What type of inhibition underpins performance on Luria's Fist-Edge-Palm task?

Varkovetski, Michael[1]., Pihkanen, Kaisla[1]., Shanker, Shanti[1,2]., Parris, Benjamin. A[1*]., & Gurr, Birgit[2].

1. Department of Psychology, Bournemouth University

2. Community Brain Injury and Adult Neuropsychology Poole Hospital, Dorset Healthcare University NHS Foundation Trust

*Address for correspondence:

Dr. Benjamin A. Parris

Department of Psychology

Faculty of Science and Technology

Bournemouth University

Talbot Campus

Fern Barrow

BH12 5BB

Email: bparris@bournemouth.ac.uk

Abstract

Objective: The Fist-Edge-Palm task is a motor sequencing task believed to be sensitive to frontal lobe impairment. The present study aimed to investigate the inhibitory processes underlying successful execution of this task.

Method: Seventy-two healthy participants were asked to perform the Fist-Edge-Palm task paced at 120bpms, 60bpms and self-paced. They also completed assessments sensitive to recently dissociated forms of inhibition (the Hayling Sentence Completion Test and the Stroop Colour-Word Test) that have recently been shown to be differentially lateralised (the right and left Prefrontal Cortex, respectively), and Cattell's Culture Fair Intelligence test.

Results: Analysis revealed that performance on the Hayling Sentence Completion Test predicted the amount of crude errors and the overall score on the Fist-Edge-Palm task, and that pacing condition had no effect on this outcome. Neither the Stroop Colour-Word Test nor Cattell's Culture Fair Intelligence Test predicted performance on the Fist-Edge-Palm task.

Conclusions: Consistent with some previous neuroimaging findings, the present findings suggest that Fist-Edge-Palm task performance relies on right lateralised inhibitory processes.

Keywords: Luria test; fist-edge-palm; executive function; prefrontal cortex; motor sequencing

Public Significance Statement:

Luria's Fist-Edge-Palm task is a well-known neuropsychological assessment employed to assess frontal lobe and psycho-motor functioning, and to detect voluntary movement disorders, but the inhibitory processes underpinning performance are not well understood. This study provides evidence indicating that right, but not left, prefrontal cortex inhibition functions underpin successful performance on Luria's task. These findings increase the clinical utility of this much-used task.

1. Introduction

Human voluntary movement is the outcome of a highly complex functional system which incorporates a multitude of cognitive processes relying on the synchronous organization and utilization of various cortical regions (Miziara, Manreza, Mansur, Reed & Buchpiguel, 2013), and as such a variety of neuropsychological assessments are critical to making fine distinctions of an individual's cognitive and motor abilities. One well-known and widely used task is the Fist-Edge-Palm task (FEP; Luria, 1966). The FEP task is a complex motor sequence task developed to assess frontal lobe and psycho-motor functioning and has been extensively utilized to detect voluntary movement disorders (Umetsu et al., 2002) and is included in numerous neuropsychological assessment batteries (Chen et al., 1995; Dubois, Slachevsky, Litvan, & Pillon, 2000; Golden, 1981; Mitsuhashi, Hirata, & Okuzumi, 2018). The task relies on fine motor coordination and a number of executive functions such as planning, updating and inhibition (Chan et al., 2015). During the FEP task, participants are required to reproduce a sequence of hand movements presented by the examiner, most commonly in the 'fist-edge-palm" arrangement. Participants are then asked to repeat the sequence of hand movements for a certain number of cycles. A single cycle is comprised of a fist with the knuckles down, followed by a cutting motion with the fingers fully extended, and concludes with a flat palm on the table with the fingers fully extended. Participants are required to break contact with the table between each change in hand movement.

Whilst there has been much work investigating the neural correlates of the FEP task (Astolfi et al., 2004; Chan, Rao, Chen, Ye & Zhang, 2006; Chan et al., 2015; Rao et al., 2008; Serrien & Brown, 2003; Umetsu et al., 2002) there is a surprising dearth of research on the cognitive mechanisms underpinning the FEP task. A central challenge inherent in correct performance of the FEP task is the inhibition of the prepotent but incorrect hand movements associated with the task. Participants must not perform the flat palm movement after the fist-with-knuckles-down movement. Yet it is clear that there are varying levels of success at implementing this form of inhibitory control (Weiner et al., 2011). Kok (1999) reviewed behavioural and psychophysiological studies and concluded that there are multiple forms of inhibition with distinct and interacting neuronal substrates. For example, Van veen and Carter (2005), and more recently Parris et al. (2019), have argued for separate neural substrates for two distinct types of inhibition in the Stroop Colour-Word Test (Stroop test). Consistently, Cipolloti et al. (2016) have recently proposed that there are several processes controlled by anatomically separable systems involved in inhibition tasks.

Cipolotti et al. (2016) systematically explored the relationship between inhibition, fluid intelligence and lesion location in a neuroimaging study employing voxel-based lesion-symptom mapping. The results from 30 frontal lobe patients of varying aetiologies showed that after accounting for fluid intelligence (as measured by Cattel's Culture Fair Intelligence test), performance on the Hayling Sentence Completion test (Hayling test; Burgess & Shallice, 1996), which requires participants to finish a sentence with a word that is not related to the sentence's meaning (e.g., *The captain wanted to stay with the sinking*….*lamp*) significantly relied on the integrity of the right Prefrontal Cortex (PFC), specifically the superior and middle frontal gyri. In contrast, performance on the Stroop test (Stroop, 1935), which requires participants to name the colour of the ink in which a word is presented (e.g., the word *red* is written in blue ink) relied on the integrity of the superior and middle frontal gyri of the left PFC. The authors noted that these findings are consistent with other findings in the literature (Aron, Robbins, & Poldrack, 2014; Demakis, 2004; Derrfuss, Brass, Neumann, & von Cramon, 2005; Geddes et al., 2014; Hodgson et al., 2007; Hornberger, Geng, & Hodges, 2011; Parris et al., 2019; Perret, 1974; Robbins, 2007; Robinson et al., 2015; Roca et al., 2010; Stuss et al., 2001; Szczepanski & Knight, 2014) and argued that lesion location is critical in producing impairments on two inhibitory tasks that despite loading similarly on verbal control, have different neurological substrates. Moreover, the authors argued that the two measures of inhibition are therefore possibly dissociable components of the executive function of inhibition, supporting Kok's (1999) conclusion that there are multiple forms of inhibition.

The aim of the present study was to investigate whether the distinct inhibitory mechanisms involved in the Hayling and Stroop tasks underpin performance on the FEP task. Given their uniquely and recently established doubly dissociated inhibitory mechanisms (Cipolotti et al., 2016), we investigated whether one or both cognitive tasks predicted FEP task performance. Whilst both the Hayling and Stroop tests are measures of lexical control, their established dissociation suggests important differences between the two tasks (see the Discussion section for a fuller consideration of this issue) and any association with FEP task performance would be informative as to the cognitive mechanisms underpinning this commonly used motor sequencing task. Following Cipolotti and colleagues (2016), a measure of fluid intelligence was included in our analysis as a measure of general cognitive ability. Fluid intelligence was included in our analysis because it has been shown to partially mediate performance on the Hayling test (Martin, Barker, Gibson, & Robinson, 2013) and could thus potentially be responsible for any relationships between Hayling and FEP performance.

