THE JOURNAL OF FORENSIC PSYCHIATRY & PSYCHOLOGY
https://doi.org/10.1080/14789949.2022.2070523

Can measures of cognitive flexibility and inhibition distinguish forensic psychiatric inpatients from prisoners?

P.W Molleman[a], L. Molleman[b,c], C. Schilder[a], B.H. Bulten[a,d] and I.A. Brazil[a,e,f]

[a]Forensic Psychiatric Centre Pompestichting, Nijmegen, The Netherlands; [b]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands; [c]Department of Social Psychology, Tilburg University, Tilburg, The Netherlands; [d]Behavioral Science Institute, Radboud University, Nijmegen, The Netherlands; [e]Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, The Netherlands; [f]Collaborative Antwerp Psychiatric Research Institute (CAPRI), Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium

ABSTRACT

We investigated whether forensic psychiatric inpatients can be distinguished from prisoners and healthy controls on the basis of their performance on cognitive tasks measuring cognitive flexibility (with a reversal learning task) and inhibition (with a stop signal task). Forensic psychiatric inpatients were expected to perform worse compared to prisoners. This study was based on pre-existing data from N = 241 males (119 forensic psychiatric inpatients, 57 prisoners and 65 healthy controls). We fitted logistic generalized linear models to group membership data using outcome measures from a stop signal task and a reversal learning task to examine whether deficiencies in inhibition and cognitive flexibility predict membership of the group of forensic psychiatric inpatients. Neurocognitive measures of cognitive flexibility, but not inhibition, had predictive value for group membership. This suggests that the capacity for cognitive flexibility may help differentiate between forensic subpopulations and healthy controls. Furthermore, our results highlighted substantial heterogeneity in cognitive performance in inhibition and cognitive flexibility within different offender groups. This study indicates that forensic subpopulations can be adequately differentiated based on neurocognitive measures of cognitive flexibility. In a broad sense, this study is an example of how neurocognitive approaches can contribute to refining diagnosis in forensic psychiatry.

ARTICLE HISTORY Received 3 December 2020; Accepted 21 April 2022

KEYWORDS Forensic psychiatry; offenders; executive functioning; inhibition; behavioral adjustment

Introduction

The emphasis in general psychiatry is currently undergoing a shift away from a single descriptive view at the syndrome level, and towards an integrative view that incorporates biological and cognitive information at the symptom level

CONTACT P.W Molleman p.molleman@pompestichting.nl Nijmegen, The Netherlands
Supplemental data for this article can be accessed online at https://doi.org/10.1080/14789949.2022.2070523.

(Owen, 2014). The premise is that insights into the neurocognitive correlates of symptoms help unearth their mechanistic underpinnings and will ultimately allow for direct treatment at those latent levels (Brazil et al., 2018; Schwartz et al., 2013). The transdiagnostic thinking that underlies this development is already present in forensic psychiatry (Martell, 1992; Schultz, 2010), as risk factors are embedded in general transdiagnostic concepts. However, the extension of this thinking to incorporate neurocognitive factors is relatively new (Buckholtz & Meyer-Lindenberg, 2012).

The transdiagnostic view offers many advantages for forensic psychiatry because individual forensic psychiatric inpatients exhibit many similarities at the behavioral level but can differ greatly in their underlying cognitive, neurobiological and personality disturbance (Spaans, Molendijk, de Beurs, Rinne, & Spinhoven, 2016). One promising transdiagnostic approach stems from the application of tools and knowledge from cognitive neuroscience in forensic psychiatry, which has led to a shift in research focus towards understanding the cognitive and neurobiological mechanisms that underlie antisocial behavior (Blair, 2013; Brazil et al., 2018). An increasing number of studies suggest that deficiencies in neuropsychological processes play an important role in the development and persistence of delinquent, antisocial and violent behavior (e.g., Fishbein et al., 2009; Kogel de, 2008).

Despite these promising results, there is a lack of studies examining whether measurements of neuropsychological functioning can help distinguish forensic psychiatric inpatients from other delinquent populations. Distinguishing forensic psychiatric inpatients from other prisoner populations is crucial for daily practice, because it can help identify underlying neurocognitive problems and devise effective treatment targeted at deficiencies at these underlying levels. Forensic psychiatric inpatients comprise individuals who have committed serious offenses and are characterized by diminished accountability at the time of the offense. Understanding neuropsychological capacities of offender groups can help design more effective treatments and reduce recidivism, as well as inform rehabilitation programs and risk and release decisions (Shumlich, Reid, Hancock, & Hoaken, 2018). In general, a deeper understanding of the antisocial behavior of individuals from any prisoner group is useful as this behavior causes substantial harm to society (Gielen & Akkermans, 2018).

It has been proposed that antisocial and criminal behaviors can be induced by deficiencies in specific executive functions (EFs) (Portnoy et al., 2013). EF encompasses higher-order cognitive functions – such as behavioral adjustment and inhibition, which facilitate goal-directed behavior and promote self-control (Hofmann et al., 2012). Comprehensive meta-analyses have demonstrated robust associations between antisocial behavior and deficiencies in EF, as indicated by performance errors in neuropsychological measures of response behavioral adjustment and inhibition (Morgan & Lilienfeld, 2000;

Ogilvie et al., 2011). Specifically, deficiencies in response inhibition and behavioral regulation appear to be consistently related to crime, delinquency, physical aggression and behavioral maladjustment (Morgan & Lilienfeld, 2000). Correspondingly, it has been suggested that well-developed EFs serve as a protective factor against antisocial and criminal behavior (Portnoy et al., 2013). In addition to criminality, the most common diagnoses among forensic psychiatric inpatients, i.e., schizophrenia and other psychotic disorders (Goossens et al., 2019), are associated with executive dysfunctions (Barch, 2005).

