for participants in the intervention group with an average of 93 days between pre- and post-test. For control group participants, post-test assessment was scheduled on average 73 days after pre-test assessment. Furthermore, prison officers were asked to evaluate prisoners' aggressive behavior on the ward twice (shortly after pre-test and right before post-test for both groups). In addition, CoVa trainers were also asked to evaluate the behavior of prisoners in the intervention group at two time periods: at the beginning and at the end of intervention.

**Measures**

**Verbal intelligence level**

A Dutch version of the National Adult Reading Test (Nederlandse Leestest voor Volwassenen [NLV]) was used as a measure of verbal IQ (Nelson, 1982; Nelson & O’Connell, 1978; Schmand, Lindeboom, & van Harskamp, 1992). The total NLV score appears to correlate highly with the Wechsler Adult Intelligence Scale total IQ score (.74) and the total verbal IQ score (.85) (Schmand et al., 1992). The NLV score is not valid for non-native Dutch speakers (personal communication B. A. Schmand, 22 December, 2010). For this
reason, participants who did not complete at least primary school in the Netherlands were not eligible to complete the NLV test (N = 14 in the intervention group and N = 8 in the control group).

**Neurocognitive functioning**

Neurocognitive tasks were selected based on the following: 1) literature concerning the relationship between cognitive dysfunctioning and antisocial behavior (Brower & Price, 2001; Morgan & Lilienfeld, 2000; Ogilvie et al., 2011), 2) domains of cognitive functioning addressed by the CoVa training (interpersonal problem solving skills, self-control, social perspective taking skills and critical/moral reasoning) and 3) studies included in the literature review that found a predictive value of specific cognitive function tasks in relation to CBT for prisoners (Fishbein et al., 2009; Mullin & Simpson, 2007). This resulted in the selection of the following neurocognitive tasks: the Stroop Color Word Task, Concept Shifting Test (CST), the D2 Cancellation Task, the WAIS-III Digit Span, the Modified Wisconsin Card Sorting Task (M-WCST), the Tower of Hanoi (TOH), the Stop it Task, the Reading the Mind in the Eyes Task, and the Controlled Oral Word Association Task (COWAT). A detailed description of the selected neurocognitive tasks can be found in the supporting information (Table S1).

**Physiological functioning**

Heart rate activity was measured using the VU-AMS ambulatory monitoring system (VU University, Amsterdam; De Geus, Willemsen, Klaver, & Van Doornen, 1995; Willemsen, De Geus, Klaver, Van Doornen, & Carroll, 1996). The device continuously recorded heart rate (HR, in beats per minute; bpm) and heart rate variability (Respiration Sinus Arrhythmia (RSA) in milliseconds; msec), which is a measure of parasympathetic functioning (Berntson, Cacioppo, & Quigley, 1991). VU-AMS software package 3.5 automatically scored all beats in the electrocardiogram (ECG). A trained junior researcher used the VU-AMS Manual version 1.1 to manually check the entire recording to detect irregularities. The dependent measures were HR reactivity (= HR during stress minus HR during baseline) and RSA reactivity (RSA during stress minus RSA during baseline). The stressful period entailed a 4-minute VU-AMS recording during the D2 Cancellation Task. Baseline heart rate activity was measured during a 5-minute relaxation moment in which the participant was sitting behind a laptop screen with headphones on and was presented with emotionally neutral pictures (e.g., nature pictures) and relaxing classical music.

**Criminal record and RISc file information**

Previous research has shown the relationship between specific sentence-related characteristics and treatment outcome of offenders (Fishbein et al., 2009; Mullin &
Simpson, 2007). For this reason, we included available sentence-related characteristics for both groups. Information with regard to index offense and the age at first judicial confrontation was derived from RISc records. The index offense was classified as either violent (e.g., robbery, attempted murder) or non-violent (e.g., theft, drug trafficking). In addition, level of cognitive deficits as evaluated by the Probation Service was also obtained from RISc files. Prisoners were assessed using the RISc instrument on the following cognitive aspects: impulsiveness, problem insight, problem management, future orientation, and thinking/learning style. Finally, the Custodial Institutions Agency (DJI) of the Ministry of Security and Justice provided information on the total number of previous imprisonments.

**Intervention outcome measures**

To explore whether neurobiological changes are associated with behavioral improvement after intervention, the effectiveness of the intervention on behavior was assessed with three different treatment outcome measures. First, prisoners’ behavior in the intervention group was assessed with the English questionnaire *Treatment Gain: Short Scale*, used by Fishbein and Sheppard (2006). The questionnaire was translated into Dutch and edited for the current study. Trainers were asked to complete the questionnaire comprising seven multiple-choice questions on participant’s skills, participation, and competence during treatment. The questionnaire was completed at two time periods: shortly after the start of treatment and before the end of treatment. Cronbach’s alpha was found to be good (.84) for both assessments.

Second, prison officers were asked to evaluate prisoners’ behavior using the *Social Dysfunction Aggression Scale* (SDAS; Wistedt et al., 1990). The SDAS is an eleven-item (scaled 0 to 4) period-based instrument that records a variety of aggressive behaviors, ranging from mild to moderate to severely aggressive behavior. The reliability of the SDAS has been found to be sufficient (Kobes, Nijman, & Bulten, 2012; Wistedt et al., 1990; Zaalberg et al., 2015). The SDAS was also completed at two time periods: at the start and end of treatment. Prison officers were asked to rate prisoners’ aggressive behavior in the past week with questions such as “To what extent did the prisoner use verbal aggression in the past week?”. This questionnaire was also completed at two time periods for prisoners in the control group.