Evidence for a right PFC locus for FEP performance in neuroimaging work (Rao et al., 2008) suggests that FEP performance might rely on similar inhibitory control mechanisms as those underpinning the Hayling test. To investigate this potential relationship and to sufficiently tax the capacities of our healthy participants we titrated task difficulty by asking participants to perform the FEP task in three pacing conditions. It was reasoned that the self-paced condition would lead to ceiling effects and so we introduced two externally paced conditions; one paced at 60bpm and one at 120bpm. It was expected that the externally-paced conditions would be harder than the self-paced condition, and of the externally-paced conditions, the faster condition (120bpm) would be harder than the slower condition (60bpm); it was reasoned that we might be more likely to observe a relationship between FEP task performance and the Hayling and / or Stroop tests in a healthy population if the task was more difficult. However, since this was not a key prediction in our investigation (indeed we were unsure as to how or whether pacing a condition would modify performance in healthy controls) it was a priori decided only to analyse the pacing conditions as separate conditions if an initial one-way ANOVA or non-parametric equivalent and appropriate follow up tests assessing differences between performance for the three levels of pacing returned a significant result. This constraint would have the effect of reducing the need for multiple analyses for each performance measure (subtle errors, crude errors, and self-corrections of those errors, and an overall FEP score).

2. Methods

2.1. Design

This study utilized a repeated measures design. Scores on the Stroop, Hayling and Cattell's Culture Fair Intelligence tests were the independent variables. The amount of subtle errors, crude errors, self-corrections, and the overall score on the FEP task were the dependent variables.

2.2. Participants

Seventy-two healthy university students (45 females and 27 males; mean age = 21years, SD = 3.30 – see Table 1) were recruited from the Psychology Research Participation System at Bournemouth University. All participants reported no neurological illness or psychiatric diseases. The Bournemouth University ethics panel approved this study. Participants received an information sheet prior to consenting and were debriefed at the end of the study. Written informed consent was obtained for every participant.

2.3. Materials

To measure left PFC performance, we used the Golden and Freshwater (2002) version of the Stroop Colour-Word Test which assesses prepotent response inhibition. The test consisted of three sections; each section arranged into five columns which consisted of 20 items each. The first section consisted of 100 words in black ink, the second section of 100 lines of 'XXXX" printed in coloured ink (blue, red and green), and the third section of 100 words "BLUE", "RED" and "GREEN" printed in an incongruent colour. In the first section, participants were instructed to read the words out loud as quickly as possible. In the second section, participants were instructed to name the colour of the ink for each item as quickly as possible. In the third section, participants were instructed to say out loud the ink colour of each word. Participants were instructed to complete each section as quickly as possible within a time limit of 45 seconds. If participants reached the end of the last column before the time limit, they were instructed to reread the page. Participants were not permitted to cover the page by any means, or to use their finger to guide their gaze. Whilst we employed the Golden and Freshwater (2002) version of the Stroop task (this was the version available to us) and used their recommended time limit, following Cipolotti and colleagues (2016) we calculated a single score based on the amount of correctly identified incongruent ink colours in the third section, within this time limit (Trenerry, Crosson, DeBoe and Leber, 1989).

The Hayling Sentence Completion Test (Burgess & Shallice, 1997), which assesses initiation speed and response suppression, is comprised of two sections. In the first section, participants orally completed 15 sentences missing the last word by generating a word which correctly completes each sentence. In the second section, participants orally completed another 15 uncompleted sentences, but were instructed to generate a word that was unconnected to the sentence in every way. Responses and response time were noted for each sentence. Following Cipolotti and colleagues (2016), we calculated two scores for the second section: the total Suppression Reaction Time and the Suppression Errors Score. The Suppression Errors score is the sum of the Total Category A Errors (errors which plausibly complete the sentence were given a score of 3) and the Total Category B Errors (errors which were somewhat connected to the sentence were given a score of 1). Whilst these scores can be scaled, doing so in a non-patient population leads to very little variability and as such we used the raw scores for all analyses.

The Cattell's Culture Fair Intelligence Test (Advanced version, Scale 3, Cattell, 1963) was used as a measure of fluid intelligence. The test is comprised of four subtests: classification, series, matrices and analogies. Each subtest was timed: three minutes for the first, four minutes for the second, three minutes for the third, and two and a half minutes for the fourth. Participants' correct answers were summed up for each subtest to give a final score which was then scaled to give an estimate of fluid intelligence.

2.4. Procedure

The FEP task lacks a standardized administration protocol and scoring scheme. Luria (1980) provided three administrative steps: first the examiner demonstrates for 10 cycles, then the patient imitates the examiner for 20 cycles, and finally the patient continues without model for 20 cycles. Despite Luria’s initial protocol, variation in administration of the FEP became evident. Several studies have not determined the amount of cycles a participant is required to complete (Rao et al., 2008; Chan et al., 2006), while other clinical studies have asked patients to perform as few as three cycles of the task (Iseki et al., 2013; Weiner et al., 2011), some six (Park & Moon, 2014; Miziara et al., 2013) and others 15 (Zaytseva et al., 2014). Given this inconsistency we selected a rough mid-way point between previous studies and elected to have our participants perform 10 cycles in each of the pacing conditions.

Participants were assessed on the FEP task using either the left or right hand. The first half of the sample was administered the FEP task using the right hand, and the second half of the sample was administered the FEP task using the left hand. Performance was counterbalanced in this way because each hand is controlled by the primary motor cortex in the contralateral hemisphere of the brain and so the relationship to tasks primarily recruiting the left (Stroop) or right (Hayling) hemisphere could otherwise potentially confound the outcome. Prior to administrating the FEP task, a simple pre-test was performed. Participants performed each of the individual motor movements within the FEP to demonstrate that no primary motor deficits were present.

Participants were first requested to observe and then imitate the examiner producing a single FEP cycle (fist-edge-palm). Participants were then asked if they understood how to perform the task correctly. Following this, the participants were asked to produce one FEP cycle on their own. Participants were then asked to perform 10 FEP cycles at three different tempos; at their own tempo, an externally paced tempo of 60 beats per minute, and an externally paced tempo of 120 beats per minute. The examiner made a video recording of the hand performing the

FEP task throughout all three tempos. Instructions regarding what to do in case of an error were explained to the participant prior to the start of the experiment. If a participant made a subtle error in technique or from hesitation/lag, they were instructed to continue through their current cycle and to begin the next cycle normally. However, if the participant made a crude error in producing the wrong hand movement, they were instructed to stop and restart that cycle, and to continue onto the next cycle normally. Additionally, participants were asked not to externally narrate they own execution of the FEP task by saying “fist-edge-palm”. The examiner kept count of the number of completed cycles for each tempo and asked the participant to stop when they completed 10 cycles. The order of tempos and which hand the participant used was counterbalanced to reduce any order and handedness effects. Due to counterbalancing, and a lack of an equal number of left hand dominant vs. right hand dominant participants, hand dominance was not accounted for in the analysis.