In this paper, we focus on two specific aspects of EF: Inhibition and cognitive flexibility. Inhibition can be defined as the ability to suppress inappropriate actions (Van Toor et al., 2010). Whether or not an individual engages in antisocial and criminal behavior depends to an important extent on their capacity to inhibit behavior, and inhibition deficiencies are considered important drivers of antisocial behavior in forensic groups (Kerr, 1997). Evidence from a study conducted in Dutch prisons suggests that inhibition deficiencies differ between violent and non-violent offenders, and are predictive of treatment outcomes during detention (Meijers et al., 2017). Moreover, prisoners with inhibition deficiencies have been shown to make less progress in detention programs and are more likely to stop treatment prematurely (Fishbein et al., 2009). Yet, there is little research considering a possible lack of inhibition as a dysfunctional neuropsychological dimension in forensic psychiatric inpatients. There is one prior study comparing forensic psychiatric inpatients and prisoners on inhibition (Shumlich et al., 2019). Based on two tasks that measured inhibition, this study found that, a higher fraction of forensic psychiatric inpatients performed worse than correctional offenders without a mental disorder.

Inhibition often helps individuals to flexibly adjust their behavior towards their goals in an ever-changing environment, and inhibition deficiencies often lead to behavior that is no longer necessary or appropriate (Verbruggen & Logan, 2008). Behavioral adjustment to such dynamic environments can be studied experimentally using 'reversal learning' tasks, in which participants are prompted to adjust their behavior after an unexpected punishment or reward by inhibiting previously learned contingencies (Cools et al., 2006; Frank et al., 2004). Many patients in forensic psychiatric settings exhibit maladaptive behavior in such tasks and have difficulties with reversal learning (I. Brazil et al., 2013; De Brito et al., 2013; Gregory et al., 2015). These results are in line with the results of other studies that found difficulties in using information about negative outcomes to change behavior across many subpopulations of antisocial adults (Brazil et al., 2009; I. Brazil et al., 2013; Budhani et al., 2006; M. A. Dargis et al., 2017; De Brito et al., 2013; Dillien, Goethals, Sabbe, & Brazil, 2018; Gregory et al., 2015; Newman, Patterson, & Kosson, 1987; Von Borries, Bulten, & Verkes, 2015). This kind of perseveration

of behavior likely reflects a form of deficient cognitive flexibility, in which an individual is unable to inhibit previously rewarded behavior that no longer yields positive outcomes.

Different groups of offenders show different types of deficiencies in cognitive flexibility and inhibition, both on the behavioral (I. Brazil et al., 2013; De Brito et al., 2013; Budhani et al., 2006) and on the neurobiological levels (e.g., Gregory et al., 2015). For example, results from an fMRI study suggest that antisocial offenders with and without psychopathy differ in how the brain processes information after a change in contingencies has taken place (Gregory et al., 2015). However, this research has predominantly focused on comparing offenders with and without psychopathy, and it remains unclear whether deficiencies in cognitive flexibility and inhibition are characteristic of offenders in general or differ between forensic psychiatric inpatients and prisoners. Shumlich et al. (2019) suggest that forensic psychiatric inpatients have a decreased inhibition compared to offenders without a mental disorder. It remains unclear whether the level of deficiency in cognitive flexibility and inhibition can differentiate between forensic psychiatric inpatients and prisoners due to the presence of psychopathology. By definition, forensic psychiatric inpatients engaged in criminal behavior and have a mental disorder (at least at the time they committed the crime they were sentenced for). Mental disorders are robustly associated with deficiencies in flexibility (M. A. Dargis et al., 2017) and inhibition (Nigg, 2016).

The aim of this explorative study is to investigate whether forensic psychiatric inpatients can be distinguished from prisoners and healthy controls on the basis of their performance on cognitive tasks measuring cognitive flexibility and inhibition. Based on the literature, we predicted that both offender groups would perform worse than non-offender controls. In addition, we examined whether there is a difference between the offender groups with respect to cognitive flexibility and inhibition. Because forensic psychiatric inpatients by definition have a mental disorder – or at least had a mental disorder at the time of committing the crime they were convicted for – and mental disorders are associated with executive dysfunction, we expected decreased cognitive flexibility and inhibition in forensic psychiatric inpatients compared to prisoners.

Methods

Participants

The current study was based on pre-existing data from 120 male forensic psychiatric inpatients, 60 male prisoners and 66 male healthy non-offender controls (Von Borries et al., 2015). Forensic psychiatric inpatients were recruited in the Pompekliniek and Oldenkotte, two forensic psychiatric hospitals, or 'TBS-clinics[1]', in the Netherlands. Data on the prisoners were collected at Vught

Penitentiary Institution, the Netherlands. The sample (N = 241) consisted of 119 'TBS'-patients or forensic psychiatric inpatients (M_{age} = 38.9, SD = 9.2), 57 prisoners (M_{age} = 31.2, SD = 8.5) and 65 healthy controls (M_{age} = 36.3, SD = 11.8; see statistical design for the procedure and the Supplement for details).

TBS is a treatment measure in Dutch criminal law that is generally imposed in combination with a prison sentence. Judges can impose TBS on a suspect for a crime that is punishable by at least four years of imprisonment. A key condition for imposing TBS is that the suspect suffered from a mental disorder at the time of the offense. Convicts can start receiving treatment in a TBS hospital after serving at least one-third of their sentence. In contrast to a prison sentence, which covers a broad spectrum of crimes ranging from severe crimes to minor offenses, TBS offenders have generally committed a more serious crime. In addition, unlike prisoners, severe psychopathology had been diagnosed in all TBS patients. The TBS status can be prolonged until the risk of reoffending is deemed to have reduced to acceptable levels due to successful treatment and risk management. In Dutch criminal law, the court imposes TBS, a prison sentence or a combination of both. This decision is taken by the court on the basis of a number of criteria: a serious offense (punishable with at least 4 years in prison), diminished accountability at the time of the offense due to a 'deficient development or morbid disorder', such as a personality disorder, psychosis or an intellectual disability, and a high probability of recidivism. TBS offenders meet at least these three criteria, and prisoners generally do not.