Finally, a three-item questionnaire was specifically designed to assess self-reported treatment gain. Participants in the intervention group were asked to answer the following questions: ‘Do you think you benefitted from the training?’, ‘To what extent do you think you are able to apply the acquired skills outside prison?’ and ‘Do you notice any changes
in the way you think/behave after completing treatment?’. The dependent variable is the total score on the three items. The internal reliability of this measure appears to be acceptable (Cronbach’s alpha = .72). A principal axis factor analysis was conducted on the three items with oblique rotation (direct oblimin). This analysis revealed that the three items loaded on one similar factor (only one factor with an eigenvalue greater than 1 was found that explained more than 60% of the variance).

Statistical Analysis
Attrition rate and imputation strategies
In both groups, missing values were identified on several variables. In the intervention group, the average percentage of missing values per variable was 16.9% (range 0 - 43.8%). Comparable results were found in the control group, with an average percentage of missing values 16.2% (range 0 - 42%). The highest percentage of missing values was detected on post-test measures. A complete case approach, in which subjects with missing values are simply excluded from analyses, would lead to a decrease in statistical power and could result in a bias if the remaining cases are not representative of the complete sample. Therefore, we decided to impute missing values with a multiple imputation (MI) technique (Rubin, 1987; Van Buuren, 2012). Details on how MI was performed can be found in the supporting information.

In order to impute data, the imputation model should have sufficient data to estimate the model. For this reason, it was decided to exclude participants with a relatively high percentage of missing values (more than 50%) (IBM, and personal communication Van Buuren, 9-2-2015). In the control group, five prisoners completed less than 50% of the measures and therefore the final sample for imputation consisted of 64 prisoners in the control group. In the intervention group, 35 prisoners did not complete treatment. Because less or no intervention effect on both behavioral and neurobiological measures was expected for non-completers, we decided not include these 35 participants in analyses. Of the remaining 86 participants, two prisoners had more than 50% missing data. Thus, 84 participants in the intervention group were found eligible for further analyses.

Analytic approach
First, independent sample t-tests were conducted to test for group differences on relevant background characteristics (Table 1). Results indicate that both groups were identical on most background characteristics. However, there appeared to be a small, but significant, difference in age between the two groups. In addition, a higher proportion of offenders were imprisoned for a violent offense in the intervention group compared to the control group (80.7% versus 59.53% respectively). To account for these differences in
background characteristics, age and index offense are treated as covariates in subsequent analyses. Furthermore, ANCOVAs were performed, using general linear models to test for intervention effects. Before analyses, assumptions of ANCOVA (e.g., normality, linearity, and homogeneity of regression slopes) were verified. In addition, data were screened for outliers using 5% trimmed mean information and by inspecting boxplots. "Extreme" outliers, as defined by SPSS, were removed from subsequent analyses. These outliers were caused by various factors, such as measurement errors (as happened with the heart rate equipment), extreme fatigue as reported by the prisoner, or prisoners’ lack of understanding of task instructions.

To account for variation in neurocognitive functioning and heart rate activity on the post-test, we examined whether subjects in both groups differed on specific confounders (country of birth, history of brain injury, smoking, physical activity, medication usage and body mass index). There appeared to be a small significant difference in the average number of cigarettes smoked per day between the intervention group, who smoked on average 6.62 (SE = .66) cigarettes per day, and the control group, who smoked on average 9.50

Table 1: Background Characteristics of Intervention and Control Group Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intervention</th>
<th>Control</th>
<th>Independent t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age</td>
<td>28.58</td>
<td>7.90</td>
<td>31.84</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>80.55</td>
<td>9.85</td>
<td>79.78</td>
</tr>
<tr>
<td>Number of previous imprisonments</td>
<td>3.55</td>
<td>3.04</td>
<td>3.59</td>
</tr>
<tr>
<td>Age at first judicial contact†</td>
<td>18.02</td>
<td>5.56</td>
<td>19.05</td>
</tr>
<tr>
<td>Recidivism risk†</td>
<td>2.89</td>
<td>0.93</td>
<td>2.98</td>
</tr>
<tr>
<td>Cognitive deficits†</td>
<td>5.43</td>
<td>1.90</td>
<td>5.37</td>
</tr>
<tr>
<td>Index offense†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Violent offense</td>
<td>80.7%</td>
<td>59.53%</td>
<td>χ²</td>
</tr>
<tr>
<td>Non-violent offense</td>
<td>19.3%</td>
<td>40.47%</td>
<td>8.046</td>
</tr>
</tbody>
</table>

* p < .05
† Derived from RISc files completed by Probation Service officers.

"Extreme" outliers are scores more than 3*IQR from the rest of the scores. IQR stands for the 'interquartile range', which is the middle 50% of the scores.
(SE = 1.23) cigarettes per day (p = .039). Furthermore, prisoners in the control group were significantly more likely to report a change in medication usage between pre and post-test compared to the intervention group ($\chi^2(1, N = 148) = 6.61, p = .01$). Because literature has suggested that both smoking and medication can affect heart rate activity (e.g., Armstrong, Keller, Franklin, & MacMillan, 2009; Portnoy et al., 2014; Scarpa, Haden, & Tanaka, 2010), these variables were included as additional covariates in subsequent analyses. In addition, some medications, such as antidepressants, might affect neurocognitive functioning (Biringer, Rongve, & Lund, 2009), and therefore change in medication usage between pre and post-test was also included as a covariate in the ANCOVAs on neurocognitive functioning.