Following the completion of all three tempos of the FEP task, the participants’ cognitive performance was assessed using the Stroop, Hayling and Culture Fair Intelligence tests. All three tests were administered in the published standard manner and administration was counterbalanced across participants.

2.5. FEP task scoring

Variation exists in how the scores were calculated in previous studies. Numerous studies scored only crude errors, such as omission or repetition of a motion (Park & Moon, 2014; Miziara et al., 2013). Other studies scored more subtle technical errors, such as flexing of the fingers during cutting motions (Weiner et al., 2011). Furthermore, some studies implemented a point system when scoring errors. In this system, the score is dependent on how many crude/subtle errors are made (Zaytseva et al., 2014), or how many successful consecutive cycles the patient completes (Iseki et al., 2013).

For this reason, in the present study, we created a new method for scoring performance. Subtle errors, crude errors, and self-corrections of those errors and an overall FEP score were used as the dependent variables and were calculated upon reviewing each participant's video recording. The amount of subtle and crude errors a participant made was scored by two researchers. If a disagreement arose on the scoring of any of the indices of performance, they would re-watch the video until an agreement was reached. A subtle error was scored when a participant produced a hand movement with poor technique, or when a hesitation/lag was evident between hand movements. Poor technique is defined as a hand movement with; a fist orientated the wrong way, an edge with the fingers curled in, or a palm that is angled more than 45° above the table. A crude error was scored when a participant produced the wrong hand motion (e.g., the participant produces a fist instead of a flat palm, following the production of an edge). The amount of self-corrections was also scored. Each subtle error was counted as one point, and each crude error, which we deemed as being a bigger and more problematic error, was counted as two points. Self-corrections were counted as .5 points. To calculate each participant's overall FEP score, the total self-corrections score (across all tempo conditions) was subtracted from the total error score (crude + subtle errors across all tempo conditions).

2.6. Statistical Analysis Plan

To determine whether either of the four predictors (Hayling test suppression error score, Hayling test suppression reaction score, Stroop test score, or fluid intelligence) were able to significantly predict participants' overall FEP score, we first aimed to conduct a multiple regression analysis including all measures as predictor variables. We also planned to conduct further multiple regression analyses to determine whether the predictors were able to predict the four dependent variables (crude errors, subtle errors, crude error self-corrections, subtle error self-corrections). However, before conducting the individual analyses of the four dependent variables, we planned to initially determine whether the values of the dependent variables significantly differed between the three tempo conditions (Self-tempo, 60bpm and 120bpm) using a one-way repeated-measures ANOVA and follow up Wilcoxon matched comparisons (with Bonferroni correction). If scores did not significantly differ between the tempo conditions, scores across tempo conditions were combined to reduce the number of analyses conducted. In the event that the DVs were not normally distributed, the non-parametric versions of the tests were used. Finally, in order to establish whether DVs were statistically independent, we planned to run a Kendall rank correlation.

3. Results

See Table 1 for descriptive data about participants, and Table 2 regarding descriptive data of the neuropsychological assessment scores. See Table 3 for descriptive data regarding FEP task measures.

Casewise diagnostics and a scatterplot revealed that one participant was an outlier with an overall FEP score of 33.5 (compared to an average of 8.94). They were removed from the analysis since it was noted during testing that they exhibited difficulties in following the rhythm of the metronome, which may have increased errors, and thus we believe that they were not an accurate representation of the target population.

3.1 Analysis of overall FEP score

To assess linearity, a scatterplot of participants' overall FEP score against each of the four predictor variables with a superimposed regression line was plotted. Visual inspection of these plots revealed a linear relationship between the overall FEP score, and each of the predictor variables. There was homoscedasticity, normality of the residuals and all variance inflation factors were below 1.27 indicating small to nil multicollinearity. With a perfect score of zero, the overall FEP score had a range of scores from zero to 28.5.

The four predictor variables accounted for 23% of the variation in participants' overall FEP score with adjusted R^2 = 19%, a medium size effect according to Cohen (1988). The four predictor variables significantly predicted the overall FEP score, $F(4, 66) = 5.03$, $p = .001$; see Figure 1. The analysis indicated that only the Hayling test suppression error score significantly predicted the overall FEP score ($\beta = .440$, $p = .003$; see Figure 2), while Hayling test suppression reaction score ($\beta = .062$, $p = .253$), fluid intelligence ($\beta = .066$, $p = .265$), and Stroop test score ($\beta = -.110$, $p = .198$) did not.

3.2 Independent analysis of each dependent variable

Several of the variables appeared to be relatively rare and significantly skewed. We used P-P plots and indices for acceptable limits of ±2 for skewness and kurtosis (Trochim & Donnelly, 2006; Field, 2009; Gravetter & Wallnau, 2014) to check the assumption of normality. The following variables were shown to be non-normally distributed: Self-Tempo Subtle Errors (Skewness = 2.911, Kurtosis = 9.977), 60bpm Subtle Errors (Skewness = 1.737, Kurtosis = 2.397), Self + 60bpm Subtle Correction (Skewness = 1.706, Kurtosis = 3.101), 120bpm Subtle Corrections (Skewness = 2.055, Kurtosis = 3.942), 60bpm Crude Errors (Skewness = 1.803, Kurtosis = 2.896), 120bpm Crude Errors (Skewness = 2.938, Kurtosis = 10.912), Self-Tempo

Crude Corrections (Skewness = 2.373, Kurtosis = 6.214), 60bpm Crude Corrections (Skewness = 2.700, Kurtosis = 7.821), 120bpm Crude Correction (Skewness = 2.572, Kurtosis = 7.574), Overall FEP Score (Skewness = 1.314, Kurtosis = 2.054). Therefore, prior to analysis, we attempted to normalize the variables using log transformations to no success, and thus continued with the non-transformed variables. As a consequence, Friedman's test and Wilcoxon Matched-Pairs tests were employed for analyses of the means.

Furthermore, upon assessing assumptions for regression it was shown that a few variables did not show homoscedastic residuals (self-tempo and 60bpm combined subtle error score, 120bpm subtle error score) and some variables' residuals deviated from normality on the Normal P-P plots (self-tempo and 60bpm combined subtle self-correction score, and 120bpm subtle self-correction score). This could lead to imprecise coefficient estimates and increases the likelihood of a model term that is significant when it is actually not. Therefore, the results from these analyses should be interpreted with caution.