For the recruitment of the forensic psychiatric patients, clinicians were asked whether selected patients meeting the inclusion criteria could be contacted for the study. If so, these patients were provided with information about the study and asked whether they would like to participate. If the patient agreed to participate, three meetings of two hours each were planned to conduct the experiment. Within the meeting, a break was planned and participants could indicate whether they needed additional breaks. Prisoners were recruited in a similar manner by contacting staff members beforehand. Controls were recruited directly within the forensic clinic among staff members and other citizens of Nijmegen and its surroundings. For control, a subset of tasks was administered, which took about four hours in total. Therefore, data collection in controls was done within one session including three short breaks. Controls had no education level higher than Higher Vocational Education (in Dutch: HBO) and no criminal history.

Data collection procedures

The test battery was administered in a fixed order, testing sessions never took longer than 2 hours . The entire test battery was collected within 2 weeks. During the sessions, the researcher remained in the room with the subject to answer any questions, but without having a direct view of the answers to be

given by the subject. The tasks were performed on an HP Compaq 6735s Notebook PC with a 15.4-inch display. Responses were recorded using either a keyboard, joystick or mouse. Participants received a financial compensation for participation. Anonymization was achieved by giving each subject a numerical code. All information and test data were stored under this code on an external, password-protected hard drive. Only the principal investigator has access to a file in which the subject number is linked to their name. There were some missing data because some offenders quit the task in the middle of a session.

The study was conducted in accordance with the principles of the Declaration of Helsinki (Medical Association (WMA) & World Medical Association (WMA), 2009) and in accordance with the Medical Research Involving Human participants

Act (WMO). The written information about the study was handed out to prospective participants prior to the start. Prospective participants had a week to consider whether or not to take part, and to give informed consent. The pre-existing study ('Automatic Neurocognitive Assessment in forensic contexts'; CMO case number: 2008/248, ABR number: NL24614.091.08) was approved by the Research Ethics Committee, CMO region Arnhem-Nijmegen, The Netherlands, and is part of a large study regarding the underlying neurocognitive mechanisms of criminal behavior (Von Borries et al., 2015).

Instruments

We used tasks from the Forensic Mental Information Processing and Neuropsychological Diagnostic System (ForMINDS; Von Borries et al., 2015) to measure participants' capacities for cognitive flexibility and inhibition. ForMINDS was developed by the Pompestichting in collaboration with the Utrecht University and Radboud University, and consisted of both questionnaires and computerized cognitive tasks that participants completed on their own. In this study, we used the reversal learning task to measure participants' capacity for cognitive flexibility (Cools et al., 2006), and the stop signal task to measure capacity for inhibition (Logan & Cowan, 1984; Verbruggen et al., 2008).

Reversal learning task

The reversal learning task can be used to determine the extent to which an individual is able to learn about, and adjust to, changes in stimulus-outcome contingencies. This task measures whether a person can adjust their behavior after an unexpected punishment or after an unexpected reward. Because of the use of stimulus-outcome learning (instead of stimulus-response learning) in combination with the reversal aspect of this task (for example, stimulus A is first followed by expected punishment and later by unexpected reward, while

for stimulus B the opposite applies), cognitive flexibility can be measured after both an unexpected punishment and an unexpected reward have occurred (Cools et al., 2006).

The task included a series of stimuli followed by unexpected rewards or unexpected punishments (counterbalanced across participants), each containing a practice block, and two experimental blocks. The practice block was followed by a reversal stage after the participant had reached the initial learning criterion of 10 correct trials. The task was terminated if this criterion was not met within 80 trials. There were 2 blocks, with 120 trials per block, so that each subject performed 4 blocks and 480 trials. A schematic of sample trail-sequences for each condition is shown in Figure S1.

Participants had to imagine they were the manager of a casino watching along with a card game via a camera in their office. This card game consisted of only two cards: the joker and the ace of spades. The participant required to predict whether, based on previous results, the imaginary player would win or lose drawing a card. The winning card changed from time to time (the 'reversal' event) and this change depended on the predictions of the participant: after achieving the learning criterion of 5–9 consecutive correct responses, a reversal of the contingencies followed and the card previously yielding a reward started to lead to punishment (and vice versa; Von Borries et al., 2015). Participants gave their responses with the left- and right-cursor keys. Participants had to select the left button if he predicted that the imaginary player would lose and select the right button when he predicted that the player would win. Participants had no influence over the outcome; the feedback given after the participants had predicted whether a win or a loss would follow did not relate to the correctness of the prediction but merely indicated whether the card was winning or losing.

In this study, we characterize a participant's capacity for cognitive flexibility with two performance measures: (a) as the proportion/mean number of switch errors per stage (= the number of switch errors divided by the number of stages) and (b) the number of stages (number of 5–9 consecutive correct trials) reached in the experimental blocks (the more stages the learning criterion was met for, the more adaptive the participant's behavior Cools et al., 2006).

Stop signal task

Response inhibition was measured with the stop signal task (Logan & Cowan, 1984; Verbruggen et al., 2008). Participants were instructed to press the left-cursor key when the arrow on their computer screen pointed left and press the right-cursor key when the arrow pointed right as quickly as possible, but inhibit their response when they see a STOP-signal and heard a sound (of a bell) just after the appearance of the arrow on the screen (Figure S2).

The task consisted of 120 GO trials during which participants had to press a key with the left- or right-index finger, and of 120 Stop trials in which participants had to withhold from responding. In Stop trials, a visual and acoustic STOP-signal was presented shortly after the Go-stimulus, and participants had to suppress their tendency to respond (Figure S2; Logan & Cowan, 1984). The Go-stimuli were presented sequentially in the center of the screen for 750 ms against a gray background.