To investigate whether the intervention group showed improved neurocognitive functioning and heart rate activity after treatment compared to the control group, separate ANCOVAs were run on every post-test score including relevant covariates (e.g., pre-test score, age, index offense, medication and smoking) and group membership (intervention group or control group) as fixed factor. Effect sizes were calculated with the Campbell Collaboration effect size calculator.

## Results

### Changes in Neurocognitive Functioning and Heart Rate Activity

Table 2 shows the means and standard deviations on all measures for both the intervention group and the control group. In addition, results of the ANCOVA analyses on post-test scores are displayed. Sample sizes differ due to the removal of extreme outliers and preselection of valid predictors for imputation (i.e., post-test TOH errors in the control group; see supporting information). Neurocognitive tasks are divided into two sections: (1) a collection of measures assumed to decrease between pre and post-test (e.g., number of errors), and (2) measures that are assumed to increase between pre and post-test (e.g., total correct). Overall, performances on tasks changed into the expected direction for both groups. However, the results show no support for a treatment effect of the CoVa training on neurocognitive functioning and heart rate activity.

As mentioned earlier, there are reasons to suspect that specific subgroups benefit more from therapy than others. One hypothesis based on previous research, is that inmates with low neurocognitive skills pre-treatment benefit less from intervention compared to those with normal to high neurocognitive functioning (Cornet, Van der Laan, Nijman, Tollenaar, & De Kogel, 2015; Fishbein et al., 2009). To test this, equal percentile groups were created based on pre-test neurocognitive functioning (low, medium, and high neurocognitive...
Table 2 Changes in Neurocognitive Functioning and Heart Rate Activity Between Pre- and Post-Test Assessment

<table>
<thead>
<tr>
<th>Measures assumed to decrease</th>
<th>Intervention</th>
<th>Control</th>
<th>Intervention</th>
<th>Control</th>
<th>ANCOVA</th>
<th>F</th>
<th>p</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop Color Word Task – interference (sec)</td>
<td>83</td>
<td>29.71(15.16)</td>
<td>64</td>
<td>29.47(11.12)</td>
<td>84</td>
<td>25.27(10.93)</td>
<td>64</td>
<td>26.33(12.03)</td>
</tr>
<tr>
<td>Concept Shifting Test – Shifting score (sec)</td>
<td>80</td>
<td>-4.21(7.14)</td>
<td>63</td>
<td>-3.71(7.60)</td>
<td>84</td>
<td>-2.51(6.83)</td>
<td>62</td>
<td>-3.53(5.51)</td>
</tr>
<tr>
<td>M-WCST – Total errors</td>
<td>84</td>
<td>11.42(9.58)</td>
<td>64</td>
<td>12.35(11.45)</td>
<td>84</td>
<td>8.49(8.29)</td>
<td>64</td>
<td>7.93(7.45)</td>
</tr>
<tr>
<td>Tower of Hanoi - Errors</td>
<td>83</td>
<td>6.10(4.52)</td>
<td>62</td>
<td>7.35(5.93)</td>
<td>83</td>
<td>5.23(5.00)</td>
<td>50</td>
<td>5.64(3.88)</td>
</tr>
<tr>
<td>Stop it – SSRT (msec)</td>
<td>84</td>
<td>221.17 (44.88)</td>
<td>63</td>
<td>272.68 (73.06)</td>
<td>84</td>
<td>219.54 (48.15)</td>
<td>64</td>
<td>251.96 (41.68)</td>
</tr>
</tbody>
</table>

| Measures assumed to increase                  |             |         |             |         |        |         |         |          |
| D-II Cancellation test – CP                  | 84           | 170.05(30.68) | 64           | 164.92(34.03) | 84      | 192.60(39.42) | 64      | 189.16(35.42) | .4800 | .510 | .117 |
| WAIS Digit Span – Total correct backward    | 84           | 5.76(1.97)    | 64           | 5.53(1.77)    | 84      | 6.19(1.82)    | 64      | 5.73(1.95)    | 1.4838 | .226 | .2021 |
| M-WCST – Categories correct                 | 84           | 4.84(1.67)    | 64           | 4.69(1.90)    | 81      | 5.25(1.31)    | 64      | 5.28(1.23)    | 1.9464 | .167 | .234 |
| Tower of Hanoi – Total planning time (sec)  | 84           | 72.95(28.53)  | 63           | 75.23(30.90)  | 84      | 74.42(28.51)  | 63      | 73.77(30.46)  | .0073 | .932 | .0142 |
| Reading the Mind in the Eyes – total correct | 84           | 19.70(3.04)   | 64           | 19.39(3.77)   | 84      | 20.23(2.70)   | 64      | 19.80(2.63)   | .3745 | .544 | .1015 |
| COWAT – total words correct                 | 84           | 27.61(9.51)   | 64           | 30.52(9.74)   | 83      | 30.83(8.81)   | 64      | 31.63(9.20)   | .9378 | .335 | .1611 |

| Heart rate measures                          |             |         |             |         |        |         |         |          |
| Heart rate – rest (bpm)                      | 84           | 69.07(9.79)  | 64           | 68.83(13.19) | 84      | 70.07(10.59) | 64      | 70.97(9.32)  | .5258 | .471 | .1203 |
| Heart rate – D-II task (bpm)                 | 84           | 76.65(10.93) | 63           | 75.07(12.69) | 84      | 76.63(11.44) | 64      | 77.03(10.43) | .6792 | .412 | .1374 |
| RSA – rest (msec)                            | 84           | 85.99(52.19)  | 63           | 78.20(44.32)  | 83      | 79.26(50.06)  | 64      | 69.93(36.23)  | .0238 | .878 | .0258 |
| RSA – D-II task (msec)                       | 82           | 63.19(39.90)  | 61           | 58.03(36.77)  | 79      | 56.41(33.62)  | 64      | 52.19(26.02)  | .3269 | .569 | .0972 |