Subtle errors: A Friedman test showed that the amount of subtle errors a participant made significantly differed between tempo conditions; χ^2 (2) = 37.862, $p < .001$. Wilcoxon matched comparisons were performed with a Bonferroni correction for multiple comparisons. Statistical significance was accepted at the $p < .0167$ level. There was a significant difference between the scores for self-tempo subtle errors and 120bpm subtle errors ($p = .001$, $r = -0.278$), and between 120bpm subtle errors and 60bpm subtle errors ($p < .001$, $r = -0.434$). There was no significant difference between self-tempo subtle errors and 60bpm subtle errors ($p = .047$, $r = 0.166$). Thus, two multiple regressions analyses were conducted; the first on a combined score of the self-tempo and 60bpm subtle errors, and the second on the 120bpm subtle errors. The results of the multiple regression analysis indicated that neither of the four predictor variables were able to predict the amount of subtle errors made in the self-tempo and 60bpm conditions (R^2 = .053, F(4,66) = .918, *p* = .459), or the 120bpm condition (R^2 = .052, F(4,66) = .909, *p* = .464).

Crude errors: A second Friedman test showed that the amount of crude errors a participant made did not significantly differ between each tempo condition; χ^2 (2) = 1.589, *p* = .452. Thus, a total crude errors score was calculated and used for the multiple regression analysis. The four predictor variables accounted for 24% of the variation in participants' total crude errors score with adjusted R^2 = 20%, a medium size effect according to Cohen (1988). The four predictor variables significantly predicted the total crude errors score, *F*(4, 66) = 5.284, *p* = .001; see Figure 3. The analysis indicated that only the Hayling test suppression error score significantly predicted the total crude error score (β = .179, *p* = .004; see Figure 4), while Hayling test suppression reaction score (β = .037, *p* = .107), fluid intelligence (β = .029, *p* = .249), and Stroop test score (β = -.029, *p* = .424) did not.

Subtle self-corrections: A third Friedman test showed that the amount of subtle self-corrections a participant made significantly differed between tempo conditions χ^2 (2) = 7.189, *p* = .027. Pairwise comparisons were once again performed with a Bonferroni correction for multiple comparisons. Statistical significance was accepted at the *p* < .0167 level. A significant difference in the amount of subtle self-corrections a participant made existed between 60bpm and 120 bpm conditions (*p* = .002, *r* = -0.254). However, no significant differences in the amount of subtle self-corrections were found between self-tempo and 60bpm conditions (*p* = .052, *r* = 0.163), or between self-tempo and 120bpm conditions (*p* = .318, *r* = -0.084). Thus, two multiple regression analyses were conducted; the first on a combined score of the self-tempo and 60bpm subtle self-correction conditions and the second on the 120bpm subtle self-correction condition. However, the results of the multiple regression analyses indicated that neither of the four predictors were able to predict the amount of subtle self-corrections made in the self-tempo and 60bpm conditions (R^2 = .035, F(4, 66) = .600, *p* = .664) or the 120bpm condition (R^2 = .044, F(4, 66) = .754, *p* = .559).

Crude Self-Corrections: A final Friedman test showed that the amount of crude self-corrections a participant made did not significantly differ between each tempo condition; χ^2 (2) = .819, *p* = .664. Thus, a total crude self-correction score was calculated and used for the multiple regression analysis. Like the analysis of subtle self-corrections, the results of the multiple regression analysis indicated that neither of the four predictors were able to predict the overall amount of crude self-corrections a participant made (R^2 = .038, F(4, 66) = .647, *p* = .631).

3.3 Correlations between errors

Lastly, a Kendall rank correlation was run to assess the relationship between the subtle and crude errors made during execution of the FEP task. A Kendall rank correlation was chosen due to the violation of the normality assumption among the variables, and because it is considered to be more robust and efficient than the Spearman correlation (Knight, 1996). No significant correlation between total crude errors and subtle errors of each tempo condition was evident. Table 4 summarises these results.

3.4 Summary of results

In summary, the analysis indicated that only the Hayling test suppression error score significantly predicted the overall FEP score (β = .440, *p* = .003; see Figure 2), while the other IVs did not. Moreover, only the Hayling test suppression error score was able to significantly predict participants total crude error score (β = .179, *p* = .004; see Figure 4).

4. Discussion

By assessing performance on the FEP task and neuropsychological assessments sensitive to recently doubly-dissociated inhibitory functions involved in the Hayling and Stroop tests, the present study was able to shed some light on the inhibitory functions underpinning FEP task performance. The Hayling test, a verbal suppression test known for its sensitivity to right PFC lesions (Cipolotti et al., 2016; Robinson et al., 2015), and in particular the suppression score associated with the test, was able to significantly and independently predict FEP task performance, a motor sequencing task, whereas Stroop test performance and fluid intelligence did not significantly predict performance on the FEP task. Other than the overall FEP score, the Hayling test suppression error score was also a significant predictor of crude error scores. There were no other significant predictive relationships between our independent and dependent variables. Overall, our findings provide complimentary cognitive evidence for the involvement of right PFC inhibitory functions in FEP task performance reported in a previous neuroimaging assay (Rao et al., 2008).

Kok (1999) argued that the executive function component of inhibition may comprise multiple forms, each with their own distinct neuronal system. Whilst Cipolotti et al. (2016) argued that their findings supported this assertion by indicating dissociable inhibitory functions, it is not clear how the two types of inhibition measured by these two tasks differ. Cipolotti et al. (2016) described the Stroop test as an inhibitory test that measures the ability to inhibit pre-potent responses, and it could be argued that the Hayling test involves semantic inhibition in that it involves supressing an appropriate semantic response. In fact, whilst the locus of the Stroop effect is commonly attributed to competition between the competing responses that are indicated by each dimension of the Stroop stimulus, it has recently been shown that the Stroop effect involves competition at various levels of processing including, but not limited to, response, semantic and task level conflict (see Augustinova, Parris & Ferrand, 2019; Parris, Augustinova & Ferrand, 2019). Moreover, the type of competition might well depend on whether the Stroop stimuli are presented in blocked or unblocked formats (Hasshim & Parris, 2017), with the former being more common in the paper version of the task (as used here). For present purposes we interpret the tasks in line with the interpretation of Cipolotti et al. who argued that the Hayling test measures inhibitory mechanisms in the right hemisphere and the Stroop task, inhibitory mechanisms in the left hemisphere. Nevertheless, more research is needed to determine what differentiates the inhibition processes involved in these two tasks.

Cipolotti et al. (2016) noted that both tests involve suppressing a dominant response, but also differ in the involvement of other complex processes such as goal maintenance in the face of a visually presented distraction in the case of the Stroop test and strategy utilisation in the Hayling test . Indeed, it could be argued that the FEP task has more in common with the Stroop test in that it requires suppression of a set of manual responses (a set number of possible colour responses in the Stroop test and set number of movements in the Hayling test). The Hayling test in contrast does not involve inhibition of a manual response and requires the inhibition of just one response. However, in the Stroop test, any of the possible response options could be the next correct response, whereas in the FEP task the next correct response is pre-determined. Maintaining the correct sequence might require the invocation of a strategy such as constantly repeating “fist-edge-palm” to oneself, just as efficient performance on the Hayling test requires strategy use (e.g., use the name of objects in the room as your unrelated response). Unfortunately, our data do not permit a conclusion as to the exact relationship between the inhibition mechanisms involved in the FEP and Hayling tests, they do however give direction for future research aimed at understanding the mechanisms underpinning the FEP task.