Outcome measures for GO and STOP trials per block were the average reaction times and the corresponding standard deviations, the percentage of choice errors and omissions in GO trials, and the percentage of correct inhibitions in STOP trials (Von Borries et al., 2015).

Since there is no observable endpoint for the response inhibition to be completed, a model for estimating the finishing point of the stop process was used to characterize our two performance measures for a participant's capacity for inhibition: (*i*) reaction times (in ms) to the stop signal averaged across all trials and (*ii*) the number of inhibition errors (Logan et al., 1984).

Statistical design

We sought to find out whether cognitive flexibility and inhibition measures predicted group membership ('prisoners' vs. 'TBS' vs. 'healthy controls'). The predictive value means that the measure is significantly related to group membership, in other words that groups differ significantly on that measure. To this end, we fitted logistic generalized linear models (GLMs) with pairwise group comparisons to group membership data (0 = prisoners; 1 = TBS; 2 = healthy controls) using outcome measures from the stop signal task (mean reaction time and number of errors) and the reversal learning task (number of switch errors and number of stages). Our *a priori* predictions for the performance on each of the measures were directional (TBS < prisoners < healthy controls). We therefore used one-tailed tests to examine group differences. Note that in the results section, reported confidence intervals of odds ratios are two-sided, while p-values are based on one-sided tests. This means that when effects are small, the 95% CI of odds ratios may include 1, while the p-value is lower than 0.05. We also included the demographic variables 'age', 'IQ' and 'education level' as predictors, each of which is known to influence neurocognitive functioning and may therefore partly explain group differences (Stern, 2006). IQ was measured with the Dutch Adult Reading Test (Schmand et al., 1991). We used the Dutch Verhage scale to classify the level of education (Verhage, 1964).

We performed the following outlier removal procedure: outliers were excluded when scores had a deviation of more than 2 standard deviations (Van Selst & Jolicoeur, 1994). This procedure led to the exclusion of 4 participants (1 forensic psychiatric inpatient and 3 prisoners). We then ran logistic

generalized linear models (GLMs) fitted to group membership (TBS, prisoner or healthy controls) for each of the predictors separately (see Tables S3-S6 in Supplementary Information). Tables S3-S6 present statistical analyses based on logistic regressions with 'group membership' as dependent variable. For ease of exposition, each model contrasts two groups. We did not fit a model using all predictors simultaneously due to missingness of data (see Table S1), and the multicollinearity between the outcome measures on the reversal learning task (number of stages and mean number/proportion of switch errors per stage; see Table S2).

Results

Here, we summarize the distributions of the neurocognitive measures of participants' capacities for cognitive flexibility (from the reversal learning task) and inhibition (from the stop signal task). For both offender groups and healthy controls, we evaluate the predictive value of each of these measures using logistic GLMs with pairwise group comparisons fitted to group membership ('TBS' vs. 'prisoners' vs. 'healthy controls').

Figure 1 shows the distribution of the four outcome variables of our study broken down by comparison group. In the reversal learning task on the number of stages, relative to healthy controls, we observe worse performance among the offender groups, especially in the group of forensic psychiatric inpatients (Figure 1A). In line with our hypothesis, healthy controls *(M* = 26.86, *s* = 13.01) achieved more stages than prisoners *(M* = 22.41, *s* = 12.38), and prisoners reached more stages than forensic psychiatric inpatients (M = 17.64, *s* = 15.11). GLMs fitted to pairs of groups confirmed that healthy controls completed significantly more stadia than forensic psychiatric inpatients (β = –0.045, *p* = 0.001, *OR* = 0.961 (95% CI:.932, .992)) and prisoners (β = –0.028, *p* = 0.047, *OR* = 0.963 (95% CI:.939, .999); see Table S3 for details). However, the predictive value of the number of completed stages disappeared after controlling for demographics (see Table S3). Also, the GLM indicated a significant difference between the offender groups on the number of completed stages (GLM: β = –0.024, *p* = 0.028, *OR* = 1.006 (95% CI:.979, 1.034); see Table S3). Importantly, after controlling for demographics, the model indicates that forensic psychiatric inpatients completed significantly fewer stages than prisoners did (GLM: β = –0.047, *p* = 0.005, *OR* = 1.048 (95% CI:1.012, 1.085)).

In the reversal learning task, forensic psychiatric inpatients made, on average, more switch errors per stage *(M* = 0.56, *s* = 0.93) in comparison to prisoners (M = 0.43, *s* = 0.82) and controls (M = 0.25, *s* = 0.35; Figure 1B). GLMs revealed significant differences between forensic psychiatric inpatients and controls (GLM: β = 0.789, *p* = 0.037, *OR* = 2.201 (95% CI:.931, 5.207)), but there was no significant difference when controlling for demographics (see Table S4). The GLM indicated no significant differences between the offender groups on the

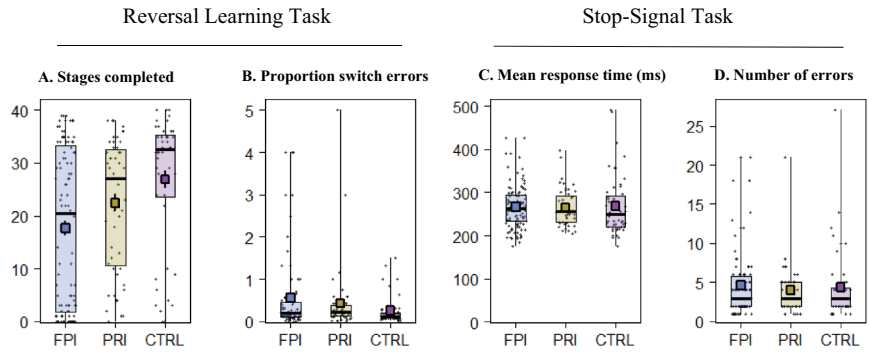

Figure 1. Distributions of neurocognitive measurements of capacities for cognitive flexibility and inhibition across forensic psychiatric inpatients (FPI), prisoners and healthy controls. Plots (A-D) show distributions of individual measures broken down by group (FPI= blue, prisoners = yellow, healthy controls = purple). Dots represent individual participants; boxplots show interquartile ranges with medians in thick lines. Squares and whiskers show means ± 1 SEM. See Table S1 for descriptive statistics.

proportion/mean number of switch errors per stage (see Table S4). However, after controlling for demographics there was a significant difference between the offender groups (GLM: $\beta = 1.481$, $p = 0.038$, $OR = 0.227$ (95% CI:.044, 1.163)).