Note. M-WCST = Modified Wisconsin Cart Sorting Test, SSRT = Stop Signal Reaction Time, CP = Concentration Performance, WAIS = Wechsler Adult Intelligence Scale, COWAT = Controlled Oral Word Association Task, RSA = Respiration Sinus Arrhythmia.

1 This result did not remain significant after correction for multiple testing.
functioning) by combining z-scores (some variables were transformed to create one overall z-score that indicated strength of neurocognitive performance). Initial ANCOVA analyses were re-analyzed including group × level of pre-test cognitive functioning interaction term. The modified Hochberg correction was applied for controlling overall type I error in the case of multiple testing (Rom, 2013). It appeared that none of the subsequent analyses revealed a significant interaction effect of group × level of cognitive functioning on post-test scores. In other words, no differences were found between intervention and control group participants on their initial level of neurocognitive functioning in relation to post-test outcome. In addition, previous literature has shown an effect of age on correctional treatment outcome (e.g., Mullin & Simpson, 2007). To further explore the existence of age-related subgroups associated with change in neurocognitive functioning after intervention, equal percentile groups were created based on age (low, medium, high age groups). Again, no significant interaction effects of group × age were detected on the post-test scores.

Furthermore, one hypothesis central to the current study was that heart rate activity would become ‘less abnormal’ in response to intervention. In other words, it was expected that prisoners with relatively low initial heart rate activity would show increased heart rate activity after intervention and prisoners with relatively high heart rate activity would show a decrease after intervention compared to the control group. To test this hypothesis, heart rate and respiration sinus arrhythmia (RSA) during rest and stressful phase were transformed to z-scores and recoded to create one index of heart rate activity. Equal percentile groups were formed (low, medium and high heart rate activity at pre-test) to test for significant interaction effects of group × initial heart rate activity on post-test heart rate measures. However, no significant interaction was found.

**Behavioral changes**

With regard to behavioral changes, evaluations of prisoners’ behavior in the intervention group as reported by trainers, significantly rose from 17.54 (SE = .39) at pre-test to 18.73 (SE = .35) at post-test (p < .00). Nevertheless, this positive change in behavior was not reflected by evaluations of the intervention group reported by prison officers, as SDAS scores were 2.87 (SE = .48) at pre-test and 2.10 (SE = .34) at post-test (p = .394). A similar pattern was observed for prisoners in the control group, where SDAS scores changed from 3.18 (SE = .51) to 2.81 (SE = .49) from pre- to post-test (p = .492).

With regard to the self-evaluation, prisoners in the intervention group responded relatively positively on question 1 (‘Do you think you benefitted from the training?’), where 66% responded ‘yes’, and question 2 (‘To what extent do you think you are able to apply the acquired skills outside prison?’), where 81% answered between ‘sufficient’ and ‘excellent’.
However, answers were mixed on the final question ('Do you notice any changes in the way you think/behave after completing treatment?'), where 45% answered 'yes' and 55% responded either 'no' or 'don't know'.

The second aim of this study was to investigate the relationship between behavioral changes and changes in neurocognitive functioning and heart rate activity after intervention. However, no significant changes were detected on neurocognitive tasks and in heart rate activity in favor of the intervention group. Nevertheless, mean group scores could have masked the presence of individual improvement on specific tasks and behavioral measures. To examine the relationship between neurobiological changes and behavioral improvement, change scores were created by subtracting pre-test from post-test scores. Table 3 presents the correlations between changes in behavioral evaluations, as reported

### Correlations Between Behavioral Evaluations and Changes in Neurocognitive Functioning and Heart Rate Activity

<table>
<thead>
<tr>
<th></th>
<th>Trainer change</th>
<th>SDAS change</th>
<th>Self-evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop Interference</td>
<td>-.050a</td>
<td>.027a</td>
<td>-.034a</td>
</tr>
<tr>
<td>CST Shifting</td>
<td>-.061a</td>
<td>-.097</td>
<td>-.145</td>
</tr>
<tr>
<td>COWAT</td>
<td>-.151</td>
<td>-.122b</td>
<td>-.018b</td>
</tr>
<tr>
<td>DII - CP</td>
<td>.032</td>
<td>.146</td>
<td>-.032</td>
</tr>
<tr>
<td>TOH errors</td>
<td>.004b</td>
<td>-.069b</td>
<td>-.140b</td>
</tr>
<tr>
<td>TOH Planning time</td>
<td>.042</td>
<td>-.108</td>
<td>-.058</td>
</tr>
<tr>
<td>Reading the Mind in the Eyes</td>
<td>.079</td>
<td>-.013</td>
<td>-.120</td>
</tr>
<tr>
<td>Digit span backward</td>
<td>.020</td>
<td>.059</td>
<td>.053</td>
</tr>
<tr>
<td>MWCST categories correct</td>
<td>-.052c</td>
<td>.079b</td>
<td>-.077c</td>
</tr>
<tr>
<td>MWCST errors</td>
<td>.136</td>
<td>-.064</td>
<td>.108</td>
</tr>
<tr>
<td>SSRT</td>
<td>-.088</td>
<td>-.116</td>
<td>-.044</td>
</tr>
<tr>
<td>Heart rate rest</td>
<td>.120</td>
<td>.105</td>
<td>.165</td>
</tr>
<tr>
<td>Heart rate stress</td>
<td>.143</td>
<td>.159</td>
<td>.146</td>
</tr>
<tr>
<td>RSA rest</td>
<td>-.009b</td>
<td>.036b</td>
<td>-.086e</td>
</tr>
<tr>
<td>RSA stress</td>
<td>-.036b</td>
<td>-.201b</td>
<td>-.174e</td>
</tr>
</tbody>
</table>