A notable limitation of the present research is that our pacing manipulation was not wholly effective. Whilst, as predicted, the 120bpm condition was shown to be more difficult in terms of errors committed, the self-paced condition was shown to be of equal difficulty to the 60bpm condition. However, the predictive relationship between the Hayling test suppression error score and FEP performance was not dependent on a particular pacing condition. Nevertheless, a future study might consider employing an even faster paced condition to induce more subtle errors and corrections. Such a manipulation might reveal the cognitive processes underpinning more refined errors.

Another limitation of the present research is that our method of calculating the Hayling test scores. To score the Hayling test, the number of category A and category B errors are summed and then scaled. Our scaled scores meant that >90% of the participants had a score of 6 which is clearly not enough variability for our measures to explain. Due to this lack of variability in the Hayling scaled scores, we did not use the scaled scores for either index of Hayling performance. In the interest of maintaining performance variability among participants, we instead used raw scores for all Hayling test analyses. Undoubtedly, this reduces the validity of our analyses. Future studies, particularly those working with clinical populations, might consider using the scaled scores for analysis.

A final limitation of the present research is that some participants completed the FEP task with their non-dominant hand. This was the case because it was reasoned that having participants complete the task with only their dominant hand would result in most participants recruiting left hemisphere motor control functions (87.5% of the participants were right handed), which might have confounded any relationship with the higher order cognitive control functions involving inhibition whose apparent laterality motivated the present research. Having some participants complete the task with their non-dominant hand might have increased the number of errors in their performance. However, the assumption that the control processes for the nondominant hand are a weaker analogue of those of the dominant arm has been argued to be erroneous and instead research suggests that there are specific advantages for each arm for different aspects of a movement where the left hemisphere specialises in planning and coordinating actions, and the right hemisphere specialises in updating actions and stopping at a goal position (Mutha, Haaland, & Sainburg, 2012). Nevertheless, future studies might consider recruiting an equal number of left hand dominant and right-hand dominant participants.

For the purposes of this research, a new protocol and scoring method for the FEP task was introduced. It is hoped that this method proves useful for future research. However, the protocol and method does present with some shortcomings; meaning it might not be suitable for all future research, particularly research involving patients. First, Luria recommended taking patients through 20 guided cycles of the task before testing their ability to do it independently. We did not do this in the present study precisely because we were using a healthy population. The inhibition mechanisms involved might change after such prolonged practice. Indeed, strategy use might be of less importance and thus could alter the inhibitory mechanisms involved (and the associated neural substrate). Second, whilst the scoring of subtle errors and self-corrections might be informative in a heathy adult population, patient populations would be more likely to make just the crude errors. Notably, however, none of the analyses involving these measures of more refined performance produced significant results, and whilst we must not draw strong conclusions based on null results, our data do point to the need for predictor variables of an equally refined nature.

In summary, our findings suggest that performance on the FEP task can be predicted by performance on the Hayling Sentence Completion test, and that a right PFC inhibitory process is key for the successful execution of the FEP task. Additionally, we believe that the novel and more robust administration protocol and scoring system will be of value to clinicians utilizing the FEP task as a diagnostic tool to measure the magnitude of impairment. Future studies should recruit clinical populations to further develop the FEP scoring system, and to assess its reliability in distinguishing different diagnoses.

References

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences, 18*(4), 177-185. doi:10.1016/j.tics.2013.12.003

Astolfi, L., Cincotti, F., Mattia, D., Salinari, S., Babiloni, C., Basilisco, A., . . . Babiloni, F. (2004). Estimation of the effective and functional human cortical connectivity with structural equation modeling and directed transfer function applied to high-resolution EEG. *Magnetic Resonance Imaging, 22*(10), 1457-1470. doi:10.1016/j.mri.2004.10.006

Augustinova, M., Parris, B. A., & Ferrand, L. (2019). The Loci of Stroop Interference and Facilitation Effects With Manual and Vocal Responses. Frontiers in Psychology, 10. doi: 10.3389/fpsyg.2019.01786

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34*(4), 263-272. doi:10.1016/0028-3932(95)00104-2

Burgess, P. W., & Shallice, T. (1997). The Hayling and Brixton Tests. Bury St Edmunds: Thames Valley Test Company.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1-22. doi:10.1037/h0046743

Chan, R. C., Rao, H., Chen, E. E., Ye, B., & Zhang, C. (2006). The neural basis of motor sequencing: An fMRI study of healthy subjects. *Neuroscience Letters, 398*(3), 189-194. doi:10.1016/j.neulet.2006.01.014

Chan, R. C., Huang, J., Zhao, Q., Wang, Y., Lai, Y., Hong, N., . . . Dazzan, P. (2015). Prefrontal cortex connectivity dysfunction in performing the Fist–Edge–Palm task in patients with first-episode schizophrenia and non-psychotic first-degree relatives. *NeuroImage: Clinical, 9*, 411-417. doi:10.1016/j.nicl.2015.09.008

Chen, E. Y., Shapleske, J., Luque, R., Mckenna, P. J., Hodges, J. R., Calloway, S., . . . Berrios, G. E. (1995). The Cambridge Neurological Inventory: A clinical instrument for assessment of soft neurological signs in psychiatric patients. *Psychiatry Research, 56*(2), 183-204. doi:10.1016/0165-1781(95)02535-2

Cipolotti, L., Spanò, B., Healy, C., Tudor-Sfetea, C., Chan, E., White, M., . . . Bozzali, M. (2016). Inhibition processes are dissociable and lateralized in human prefrontal cortex. *Neuropsychologia, 93*, 1-12. doi:10.1016/j.neuropsychologia.2016.09.018

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Demakis, G. J. (2004). Frontal Lobe Damage and Tests of Executive Processing: A Meta-Analysis of the Category Test, Stroop Test, and Trail-Making Test. *Journal of Clinical and Experimental Neuropsychology, 26*(3), 441-450. doi:10.1080/13803390490510149

Derrfuss, J., Brass, M., Neumann, J., & Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: Meta-analyses of switching and Stroop studies. *Human Brain Mapping, 25*(1), 22-34. doi:10.1002/hbm.20127

Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB: A frontal assessment battery at bedside. *Neurology, 55*(11), 1621-1626. doi:10.1212/wnl.55.11.1621

Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE Publications.

Geddes, M. R., Tsuchida, A., Ashley, V., Swick, D., & Fellows, L. K. (2014). Material-specific interference control is dissociable and lateralized in human prefrontal cortex. *Neuropsychologia, 64*, 310-319. doi:10.1016/j.neuropsychologia.2014.09.024

Golden, C. J. (1981). The Luria-Nebraska Children's Battery: Theory and formulaiton. In G. W. Hynd & J. E. Orbrzut (Eds.), *Neuropsychological assessment and the school-age child: Issues and procedures* (pp. 277-302). New York, NY: Grune & Stratton.

Golden, C. J., & Freshwater, S. M. (2002). Luria-Nebraska Neuropsychological Battery. In W. I. Dorfman & M. Hersen (Eds.), *Understanding Psychological Assessment* (pp. 59-75). Boston, MA: Springer.

