# The relationship between empathy and executive functioning: An ERP analysis

**R. W. Springintveld**

SNR: 2043725

Bachelor's thesis

Cognitive Neuropsychology

Supervisor: W. De Baene

2nd reader: J. Stekelenburg

School of Social and Behavioral Sciences

Department of Cognitive Neuropsychology

Tilburg University

January, 2023

**Abstract**

Theoretical models suggest that executive functioning may regulate empathy, with inhibition in particular. Inhibitory control would act by suppressing one's own feelings in order to attain another's. Despite the fact that their correlation has been well established, this relationship has not been investigated with event-related potentials (ERP's) yet. This thesis used neurophysiological components to investigate whether differences in empathy were associated with differences in inhibitory control. Sixty participants (mean age= 21.1) were asked to complete the self-report questionnaire Empathy Quotient (EQ-scale) and perform the Stroop Colour and Word Task (SCWT), which showed coloured words in congruent and incongruent conditions. Continuous EEG data were recorded during the SCWT. Incongruent trials evoked significantly larger amplitude differences in N450 and the Late Positive Component (LPC), slower reaction times and higher error rates. However, an ANCOVA with empathy as covariate showed no significant effects of congruency in all four variables. Results suggest that empathy does not relate with inhibition. These results bring importance to assuming a relationship between executive functioning and empathy and emphasize that we should be careful to copy findings in psychopathology to healthy people. Future research could investigate whether these null-results can be replicated when differentiating between cognitive and affective empathy and controlling for gender and age.

*Keywords: executive functions, empathy, Stroop, inhibition, Event-related Potential (ERP), N450, Late Positive Component (LPC)*

## Introduction

The importance of empathy shows in its functions of assisting in social interactions like motivating parental care, enabling prosocial behaviour and inhibiting aggression (Decety, 2015). Children of 8 to 10 months already show moderate empathic concern in order to explore and understand distress in their surroundings (Hanania et al., 2011; Decety, 2021). Furthermore, empathy contributes to moral reasoning connected to caring for others and guilt (Decety, 2021; Decety & Howard, 2013)).

Empathy has been considered a multidimensional construct, but researchers have been debating about how many and which components it contains. Two, three or even four component have been identified in literature (Zaki, 2014; Decety, 2021), but most literature agrees on it being a two-sided construct containing an affective (emotional) and a cognitive (rational) component (Yan et al., 2020). For the purpose of this study, empathy is defined like Baron-Cohen & Wheelwright (2004) did: "*Empathy is the drive or ability to attribute mental states to another person/animal, and entails an appropriate affective response in the observer to the other person's mental state.*" (p. 168). Affective empathy is seen as an automatic, fast response associated with emotions and the feeling of an emotional experience of others whilst cognitive empathy is seen as a controlled, slow response, concerned with understanding what someone is thinking and taking their perspective (Zeng et al., 2021; see: dual process theory; Heyes, 2018 as well). Several theories hold that the medial prefrontal cortex is involved in integrating cognitive circuits with affective circuits that are distributed within the brain (Shamay-Tsoory et al., 2004; Decety, 2015).

Empathy is often measured via questionnaires. The Empathy Quotient (EQ) scale is one of those tests, invented by Baron-Cohen & Wheelwright (2004). This test has included both the cognitive and affective component of empathy, yet it cannot measure them independently. The EQ is easy to assess, score and interpret since it relies on a total score, has strong construct validity since it has been tested against various questionnaires and has been proven to be reliable over many cultures and populations (De Lima et al., 2021). Another popular questionnaire is the Interpersonal Reactivity Index (IRI), also proven to be reliable over various populations, yet this scale depends on four subscales measuring affective and cognitive empathy, empathy for fictional characters and personal distress, and cannot be scored with one dimension (Melchers et al., 2016), making it more difficult to interpret.

Executive functions include the ability to carry out goal-directed plans, goal formation, planning, effective performance and cognitive flexibility (Diamond, 2013). Extensive research has established that executive functions are one of the last developing functions in the brain and are the first to degenerate in senescence (Jurado & Rosselli, 2007). It was first theorized that executive functions were a higher order of the frontal lobe, now they have been found to be of essentiality for our daily life since they mediate the ability to organize our thoughts in a goal-directed way (Ardila, 2008). Inhibition is an important part of executive functioning, since it is concerned with social behaviour, learned behaviour and filtering information (Jurado & Rosselli, 2007). Inhibition can be interpreted as suppressing one process to originate another; often associated with the suppression of an automatic process to initiate another process that requires more cognitive capacities.