Taken together, these results indicate that cognitive flexibility, as measured by the number of completed stages and the proportion/mean number of switch errors per stage in the reversal learning task, has predictive value in distinguishing the offender populations from controls and between forensic subpopulations. For both forensic psychiatric inpatients and prisoners and healthy controls, we observed a wide distribution of completed stages and switching errors in the reversal learning task, suggesting strong within-group heterogeneity in the ability to adjust to changing circumstances.

The stop signal task did not reveal significant neurocognitive predictors for membership of the offender group or healthy controls. Mean response times in this task did not predict group membership (Figure 1C, Table S1 and S5). The number of errors did not significantly predict group membership either (see, Figure 1D; Table S1; Table S6).

These results indicate that inhibition, as measured by response time and number of errors in the stop signal task, had no predictive value for distinguishing between forensic psychiatric inpatients, prisoners and healthy controls. Again, all groups showed a wide distribution of outcome variables, suggesting strong within-group heterogeneity in their ability to adjust to changing circumstances.

Analysis of demographic variables indicated that the forensic psychiatric inpatients were slightly older than the prisoners, but the offender groups were similar on other potentially relevant demographic variables (IQ and education level; see Table S1).

Discussion

In the present study, we evaluated the predictive power of cognitive flexibility and inhibition for distinguishing forensic psychiatric inpatients from prisoners and healthy controls. Our results indicated more reversal learning problems in offenders compared to non-offenders. Also, our findings revealed that our neurocognitive measures of cognitive flexibility (in a reversal learning task), but not inhibition (in a stop signal task), had predictive value for group membership. Therefore, our results provide support for our directional hypothesis (FPI < prisoners < healthy controls) regarding cognitive flexibility but do not support our hypotheses regarding inhibition.

The predictive power of our outcome measures in the reversal learning task (see, Figure 1 and Tables S3-S4) supports our hypothesis of decreased cognitive flexibility in forensic psychiatric inpatients compared to prisoners. Therefore, our findings suggest that deficiencies in cognitive flexibility significantly predict group membership and that reversal learning tasks may also help differentiate between forensic subpopulations. This is in line with the study by M. A. Dargis et al. (2017), which indicated that mental disorders are robustly associated with deficiencies in cognitive flexibility. We acknowledge that the predictive power of our neurocognitive measures is relatively weak (see Tables S3-S6); age, education and IQ seem to have higher predictive power in most cases, which was not unexpected (Selby, Airy-Eggertsen, & Laver, 1997). In the reversal learning task, healthy controls progressed through significantly more stages than forensic psychiatric inpatients. Our results suggest that forensic psychiatric inpatients and prisoners show more reversal learning problems in comparison to healthy controls. Also, forensic psychiatric inpatients and prisoners had a smaller number of completed stages and a higher proportion of switch errors per stage compared to healthy controls. However, there was no significant effect after controlling for demographics on both measures. Taken together, these results suggest that neurocognitive measures of cognitive flexibility may not help distinguish offenders from non-offenders (see Tables S3-4).

The present findings indicate that the inhibition measures obtained with the stop signal task may not be adequate for differentiating among offender populations. We predicted diminished inhibition in offender groups compared to healthy controls and in forensic psychiatric inpatients compared to prisoners, because of the association between mental disorders and

executive dysfunction. Our hypotheses were not supported; the outcome measures indexing inhibition did not appear to be significant neurocognitive predictors for membership in the offender group or healthy controls.

Our results showed a substantial amount of heterogeneity within each group: some forensic psychiatric inpatients had well-developed inhibition capacities, while some prisoners had major inhibition deficiencies. Such findings suggest that the group means of the outcome variables are not representative of individuals within the groups because there is a wide distribution. The broad distribution causes an interesting pattern in the data; group differences are caused by demographics or arise mainly due to different proportions of participants who perform extremely poorly. This may explain why our findings do not converge with those of a recent study reporting that forensic psychiatric inpatients displayed poor behavioral inhibition compared to correctional offenders (without a mental disorder; Shumlich et al., 2019). This difference may stem from the fact that Shumlich et al. (2019) compared group frequencies of individuals with versus without EF impairments (relying on cut-off levels rather than exploiting all measured variations in performance). Our results for the stop signal task suggest that offender groups do not differ from each other on the measures of inhibition, but also highlight a large amount of individual variability with regard to inhibition (see, Figure 1 and S1 and Tables S5-6). The heterogeneity in our results highlights the importance of developing tailored treatment interventions for each individual case. In addition, it is possible that real differences between the offender groups are partly masked by differences in the environments in which they were detained. TBS patients are exposed to a relatively less deprived environment compared to regular prison, and the treatment offered in forensic psychiatric clinics may have a positive influence on inhibitory control. This might partly explain why no differences in inhibition were found between the offender groups. It is possible that this positive effect may be smaller for regular prisoners, as Meijers et al. (2018) found that three months of imprisonment in an impoverished environment can lead to reduced self-control.