Gravetter, F. and Wallnau, L. (2014) Essentials of Statistics for the Behavioral Sciences. 8th Edition, Wadsworth, Belmont, CA.

Hasshim, N. & Parris, B. A. (2017). Trial type mixing substantially reduces the response set effect in the Stroop task. *Acta Psychologica, 189,* 43-53.

Hodgson, T. L. Chamberlain, M., Parris, B. A., James, M., Gutowski, N., Husain, M. & Kennard, C. (2007). The role of the ventrolateral frontal cortex in inhibitory oculomotor control. *Brain, 130*(6), 1525-37.

Hornberger, M., Geng, J., & Hodges, J. R. (2011). Convergent grey and white matter evidence of orbitofrontal cortex changes related to disinhibition in behavioural variant frontotemporal dementia. *Brain, 134*(9), 2502-2512. doi:10.1093/brain/awr173

Iseki, C., Takahashi, Y., Wada, M., Kawanami, T., & Kato, T. (2013). Subclinical Declines in the Verbal Fluency and Motor Regulation of Patients with AVIM (Asymptomatic Ventriculomegaly with Features of Idiopathic NPH on MRI): A Case-controlled Study. *Internal Medicine, 52*(15), 1687-1690. doi:10.2169/internalmedicine.52.8914

Kok, A. (1999). Varieties of inhibition: Manifestations in cognition, event-related potentials and aging. *Acta Psychologica, 101*(2-3), 129-158. doi:10.1016/s0001-6918(99)00003-7

Luria, A. R. (1966). *Higher cortical functions in man*. Basic Books.

Luria, A. R. (1980). *Higher cortical Functions in man*. Basic Books.

Martin, A. K., Barker, M. S., Gibson, E. C., & Robinson, G. A. (2019). Response initiation and inhibition and the relationship with fluid intelligence across the adult lifespan. Archives of Clinical Neuropsychology. doi: 10.1093/arclin/acz044

Mikadze, Y. V. (2014). The principles of plasticity in Lurian neuropsychology. *Psychology & Neuroscience, 7*(4), 435-441. doi:10.3922/j.psns.2014.4.02

Mitsuhashi, S., Hirata, S., & Okuzumi, H. (2018). Role of inner speech on the Luria hand test. *Cogent Psychology, 5*(1). doi:10.1080/23311908.2018.1449485

Miziara, C. S., Manreza, M. L., Mansur, L., Reed, U. C., & Buchpiguel, C. A. (2013). Sequential motor task (Lurias Fist-Edge-Palm Test) in children with benign focal epilepsy of childhood with centrotemporal spikes. *Arquivos De Neuro-Psiquiatria*, *71*(6), 380-384. doi:10.1590/0004-282x20130043

Mutha, P. K., Haaland, K. Y., & Sainburg, R. L. (2012). The effects of brain lateralisation on motor control and adaptation. *Journal of Motor Behavior, 44*(6), 455-469.

Park, H., & Moon, S. Y. (2014). Usefulness Of Fist-Edge-Palm Test In A Dementia Clinic. *Alzheimers & Dementia, 10*(4). doi:10.1016/j.jalz.2014.05.442

Parris, B. A., Augustinova, M. & Ferrand. L. (2019). Editorial: The Locus of the Stroop effect. *Frontiers in Psychology*, *10*: 2860. doi: 10.3389/fpsyg.2019.02860

Parris, B. A., Wadsley, M. G., Hasshim, N., Augustinova, M. & Ferrand. L. (2019). An fMRI study of response and semantic conflict in the Stroop task. *Frontiers in Psychology*, *10*: 2426. doi: 10.3389/fpsyg.2019.02426

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia*, *12*(3), 323-330. doi:10.1016/0028-3932(74)90047-5

Rao, H., Di, X., Chan, R. C., Ding, Y., Ye, B., & Gao, D. (2008). A regulation role of the prefrontal cortex in the fist-edge-palm task: Evidence from functional connectivity analysis. *NeuroImage*, *41*(4), 1345-1351. doi:10.1016/j.neuroimage.2008.04.026

Robbins, T. (2007). Shifting and stopping: Fronto-striatal substrates, neurochemical modulation and clinical implications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 917-932. doi:10.1098/rstb.2007.2097

Robinson, G. A., Cipolotti, L., Walker, D. G., Biggs, V., Bozzali, M., & Shallice, T. (2015). Verbal suppression and strategy use: A role for the right lateral prefrontal cortex? *Brain*, *138*(4), 1084-1096. doi:10.1093/brain/awv003

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., . . . Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain*, *133*(1), 234-247. doi:10.1093/brain/awp269

Serrien, D. J., & Brown, P. (2003). The integration of cortical and behavioural dynamics during initial learning of a motor task. *European Journal of Neuroscience*, *17*(5), 1098-1104. doi:10.1046/j.1460-9568.2003.02534.x

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662. doi:10.1037/h0054651

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia*, *39*(8), 771-786. doi:10.1016/s0028-3932(01)00013-6

Szczepanski, S. M., & Knight, R. T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron, 83*(5), 1002-1018.

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *Stroop neuropsychological screening test manual*. Lutz, FL: Psychological Assessment Resources.

Trochim, W.M. and Donnelly, J.P. (2006). *The Research Methods Knowledge Base*. 3rd Edition, Atomic Dog, Cincinnati, OH.

Umetsu, A., Okuda, J., Fujii, T., Tsukiura, T., Nagasaka, T., Yanagawa, I., . . . Suzuki, K. (2002). Brain Activation during the Fist-Edge-Palm Test: A Functional MRI Study. *NeuroImage, 17*(1), 385-392. doi:10.1006/nimg.2002.1218

Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *Neuroimage, 27*(3), 497-504.

Weiner, M. F., Hynan, L. S., Rossetti, H., & Falkowski, J. (2011). Lurias three-step test: What is it and what does it tell us? *International Psychogeriatrics, 23*(10), 1602-1606. doi:10.1017/s1041610211000767

Zaytseva, Y., Korsakova, N., Gurovich, I. Y., Heinz, A., & Rapp, M. A. (2014). Luria revisited: Complex motor phenomena in first episode schizophrenia and schizophrenia spectrum disorders. *Psychiatry Research, 220*(1-2), 145-151. doi:10.1016/j.psychres.2014.08.009

Table 1. Participant information

		Mean	Standard Deviation	Group Size
Age		21	3.3	72 (100%)
Education Level (Years)		15.2	1.1	72 (100%)
Gender	***Male***	-	-	27 (38%)
	Female	-	-	45 (62%)
Handedness	***Right***	-	-	63 (89%)
	Left	-	-	9 (11%)

Table 2. Performance data for the neuropsychological measures employed

	Mean	**Standard Deviation**
Fluid Intelligence Score	114.5	11.64
Hayling Suppression Reaction Time Score	15.19	13.98
Hayling Suppression Error Score	7.25	5.16
Stroop Score	51.46	8.06