The Stroop task (Stroop, 1935) is often used to test EF and it focuses on inhibition. There are many variations of the Stroop task, involving images or sounds or words. The Stroop Colour and Word task (SCWT) involves words of different colours in different trials. In incongruent trials, the word does not match the colour it has, in congruent trials it does (see: Figure 1 for an example). Participants are asked to answer which colour the letters have, not which word they can read. Incongruent trials are thought to reflect the cognitive task that is required in inhibiting the initial and more automatic response of reading the word. The SCWT has been shown to exert inhibition in higher cognitive processes than only the inhibition of motor responses (Morooka et al., 2012). This inhibition can be seen in longer response times and higher error rates in incongruent trials, but is also reflected in certain ERP components (Sahar et al., 2022).

EEG research has identified ERP components of inhibition specifically. The N450 has been found to mark interference suppression, reflected in larger negativity at centro-parietal sites in incongruent trials compared to congruent trials between 325 and 465 ms after stimulus onset (Heidlmayr et al., 2020). This component is associated with activation in the ACC and PFC, both critical locations for EF (Egner & Hirsch, 2005; Heidlmayr et al., 2020). In the Stroop test specifically, this component reflects inhibitory control of stimulus and response conflict (Coderre et al., 2011; Ahumada-Méndez et al., 2022).

Another component, called the Late Positive Component (LPC) has been associated with inhibition as well. This component has shown to reflect response inhibition in specifically the Stroop task with more positivity in incongruent than congruent trials between 600 and 900 ms after stimulus onset (Coderre et al., 2011). This LPC shows most activity around centro-parietal electrodes. It must be noted that the LPC has been associated with semantic processing (Liu et al., 2014) and conflict resolution (Heidlmayr et al., 2020) as well. Yet, there is enough evidence to conclude that the LPC reflects response inhibition in healthy participants reliably (Liotti et al., 2000; Sucec et al., 2018; Guo et al., 2018).

Empathy has been found to refer to multiple closely linked mental states rather than one monolithic cognitive function, including but not limited to, subjective experience of emotions, self-awareness, the recognition of emotions in others and the capacity to anticipate an emotion in oneself and in others (Seitz et al., 2006). This last concept implies a regulation of own emotions to find an appropriate response to other’s emotions. This would require an inhibition of own emotional responses, cognitive flexibility to switch between own perspective and other's and working memory to maintain well- processing of emotional information, which are subcomponents of executive functioning (Firat, 2019). Another idea is that executive functions are the regulator of empathy and decide whether we are motivated by it (Decety et al., 2016).

Both the cognitive and the affective component of empathy can be theorized to link with EF because of the projections in the frontal lobe (cognitive) but also to the subcortical areas (affective) and their supposed integration in the medial frontal cortex (Shamay-Tsoory et al., 2004). The correlation between them has been well established. Yan et al. (2020) performed a meta-analysis of 18 papers containing 67 samples and tried to identify the correlation of the subcomponents of empathy and EF. They first found that the correlation between cognitive empathy and EF is stronger than with affective empathy, yet both were significant. Secondly, they found that of the EF components working memory, cognitive flexibility and inhibitory control, only inhibitory control significantly correlated to both cognitive and affective empathy. These findings corresponded to the theories that empathy depends on whether one can distinguish the awareness of self from other (inhibition of self-awareness), or that empathy is essential to regulate and reduce perceived distress (inhibitory control of own stress vs others) (Yan et al., 2020).

Zeng et al. (2021) tried to discover the causal relation of empathy and EF in children; They let 147 children from around five years old perform three tasks aiming at inhibition, working memory and cognitive flexibility. For the inhibition task, they used a Go/NoGo task. For the empathy task, children had to watch a video and answer to questions how they felt and thought the other person would feel . After one year, the researchers performed this experiment again on 74 of the 147 children and compared the results of different years with each other. They found that empathy was closer related to inhibition than to working memory or cognitive flexibility and that cognitive empathy was stronger correlated with executive functioning than affective empathy. They contributed this last result to the idea that affective empathy is more hereditary and animalistic, therefor automatic, than cognitive empathy which can inhibit or excite empathic responses (Zeng et al., 2021).