One limitation of the present study is that we only tested differences between groups on a small set of neurocognitive measures. It is advisable to use a broader test battery to measure cognitive flexibility and inhibition in the future. In the larger ForMINDS study, it was not assessed whether performance on the neuropsychological tests employed reflects offenders' true level of neurocognitive functioning, or if performance was impacted by the patients' directed efforts to over-engage in the tasks. So, we have to acknowledge the potential impact of performance validity on the accuracy of the test scores. Also, there is growing evidence suggesting that reaction times and accuracy measures obtained using a stop signal task can only scratch the surface when it comes to estimating inhibition capacity (Bissett & Logan,

2011). Thus, for a more complete picture of inhibitory capacities, our outcome measures should ideally be complemented by alternative and informative data processing (Bisset & Logan, 2012a, 2012b).

Moreover, motivational processes may also contribute to antisocial behavior in addition to cognitive flexibility and inhibition. This aspect was not included in the used tasks but might be of value for distinguishing forensic psychiatric patients from prisoners. A questionnaire that measures motivational drives is the Behavioral Inhibition System/Behavioral Activation System (BIS/BAS) Scale. The BIS/BAS scale is a 20-item self-report questionnaire that contains four scales that measure behavioral activation or inhibition (Carver & White, 1994). The BIS scale assesses sensitivity to punishment and the BAS scales reflect three interrelated types of behavioral activation (drive, fun seeking and reward responsiveness). Hoppenbrouwers et al. (2015) found a positive link between antisocial behavior and the Behavioral Activation System (BAS) in offenders: Impulsive and stimulation-seeking motivational behavior was linked with BAS activation. Indeed, excessive BAS activation has frequently been linked to impulsivity (i.e., Wallace et al., 2009) and aggression (Wingrove & Bond, 1998). Moreover, a recent study (Molleman et al., 2021) showed that the BIS/BAS self-report questionnaire seemed to dissociate between forensic psychiatric inpatients and prisoners. One next step would be to examine if the BIS/BAS also captured part of the group differences in the likelihood of recidivism or treatment responsivity (Molleman et al., 2021).

Another issue is that we had no access to potentially important information from the prisoner group, like the prevalence of psychopathology and the type and severity of the offenses. As these factors are known to be negatively associated with capacities for cognitive flexibility and inhibition, including them could have helped explain why we did not find differences between the groups on all our outcome measures and why the groups varied so much internally on our main outcome measures. It is possible that the large amount of heterogeneity is related to diagnostic status and/or charge-specific factors. It is likely that neurocognitive function differs by diagnosis. For future research, it would be interesting to understand whether the within-group heterogeneity among forensic psychiatric inpatients might be associated with different personality disorders and/or other psychiatric diagnoses, as well as prescribed psychotropic medication. Identifying such associations might be valuable in understanding and treating forensic psychiatric inpatients in jurisdictions outside of the Netherlands.

Yet another limitation is that we do not know whether the participants used psychotropic medication during measurements of the inhibition tasks used. This can make the group differences appear smaller on these tasks, and it is plausible that the reported findings are largely explained by this

confounding factor. Psychotropic medication can make forensic psychiatric inpatients less impulsive. There are also different types of psychotropic medication that have different effects on neurocognitive functioning. Future studies should investigate the effects of different types of psychotropic medication on neurocognitive functioning in forensic psychiatric inpatients.

Another limitation in our study, is the lack of a fourth group consisting of non-forensic psychiatric patients. Future research should compare the forensic psychiatric population with a civil psychiatric population in order to identify unique executive dysfunction in individuals with mental disorders who commit crime. Indeed, absent a psychiatric comparison group, our data do not rule out the possibility that any observed neuropsychological differences resulted from the specific underlying psychiatric limitations and not the forensic characteristics of these (psychiatric) participants. For future research, it is important to include a non-offender psychiatric comparison group to establish whether there are differences in neurocognitive performance when compared with the two offender groups. The results of such a study might help identify risk factors for mentally ill people at risk of committing severe crimes.

Despite having matched the groups on age and sex, we acknowledge that sampling bias could partly explain the group differences found. For example, we had some missing data due to participants quitting the task in the middle of a session. This could be related to a sampling bias because these missing data were from the offender groups. As a result, the offenders' inhibition capacity and cognitive flexibility may have been overestimated, because the most inhibited offenders may not have been included in the sample. It should also be kept in mind that some of our null-findings could be related to the size of the included samples, as some effects may require even larger sample sizes to be detected reliably.

Nevertheless, our results suggest that inhibition and behavior regulation deficiencies are present to various degrees in both the forensic psychiatric patients and prisoners. It is important to investigate whether neurocognitive functions predict real-world outcome measures, such as recidivism or violent incidents, in follow-up studies.

In sum, neurocognitive measures of participants' cognitive flexibility had significant predictive power for distinguishing forensic psychiatric inpatients from prisoners and differentiate offenders from non-offenders. Moreover, our study highlighted that there is a large amount of heterogeneity in cognitive performance when it comes to cognitive flexibility and inhibition within different offender groups. The presence of such a large amount of heterogeneity suggests that it may be important to consider variations in neurocognitive characteristics in order to develop tailored interventions for each individual (Baskin-Sommers & Brazil, 2022; Baskin-Sommers et al., 2015; Brazil et al., 2018).

Note

1 TBS or 'terbeschikkingstelling' means 'disposal to be treated on behalf of the state'. In the Netherlands, one can be admitted to a TBS-clinic through an entrustment act when a person commits a serious offense that is likely to be the result of his mental disorder. Therefore, the offender is not liable (or has diminished responsibility) for this crime.