Table 3. Descriptive statistics of the FEP scores

		Mean	SD
Subtle Errors	*Self-tempo*	1.36	2.31
	60bpm	0.63	1.01
	120bpm	2.09	1.86
	Total	8.05	3.52
Crude Errors	*Self-tempo*	1.59	1.18
	60bpm	1.67	1.25
	120bpm	2.66	2.14
	Total	5.92	2.82
Subtle	*Self-*	0.71	0.63

Table 4. Table summarizing the results of the Kendall correlation analysis

		Self-tempo + 60bpm Subtle Errors	120bpm Subtle Errors	Total Crude Errors
Self-tempo + 60bpm Subtle Errors	*Correlation Coefficient*	1.000	.336**	.123
	Sig. (2-tailed)	.	.000	.187
	N	71	71	71
120bpm Subtle Errors	*Correlation Coefficient*	.336**	1.000	.091
	Sig. (2-tailed)	.000	.	.332
	N	71	71	71
Total Crude Errors	*Correlation Coefficient*	.123	.091	1.000
	Sig. (2-tailed)	.187	.322	.
	N	71	71	71

**. Correlation is significant at the 0.001 level (2-tailed).

Figure Captions

Figure 1. Scatter plot depicting the multiple regression model for Overall performance on the Fist-Edge-Palm test.

Figure 2. Scatter plot depicting the correlation between the overall FEP score and the Hayling suppression error score.

Figure 3. Scatter plot depicting the multiple regression model for Total Crude Errors on the Fist-Edge-Palm test.

Figure 4. Scatter plot depicting the correlation between the total crude error score and the Hayling suppression error score.

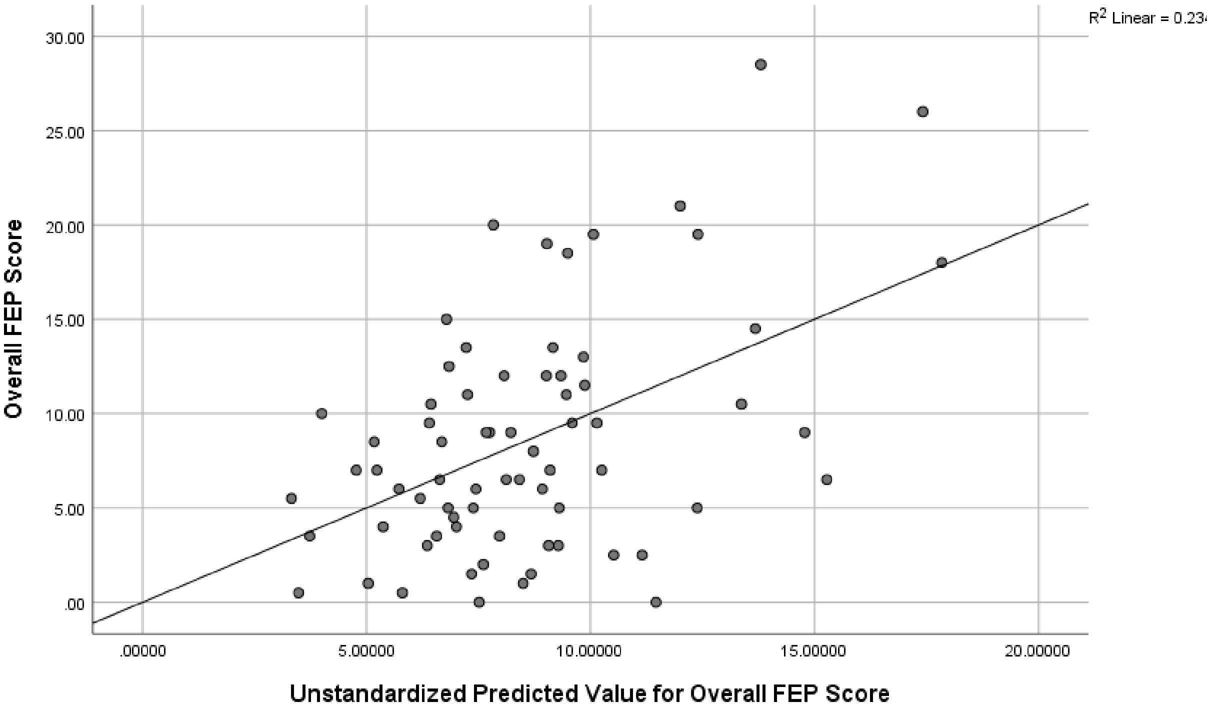

R2 Linear = 0.234
30.00
25.00
20.00
15.00
10.00
5.00
.00
Overall FEP Score
.00000
5.00000
10.00000
15.00000
20.00000
Unstandardized Predicted Value for Overall FEP Score

Figure 1.