**Note.** CST = Concept Shifting Test, COWAT = Controlled Oral Word Association Task, CP = Concentration Performance, TOH = Tower of Hanoi, MWCST = Modified Wisconsin Cart Sorting Test, SSRT = Stop Signal Reaction Time, RSA = Respiration Sinus Arrhythmia.

N = 84 unless otherwise stated.
aN = 83; bN = 82; cN = 81; dN = 80.
eAn increase in SDAS indicates an increase in aggressive behavior on the ward.
by the trainer, the prison officer, and the prisoners themselves, and neurocognitive/heart rate activity changes. Overall, no significant correlations were found. Thus, the results indicate that the significant behavioral improvement after intervention as reported by trainers was not reflected in neurocognitive or heart rate activity improvement.

**Discussion**

The aim of the present study was to examine changes in neurocognitive functioning and heart rate activity in male adult prisoners following cognitive skills training compared to a waitlist control group. It was hypothesized that prisoners who completed cognitive skills training would show improved neurocognitive functioning and a change in heart rate activity towards ‘less abnormal’. Furthermore, it was expected that treatment-induced neurobiological changes were related to behavioral improvement. Overall, the empirical results of the current study did not support these hypotheses.

Results did not provide evidence for a change in neurocognitive functioning in favor of the intervention group. Although these results do not confirm our first hypothesis, they are in line with a similar study conducted by Ross (2012). In her dissertation, Ross’ sample comprised 119 prisoners, of which 69 completed an intervention (e.g., violence prevention program or substance abuse program) and 50 inmates served as a control group. Overall, no changes in neurocognitive functioning were detected among prisoners who completed intervention. Ross (2012) argued that one possible reason for the absence of change among these prisoners was a lack of impairments in neurocognitive functioning at pre-test. A lack of neurocognitive deficits might suggest that there is no, or at least reduced potential for improvement. In the current study, data from RISc files completed by Probation Service officers supported the presence of cognitive deficits among prisoners. However, the absence of data from prisoners not selected for intervention or a healthy control group makes it difficult to conclude whether impairments on neurocognitive tasks were also present. To verify if the current sample suffered from impaired neurocognitive functioning at pre-test, a post-hoc literature comparison was performed.³

The balance of the evidence suggests that prisoners in the current sample generated substantially fewer words on the COWAT compared to healthy controls and other offender samples. This phonemic verbal fluency task requires participants to retrieve words from their lexicon memory while remembering the instructions and earlier responses in

³ Results from this literature comparison are available from the authors upon request.
working memory. In addition, participants are required to suppress irrelevant responses and repetition. In general, the task requires a complex set of cognitive skills and involves executive control processes (e.g., the ability to regulate one’s thoughts and direct behavior toward a general goal) (Henry & Crawford, 2004; Henry, Crawford, & Phillips, 2004; Lezak, Howieson, Bigler, & Tranel, 2012; Shao, Janse, Visser, & Meyer, 2014). Overall, the results are in contrast with those of Ross (2012), who did found no impairments in prisoners’ neurocognitive functioning pre-treatment. The impaired verbal fluency skills in the current sample might suggest that there was potential for improvement.

Another possible explanation for the absence of neurocognitive improvement after intervention is related to the selection of the neurocognitive tasks we used. It is conceivable that the selected neurocognitive skills were not targeted during intervention, and therefore may not have undergone change. Although all measures were selected with care and based on several arguments, there is no golden standard how to best assess neurocognitive deficits in an accurate and valid manner (Chan, Shum, Touloupolou, & Chen, 2008). Nevertheless, the fact that the current sample displayed impaired verbal fluency skills has clinically relevant implications. This finding suggests that in order to be more effective, current intervention programs should be tailored towards specific neurocognitive deficits, as it has been shown that these skills, such as impaired verbal fluency skills, are consistently related to antisocial behavior (Morgan & Lilienfeld, 2000; Ogilvie et al., 2011).