Acknowledgments

I.A. Brazil was supported by a VENI grant (451-15-014) awarded by the Netherlands Organisation for Scientific Research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Airy-Eggersten, A. S., Selby, M. J., & Laver, G. D. (1997). Comparison of neuropsychological performance in forensic and non-forensic populations. *Archives of Clinical Neuropsychology*, *12*(4), 281. https://doi.org/10.1016/s0887-6177(97)83432-4

Barch, D. M. (2005). The cognitive neuroscience of schizophrenia. *Annu. Rev. Clin. Psychol*, *1*(1), 321–353. https://doi.org/10.1146/annurev.clinpsy.1.102803.143959

Baskin-Sommers, A., & Brazil, I. A. (2022). The importance of an exaggerated attention bottleneck for understanding psychopathy. *Trends in Cognitive Sciences*, *26*(4), 325–336. https://doi.org/10.1016/j.tics.2022.01.001

Baskin-Sommers, A. R., Curtin, J. J., & Newman, J. P. (2015). Altering the Cognitive-Affective Dysfunctions of Psychopathic and Externalizing Offender Subtypes With Cognitive Remediation. *Clinical Psychological Science*, *3*(1), 45–57. https://doi.org/10.1177/2167702614560744

Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: control adjustments in the stop-signal paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(2), 392–404.

Bissett, P. G., & Logan, G. D. (2012a). Post-stop-signal adjustments: Inhibition improves subsequent inhibition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(4), 955–966. https://doi.org/10.1037/a0026778

Bissett, P. G., & Logan, G. D. (2012b). Post-stop-signal slowing: strategies dominate reflexes and implicit learning. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(3), 746–757. https://doi.org/10.1037/a0025429

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews. Neuroscience*, *14*(11), 786–799. https://doi.org/10.1038/nrn3577

Brazil, I. A., de Bruijn, E. R. A., Bulten, B. H., von Borries, A. K. L., van Lankveld, J. J. D. M., Buitelaar, J. K., & Verkes, R. J. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biological Psychiatry*, *65*(2), 137–143. https://doi.org/10.1016/j.biopsych.2008.08.011

Brazil, I., Maes, J., Scheper, I., Bulten, B., Kessels, R., Verkes, R., & de Bruijn, E. (2013). Reversal deficits in individuals with psychopathy in explicit but not implicit learning conditions. *Journal of Psychiatry & Neuroscience*, *38*(4), E13–E20. https://doi.org/10.1503/jpn.120152

Brazil, I. A., van Dongen, J. D. M., Maes, J. H. R., Mars, R. B., & Baskin-Sommers, A. R. (2018). Classification and treatment of antisocial individuals: From behavior to biocognition. *Neuroscience and Biobehavioral Reviews*, *91*, 259–277. https://doi.org/10.1016/j.neubiorev.2016.10.010

Buckholtz, J. W., & Meyer-Lindenberg, A. (2012). Psychopathology and the human connectome: toward a transdiagnostic model of risk for mental illness. *Neuron*, *74*(6), 990–1004. https://doi.org/10.1016/j.neuron.2012.06.002

Budhani, S., Richell, R. A., & Blair, R. J. R. (2006). Impaired reversal but intact acquisition: Probabilistic response reversal deficits in adult individuals with psychopathy. *Journal of Abnormal Psychology*, *115*(3), 552–558.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Cools, R., Altamirano, L., & D'Esposito, M. (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*, *44*(10), 1663–1673. https://doi.org/10.1016/j.neuropsychologia.2006.03.030

Dargis, M. A., Mattern, A. C., & Newman, J. P. (2017). Set-congruent priming stimuli normalize the information processing of psychopathic offenders. *Journal of Psychopathology and Behavioral Assessment*, *39*(2), 209–219. https://doi.org/10.1007/s10862-017-9587-9

De Brito, S. A., Viding, E., Kumari, V., Blackwood, N., Hodgins, S., & Soriano-Mas, C. (2013). Cool and hot executive function impairments in violent offenders with antisocial personality disorder with and without psychopathy. *PLoS ONE*, *8*(6), e65566. https://doi.org/10.1371/journal.pone.0065566

Dillien, T., Goethals, K., Sabbe, B., & Brazil, I. A. (2018). Reinforcement learning in child molesters. *Criminal Behaviour and Mental Health*, *29*(1), 31–42. https://doi.org/10.1002/cbm.2097

Fishbein, D., Sheppard, M., Hyde, C., Hubal, R., Newlin, D., Serin, R., Chrousos, G., & Alesci, S. (2009). Deficits in behavioral inhibition predict treatment engagement in prison inmates. *Law and Human Behavior*, *33*(5), 419–435. https://doi.org/10.1007/s10979-008-9163-7

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By Carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Gielen, W. J. M., & Akkermans, M. M. P. (2018). *Criminaliteit en slachtofferschap*. S. Kalidien & H. Eggen. Criminaliteit en Rechtshandhaving.

Goossens, I., Nicholls, T. L., Charette, Y., Wilson, C. M., Seto, M. C., & Crocker, A. G. (2019). Examining the high-risk accused designation for individuals found not criminally responsible on account of mental disorder. *Canadian Psychology/Psychologie Canadienne*, *60*(2), 102–114.

Gregory, S., Blair, R. J., Ffytche, D., Simmons, A., Kumari, V., Hodgins, S., & Blackwood, N. (2015). Punishment and psychopathy: A case-control functional MRI investigation of reinforcement learning in violent antisocial personality disordered men. *The Lancet Psychiatry*, *2*(2), 153–160. https://doi.org/10.1016/S2215-0366(14)00071-6

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. https://doi.org/10.1016/j.tics.2012.01.006

Hoppenbrouwers, S. S., Neumann, C. S., Lewis, J., & Johansson, P. (2015). A latent variable analysis of the psychopathy checklist–revised and behavioral inhibition system/behavioral activation system factors in north American and Swedish offenders. *Personality Disorders: Theory, Research, and Treatment*, *6*(3), 251–260.