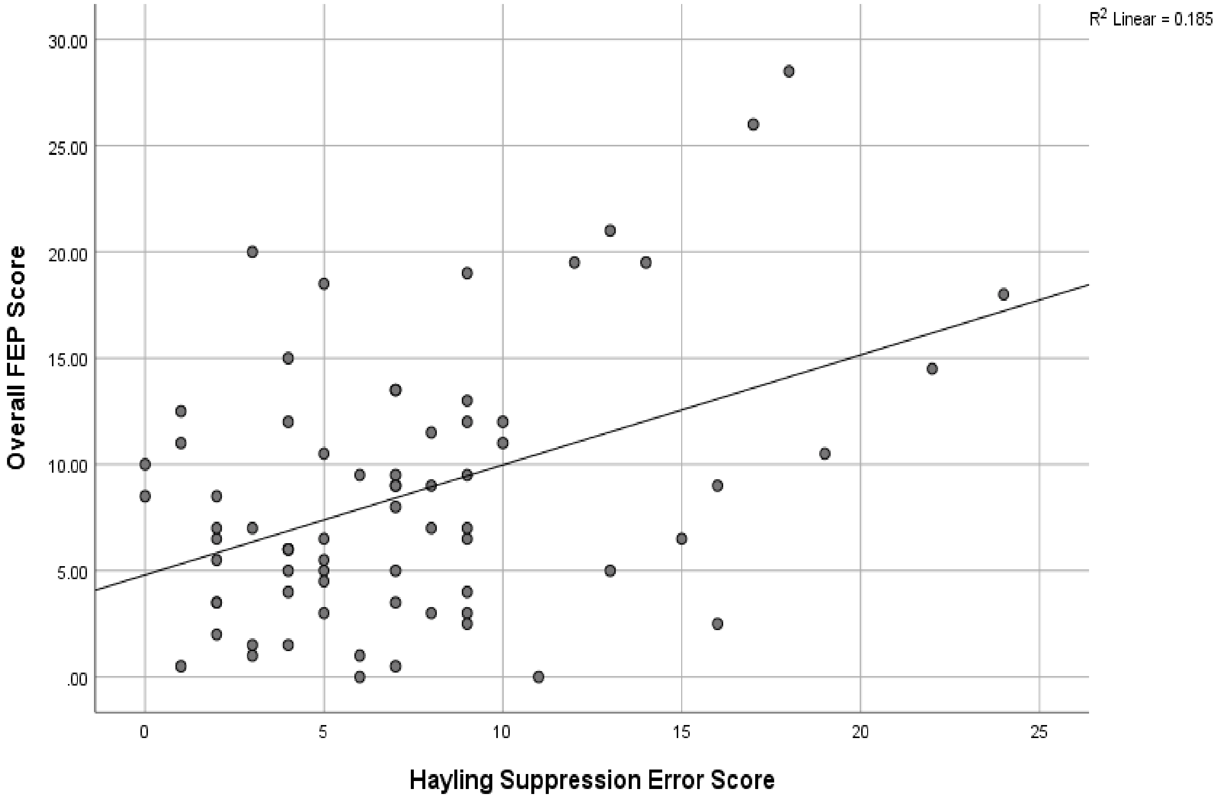
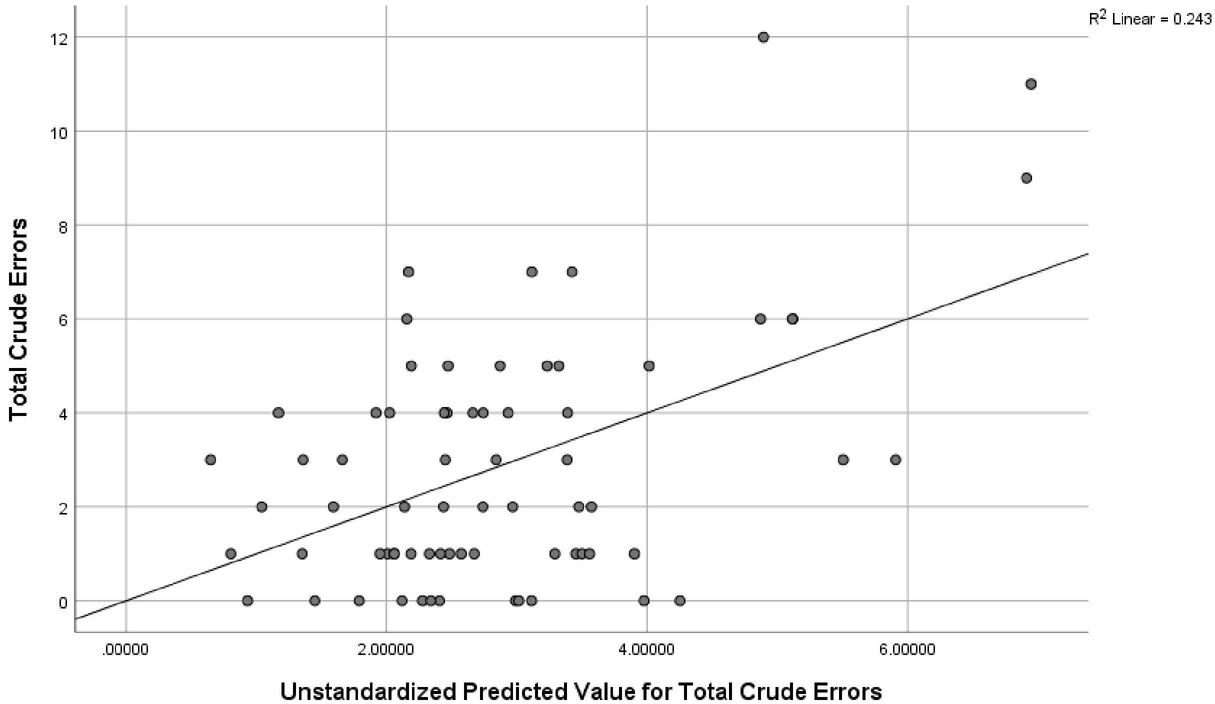

R2 Linear = 0.185
Overall FEP Score
30.00
25.00
20.00
15.00
10.00
5.00
.00
0
5
10
15
20
25
Hayling Suppression Error Score

Figure 2.



R2 Linear = 0.243
Total Crude Errors
12
10
8
6
4
2
0
.00000
2.00000
4.00000
6.00000
Unstandardized Predicted Value for Total Crude Errors

Figure 3.

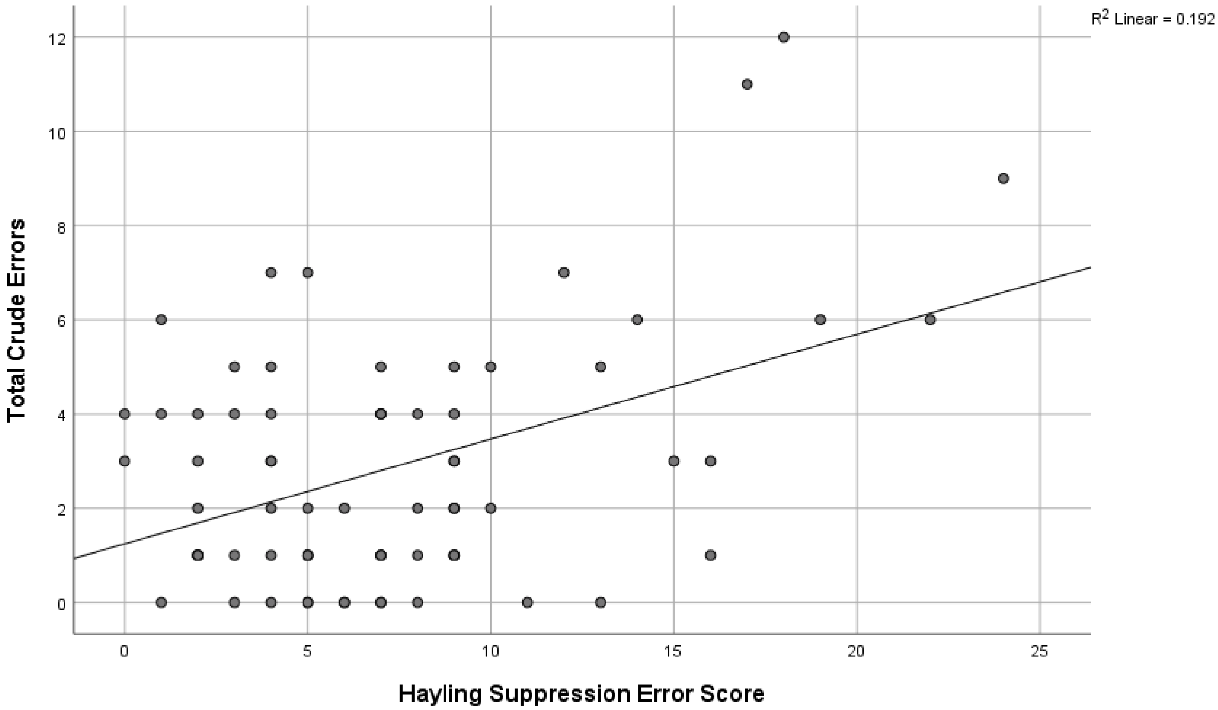

Total Crude Errors
Hayling Suppression Error Score
R2 Linear = 0.192
0
2
4
6
8
10
12
0
5
10
15
20
25

Figure 4.