The effectiveness of the current intervention program is a hot-topic, since research evaluations of the ETS program have provided mixed results regarding its success in reducing reoffending rates. While there is evidence for a 14% reduction in recidivism rates (Friendship, Blud, Erikson, Travers, & Thornton, 2003), there are also studies that show no difference in recidivism rates between ETS participants and control subjects (see for overview Gobbett & Sellen, 2014). In addition, a recent study of more than 1,400 prisoners in the Netherlands found that inmates reported a small significant improvement in interpersonal problem solving, self-control, and social perspective-taking skills, but no significant improvement in critical and moral reasoning skills after the CoVa training (Buysse & Loef, 2012). However, definite conclusions about the effectiveness of the CoVa training cannot be drawn from the study by Buysse and Loef because no control group was included. In addition, recidivism data after the CoVa training are currently not available in the Netherlands.⁴

⁴ At this moment the Research and Documentation Centre, Ministry of Security and Justice, the Hague, the Netherlands, is currently analysing recidivism rates of individuals who completed the CoVa training.
The study by Buysse and Loef (2012) also revealed that one third of the provided CoVa trainings between 2008 and 2012 did not meet the treatment integrity criteria and that 32% of the inmates did not fully meet the CoVa suitability criteria. A relatively low adherence to the suitability criteria was also found in a study by Sadlier (2010) where only 58% of the ETS participants fully met the targeting criteria. Overall, the efficiency of the CoVa training appears to be limited and behavioral results from the current study seem to support this finding. While trainers did report some improvement in prisoners’ behavior following treatment, this was not reflected by behavioral evaluations completed by prison officers. Furthermore, although inmates were relatively positive about the content of the training, the majority did not experience a change in the way they thought or behaved after intervention. All in all, the relatively low efficiency of the CoVa training could explain the absence of change in behavioral measures and neurobiological factors and the nonexistence of a relationship between these two concepts.

In response to the mixed success rates, the ETS program is currently being gradually phased out of delivery in England and Wales to be replaced with a newly developed Thinking Skills Program (TSP; Gobbett & Sellen, 2014; Riddly, 2010). The TPS program differs from ETS in several ways (Riddly, 2010). For example, the program adopts a greater and more inclusive focus on emotions and generalizes central cognitive skills to everyday life situations. In addition, the program focuses more on personal needs and strengths by incorporating individual sessions alongside 15 group sessions. A first evaluation study on program effectiveness seems to support greater self-reported behavioral improvements among prisoners who completed TPS compared to those allocated to ETS (Gobbett & Sellen, 2014). Recently, the Dutch Probation Service has proceeded to implement these adaptations to the CoVa training, resulting in ‘CoVa 2.0’. Unfortunately, it is too early to verify whether CoVa 2.0 has more potential to improve offenders’ behavior than the original version of the training.

Another important factor that may have hampered significant behavioral and neurobiological changes in our sample concerns the intervention format. The current cognitive skills training program comprises 20 sessions divided over 10 weeks. This is in accordance with a typical course of CBT programs for anxiety or depression which usually requires up to 15 sessions (Beck, 1995). However, it has been suggested that for more chronic behavior problems, such as personality disorders, 12 months or more is necessary for clinical improvement to occur (Davidson, 2007). Prisoners often suffer from severe and complex mental, behavioral, and physical problems (WHO, 2007). In addition,

---

5 All prisoners in the current sample completed the original CoVa training.
a meta-analysis comprising over 18,000 male adults prisoners revealed that 65% suffer from a personality disorder, including antisocial personality disorder (47%) (Fazel & Danesh, 2002). This information raises the question of whether the duration of existing correctional training programs facilitates long-term behavioral changes.

Finally, the prison setting itself may also play an important role in the effectiveness of intervention programs. Prisons are clear examples of impoverished and sedentary environments. For example, prison life is characterized by physical inactivity and little control over daily life routines (Cashin, Potter, & Butler, 2008; Woodall, Dixey, & South, 2014). Results from both animal and human studies have shown that an impoverished environment negatively affects prefrontal cortex functioning, leading to reduced neurocognitive skills (Melendez, Gregory, Bardo, & Kalivas, 2004; Mendes et al., 2013; Volkers & Scherder, 2011). If the impoverished nature of prison environments has a negative impact on neurocognitive functioning then this might counteract the potential positive effects of cognitive skills training. Therefore, one assumption is that comprehensive enrichment of the prison setting (including more social challenges and more control over daily activities) combined with prolonged correctional therapy might produce greater effects in improving neurocognitive functions and, as a result, reduce recidivism rates (Meijers et al., 2015).

Our second hypothesis was that prisoners with relatively low initial heart rate activity would show increased heart rate activity after intervention, and prisoners with relative high heart rate activity would show decreased heart rate activity following intervention compared to the control group. However, the results did not support this ‘normalization’ of heart rate activity. To verify whether the absence of physiological change was due to deviant heart rate levels at pre-test, another literature comparison was performed. Only resting heart rate levels were investigated since this is the most robust physiological indicator examined in samples with antisocial behavior (Lorber, 2004; Ortiz & Raine, 2004; Portnoy & Farrington, 2015). It appeared that the current sample was characterized by a slightly reduced resting heart rate compared to healthy controls as well as compared to other offender samples. It has been suggested that indices of low physiological arousal, such as low resting heart rate, are predictive of less benefit from behavioral intervention programs aimed to reduce antisocial behavior (Cornet et al., 2014). This might explain the nonexistence of neurobiological as well as behavioral improvement after intervention. However, the average heart rate levels of both groups (69.07 bpm for the intervention group and 68.83 bpm in the control group) are still

---

6 Results from this literature comparison are available from the authors upon request.
within a clinically healthy range.\textsuperscript{7} In addition, the performed literature comparison did not control for influential factors such as medication usage, psychiatric disorders, and BMI, which are assumed to affect heart rate levels. Overall, definite conclusions cannot be drawn on the effect of initial heart rate activity on the absence of physiological changes after intervention.