Kerr, M. (1997). Boys' Behavioral Inhibition and the Risk of Later Delinquency. *Archives of General Psychiatry*, *54*(9), 809. https://doi.org/10.1001/archpsyc.1997.01830210049005

Kogel de, C. H. (2008). *De hersenen in beeld: Neurobiologisch onderzoek en vraagstukken op het gebied van verklaring, reductie en preventie van criminaliteit* (pp. 270). WODC: O&B.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295–327. https://doi.org/10.1037/0033-295x.91.3.295

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(2), 276–291. https://doi.org/10.1037/0096-1523.10.2.276

Maes, J. H. R., & Brazil, I. A. (2013). No clear evidence for a positive association between the interpersonal-affective aspects of psychopathy and executive functioning. *Psychiatry Research*, *210*(3), 1265–1274. https://doi.org/10.1016/j.psychres.2013.09.028

Martell, D. A. (1992). Forensic neuropsychology and the criminal law. *Law and Human Behavior*, *16*(3), 313–336.

Medical Association (WMA), W., & World Medical Association (WMA). (2009). Declaration of Helsinki. ethical principles for medical research involving human subjects. *Jahrbuch für Wissenschaft und Ethik*, *14*(1), 223–238. https://doi.org/10.1515/9783110208856.233

Meijers, J., Harte, J. M., Meynen, G., & Cuijpers, P. (2017). Differences in executive functioning between violent and non-violent offenders. *Psychological Medicine*, *47*(10), 1784–1793. https://doi.org/10.1017/S0033291717000241

Meijers, J., Harte, J. M., Meynen, G., Cuijpers, P., & Scherder, E. J. A. (2018). Reduced self-control after 3 months of imprisonment; A pilot study. *Frontiers in Psychology*, *9*, 69–70. https://doi.org/10.3389/fpsyg.2018.00069

Molleman, P. W., Driessen, J. M. A., Schilder, C., Bulten, B. H., & Brazil, I. A. (2021). Behavioral inhibition and activation system factors in offenders and non-offenders. *International Journal of Forensic Mental Health*, 1–13. https://doi.org/10.1080/14999013.2021.1947422

Morgan, A. B., & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, *20*(1), 113–136. https://doi.org/10.1016/S0272-7358(98)00096-8

Nigg, J. T. (2016). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*(4), 361–383. https://doi.org/10.1111/jcpp.12675

Ogilvie, M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. *Criminology*, *49*(4), 1063–1107. https://doi.org/10.1111/j.1745-9125.2011.00252.x

Owen, M. J. (2014). New approaches to psychiatric diagnostic classification. *Neuron*, *84*(3), 564–571. https://doi.org/10.1016/j.neuron.2014.10.028

Portnoy, J., Chen, F. R., & Raine, A. (2013). Biological protective factors for antisocial and criminal behavior. *Journal of Criminal Justice*, *41*(5), 292–299. https://doi.org/10.1016/j.jcrimjus.2013.06.018

Schmand, B., Bakker, D., Saan, R., & Louman, J. (1991). The Dutch Reading Test for Adults: A measure of premorbid intelligence level. *Tijdschrift voor Gerontologie En Geriatrie*, *22*(1), 15–19.

Schultz, I. Z. (2010). Neurocognitive Disorders in DSM-V: forensic perspective. *Psychological Injury and Law*, *3*(4), 271–288. https://doi.org/10.1007/s12207-010-9090-z

Schwartz, C. E., Holt, D. J., & Rosenbaum, J. F. (2013). Methodological challenges of moving beyond DSM categories: transdiagnostic studies of frontolimbic connectivity changes in vulnerable individuals. *Biological Psychiatry*, *74*(4), 240–241. https://doi.org/10.1016/j.biopsych.2013.05.030

Shumlich, E. J., Reid, G. J., Hancock, M., & Hoaken, N. S. (2018). Executive dysfunction in criminal populations: comparing forensic psychiatric patients and correctional offenders. *International Journal of Forensic Mental Health*, *18*(3), 243–259. https://doi.org/10.1080/14999013.2018.1495279

Spaans, M., Molendijk, M. L., de Beurs, E., Rinne, T., & Spinhoven, P. (2016). Self-reported personality traits in forensic populations: A meta-analysis. *Psychology, Crime & Law*, *23*(1), 56–78. https://doi.org/10.1080/1068316x.2016.1220555

Stern, Y. (2006). Cognitive Reserve and Alzheimer Disease. *Alzheimer's Disease and Associated Disorders*, *20*(2), 112–117. https://doi.org/10.1097/01.wad.0000213815.20177.19

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A*, *47*(3), 631–650. https://doi.org/10.1080/14640749408401131

van Toor, D., Roozen, H. G., Evans, B. E., Rombout, L., Van de Wetering, B. J. M., & Vingerhoets, A. J. J. M. (2010). The effects of psychiatric distress, inhibition, and impulsivity on decision making in patients with substance use disorders: A matched control study. *Journal of Clinical and Experimental Neuropsychology*, *33*(2), 161–168.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods*, *40*(2), 479–483. https://doi.org/10.3758/BRM.40.2.479

Verhage, F. (1964). *Intelligentie en leeftijd*. Koninklijke Van Gorcum.

Von Borries, A. K. L., Brazil, I. A., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., & de Bruijn, E. R. A. (2009). Neural correlates of error-related learning deficits in individuals with psychopathy. *Psychological Medicine*, *40*(9), 1559–1568.

Von Borries, A. K. L., Bulten, B. H., & Verkes, R. J. (2015). ForMINDS: Forensic Mental Information Processing and Neuropsychological Diagnostic System. Unpublished Manuscript not submitted for publication, Pompestichting Nijmegen, The Netherlands.

Wallace, J. F., Malterer, M. B., & Newman, J. P. (2009). Mapping Gray's BIS and BAS constructs onto factor 1 and factor 2 of hare's psychopathy checklist – Revised. *Personality and Individual Differences*, *47*(8), 812–816. https://doi.org/10.1016/j.paid.2009.06.019

Wingrove, J., & Bond, A. J. (1998). Angry reactions to failure on a cooperative computer game: The effect of trait hostility, behavioural inhibition, and behavioural activation. *Aggressive Behavior*, *24*(1), 27–36. https://doi.org/10.1002/(SICI)1098-2337(1998)24:1<27::AID-AB3>3.0.CO;2-P