Two unexpected sample characteristics were discovered. First, as displayed in Table 1, both the intervention and control group showed a relatively low verbal intelligence level (ranging between 79-81), whereas a similar study conducted by Mullin and Simpson (2007) found an average verbal intelligence level of 93 ($SD = 12$) among prisoners selected for ETS. One of the CoVa inclusion criteria is an average IQ over 90. Prisoners with lower intellectual abilities are generally assigned to an adjusted version of the training (‘CoVa-plus’). However, an official IQ measurement is currently not part of the pre-treatment selection procedure (Ferwerda, van Wijk, Arts, & Kuppens, 2009). A serious practical implication might result from these findings, since there is literature indicating that treatment success is influenced by prisoners’ verbal intelligence level (Andrews & Dowden, 2007; Dowden & Serin, 2001). This suggests that the success rate of the cognitive skills training might benefit from better selection procedures concerning the intellectual abilities of the offenders. In addition, Table 1 revealed that participants in the control group were much more frequently confined for a non-violent offense and less frequently for a violent offense compared to the intervention group. With the limited information available from file records, it is unclear what might have caused these group differences.

Additional limitations central to the current study should be considered. First, of the 69 prisoners in the control group, 19 were offered different training programs between pre and post-test assessment. Examples of these included “Budgeting” and “Lifestyle” programs, with the latter aimed to reduce substance abuse and gambling behavior. The completion of different training programs by participants in the control group might have reduced potential group differences in neurobiological functioning on post-test measurements. A second limitation that should be considered is the reliability and sensitivity of the behavioral measures employed in the current study. For example, the self-evaluation measure has not been previously validated. In addition, the mean SDAS score of 2.87 at pre-test in the intervention group is low compared to other studies in forensic or psychiatric samples (e.g., Rossberg & Friis, 2003; Zaalberg, Nijman, Bulten, Stroosma, & van der Staak, 2010; Zaalberg et al., 2015), with the mean SDAS scores being

\textsuperscript{7} http://www.heart.org/HEARTORG/Conditions/More/MyHeartandStrokeNews/All-About-Heart-Rate-Pulse_UCM_438850_Article.jsp
9.9, 5.1, and 12.2 respectively. This might have reduced the possibilities for a (statistically significant) decline of these scores in the present sample. A final limitation concerns the examination of psychophysiological changes. In the current study, only heart rate activity was measured. However, there is much stronger empirical support for normalization of cortisol levels, which is considered as the stress hormone, among children and adolescents with antisocial behavior in response to intervention (Cornet, De Kogel, Nijman, Raine, & Van der Laan, 2015). We decided not to collect saliva samples for cortisol analysis to reduce the demand and increase prisoners’ willingness to participate in the current study. For this reason, it remains unclear whether a cognitive skills training program is able to change psychophysiological characteristics in prisoners. Future research should verify if cortisol levels are able to ‘normalize’ in response to behavioral intervention for adults with antisocial behavior.

Overall, the limitations suggest that the absence of change in both behavioral as well as neurobiological measures do not automatically imply that there was no improvement after intervention at all. The quality and selection of behavioral and neurobiological measures might have hampered the detection of improvement. In addition, changes in neurological structures may take longer than changes in behavior and emotional reactivity and therefore, it is possible that the lack of neurobiological change may be due to the short space between the two assessment intervals (approximately two to three months) (Johnco, Wuthrich, & Rapee, 2014). For this reason, it is recommended to include follow-up assessments in future research to verify whether neurobiological changes occur on the long-term.

In sum, it appears that we are still far away from a thorough understanding of why, how, and for whom correctional interventions are effective. Nevertheless, the current study guided recommendations for future research (i.e., reconsideration of neurobiological measurements) and for forensic practice. Results revealed that correctional rehabilitation might benefit from modifying treatment selection procedures (i.e., including verbal intelligence assessment), treatment content and duration, as well as enrichment of the prison setting itself. Unfortunately, limited manpower, time and financial resources probably complicate the incorporation of the suggested adaptations. However, there are alternative options that are worthwhile to explore, such as the usage of virtual reality (VR). In the Netherlands, VR is already used in forensic special care hospitals to train patients to control aggressive impulses in real life situations. VR has the potential to

https://decorrespondent.nl/2562/Gaat-Virtual-Reality-ons-betere-mensen-maken-/262656240-64da7127
increase treatment intensity/duration by enhancing prisoners' cognitive skills in their own prison cells supplementary to group therapy meetings. Furthermore, VR training programs could be tailored towards the inmate's individual neurocognitive needs/deficits (including verbal fluency impairments), and could also enrich the prison setting by exposing inmates to social challenges or real life situations. The correctional field might benefit from innovative developments like these, along with the implementation of renewed rehabilitation programs (e.g., CoVa 2.0), in an attempt to effectively enhance (neuro) cognitive deficits and, eventually, reduce recidivism rates.
References


Buysse, W., & Loef, L. (2012). Effectiveness of the cognitive skills treatment (CoVa) for offenders. Amsterdam: DSP-groep.


Chapter 5


Supplementary Material

Multiple Imputation Procedure

The goal of multiple imputation (MI) is to create several complete data sets, by replacing missing values with plausible data values (Rubin, 1987; Van Buuren, 2012). One assumption for MI is that missing values are missing completely at random, or at least at random. Missing value analysis was conducted to investigate the pattern of missing data. For both groups, Little's Missing Completely At Random (MCAR) test was not significant (intervention group: $\chi^2(8140) = 5643.428, p = 1.000$; control group: $\chi^2(3359) = 1320.984, p = 1.000$), indicating that there was no particular pattern of missing data.

The specific MI model used was the switching regression approach by Van Buuren, Boshuizen, and Knook (1999). The MI model was built up by including pre-test scores, post-test scores, potential confounders (e.g., physical activity) and behavioral outcome measures. Derived variables, such as the Stroop interference score, were not imputed directly, but underlying scores were imputed (time to complete Stroop II and Stroop III) and composed to the derived variable after imputation. Imputations were conducted separately for the intervention group and the control group. After imputation, the two datasets were combined using SPSS 19 for further analyses.

The number ($m$) of imputed datasets was set at 10. In addition, the number of iterations is usually between 5 and 20 and we decided to set this at 20 (Van Buuren, 2012). Predictive Mean Matching (PMM) option was used since this imputation method provides realistic values and has been found to produce the least biased estimates and better model performance measures (Marshall, Altman, & Holder, 2010). The statistical program R version 3.1.1 with library mice version 2.22 was used to complete the imputation model while performing a pre-selection of covariates predictive of missingness. Covariates with a correlation of less than .2 with the predicted covariate were excluded from the corresponding prediction equations to improve numerical stability. The average number of predictors per variable was 14.87 for the intervention group and 12.37 for the control group.

After imputation, post-test values (e.g., self-evaluation and post-test scores on neurocognitive tasks) for participants who dropped out of intervention ($N = 35$) were omitted from further analyses. This method is an acceptable and practical solution for dealing with imputed values that are considered to be ‘invalid’ (personal communication Van Buuren, 16-2-2015).
The combination of F-statistics from a multiple imputation in the case of finite sample sizes requires corrected degrees of freedom, as suggested by Reiter (2007). In order to apply this, a software implementation written in SPSS MACRO by Van Ginkel (2008) was used.

To perform ANCOVA analyses, mimul2 package was employed for combining multivariate estimates in multiple imputation (Van Ginkel, 2014). In addition, chi-square tests were combined as described by Li, Meng, Raghunathan, and Rubin (1991) and standard deviations were computed using Ruben’s rules.

### Selected Neurocognitive Tasks

<table>
<thead>
<tr>
<th>Neurocognitive task</th>
<th>Assesses</th>
<th>Goal</th>
<th>Dependent variables</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop Color Word Task</td>
<td>Interference control</td>
<td>To inhibit an automatic response (reading words) in order to execute a more controlled process (naming colors)</td>
<td>Interference score in seconds</td>
<td>(Hammes, 1978; Stroop, 1935)</td>
</tr>
<tr>
<td>Concept Shifting Test (CST)</td>
<td>Cognitive flexibility</td>
<td>To cross out numbers, letters and combined numbers/letters as quickly as possible in correct order</td>
<td>Shifting score in seconds</td>
<td>(van der Est, van Boxtel, van Breukelen, &amp; Jolles, 2006)</td>
</tr>
<tr>
<td>D2 Cancellation Task</td>
<td>Selective attention and concentration</td>
<td>To mark as many targets (’d’ with two marks) as possible within a certain time frame while ignoring distractors</td>
<td>Correct targets minus mistakes (CP)</td>
<td>(Brickenkamp, 2007)</td>
</tr>
<tr>
<td>WAIS-II Digit Span Backward</td>
<td>Immediate short-term memory</td>
<td>To memorize and recall as many digit spans back-to-front as possible</td>
<td>Total correctly produced digit spans back-to-front</td>
<td>(Wechsler, 2000)</td>
</tr>
<tr>
<td>Modified Wisconsin Card Sorting Task (M-WCST)</td>
<td>The ability to alternate between response sets and measure of response modulation</td>
<td>To sort a pack of cards according to a changing rule pattern provided by experimenter’s feedback</td>
<td>Total achieved categories, perseverative errors</td>
<td>(Nelson, 1976)</td>
</tr>
<tr>
<td>Tower of Hanoi (TOH)</td>
<td>Planning skills</td>
<td>To use a certain number of moves to arrange discs from initial position into a goal position while complying to specific rules</td>
<td>Planning time in seconds, errors</td>
<td>(Simon, 1975)</td>
</tr>
<tr>
<td>Stop it Task</td>
<td>Inhibitory control</td>
<td>To stop an ongoing behavior (responding to figures) in response to a beeping signal</td>
<td>Stop signal reaction time (SSRT)</td>
<td>(Verbruggen, Logan, &amp; Stevens, 2008)</td>
</tr>
<tr>
<td>Reading the Mind in the Eyes (child version)</td>
<td>Social empathy</td>
<td>To select one of four words that best describes a person’s thoughts or feelings</td>
<td>Total correct</td>
<td>(Baron-Cohen, Jolliffe, Mortimore, &amp; Robertson, 1997; Baron-Cohen, Wheelwright, Hill, Raste, &amp; Plumb, 2001)</td>
</tr>
<tr>
<td>Controlled Oral Word Association Task (COWAT)</td>
<td>Verbal fluency</td>
<td>To generate orally as many words as possible beginning with letters D, A and T within one minute, while adhering to several rules</td>
<td>Total correctly named words</td>
<td>(Borkowski, Benton, &amp; Spreen, 1967)</td>
</tr>
</tbody>
</table>
References