No literature has been identified that examined the relationship between inhibition and empathy using EEG. The utilisation of psychophysical measurement of the inhibitory process in the brain will give the opportunity to check whether inhibition is effectively active and might give a temporal dimension to our understanding of the relationship between inhibition and empathy. As the theory states that empathy requires a form of inhibition of one's own feelings to be able to take another's perspective first, this would mean that people who are better in inhibiting their own responses, would have higher measures in empathy than people who have less inhibitory control. Better inhibitory control would be reflected in higher amplitudes in the ERP components associated with inhibition. Additionally, investigation of the relationship has started from psychopathological concerns, for example whether people on the Autism spectrum would have less empathy. Development of empathy in children has also been highly associated with development of the EF areas. However, investigations of whether this relationship can be generalized to healthy adult populations remain scarce.

In order to deepen the relationship and examine whether results can be generalized to further populations, this thesis was conducted. The construct of empathy will be assessed with the Empathy Quotient (EQ) and inhibition will be measured by the Stroop Colour and Word test and the N450 and LPC component. The research question central to this thesis is whether differences in EQ scores correlate with performance on the Stroop task. In order to establish whether the desired Stroop interference effect is present, it was hypothesized; (H1) Incongruent trials will have slower reaction times and higher error rates than congruent trials. (H3) Incongruent conditions will have higher amplitudes in the N450 and LPC component than congruent conditions. Furthermore, (H2) higher EQ scores will correlate to faster reaction times and less errors and (H4) higher EQ scores will correlate with higher amplitudes in N450 and LPC components.

## Method

*Participants*

60 undergraduate students from Tilburg University were recruited (age range= 17-53, mean 21.1, men=13, women=46, unidentified=1; Native language Dutch= 29, Other= 31). Via the SONA-system, participants could enrol in the experiment in exchange for one course credit. The participants were all undergraduate psychology students. After reading the information letter, all participants gave consent prior to the experiment. The experiment was approved by the Ethical Review Board of the Tilburg School of Social and Behavioral Sciences.

*General description*

Participants first completed the Empathy Quotient (EQ-60) (Baron-Cohen & S. Wheelwright, 2004) through the online SONA system. After filling in the questionnaire, participants were briefed about the real focus of the scale. Afterwards, participants registered for the EEG experiment. To measure executive functioning/inhibition, the Stroop Colour and Word Task (SCWT) was deployed. During this task, EEG was recorded. The EQ took around 30 minutes to answer and the EEG recording took around two hours.

*Questionnaire*

To measure individual differences in empathy, the Empathy Quotient (EQ-60) (Baron-Cohen & Wheelwright, 2004) was deployed. This scale was originally developed to test ADHD patients but has been proven reliable for neurotypical participants (Allison et al., 2011). Sixty items of self-report statements were scored on a four-point Likert scale ranging from “strongly agree” to “strongly disagree”. The computer program randomized the questions. Forty items focusing on empathy were presented, where 19 items were aimed at a response of “disagree” and 21 items at a response of "agree" in order to avoid a response bias (Baron-Cohen & Wheelwright, 2004). In the questionnaire, cognitive and affective abilities of empathy were not separated. Those forty statements vary from "*I really enjoy caring for other people*" to "*I am able to make decisions without being influenced by people's feelings.*" To distract the participant from the purpose of measuring empathy, twenty filler questions were added with statements about general behaviour like "*I enjoy having discussions about politics''* or "*I like to take risks*" and the questionnaire was given the title "The Cambridge Behaviour Scale". Depending on the language of the participant, the English or Dutch version of the scale was employed (Translation: De Corte et al., 2006), which has been proven to be reliable and valid (Groen et al., 2016). The total score is the sum of all items, where higher scores indicate higher empathy. In psychometric analysis, the EQ has been found to be a valid measure of empathy (Muncer et al., 2006; Allison et al., 2011) and has been proven to rely on one dimension which supports the validity of using a total score of one scale (Allison et al., 2011). Interestingly, in most analyses an effect of gender has been found where women score significantly higher on the EQ than men (Muncer et al., 2006; Allison et al., 2011). However, since this thesis used a sample with relatively few men (13 out of 60), this effect will not be investigated as the results would not be reliable.

*Design*

Stimuli were presented in the centre of a computer screen (1920px X 1080px) located at a distance of approximately 60 cm. A QWERTY computer keyboard was placed directly in front of the screen and the participant.

*Task description*

The experimental stimuli were shown on a grey background (code: #aaaaaa) and included 16 combinations of word and colour. The words "RED", "GREEN", "YELLOW" and

“BLUE” were shown in colours indicated by the word (congruent condition) or a mismatched colour (incongruent condition). The colours were identified with RGB values red(#ff0000), green(#00ff00), blue(#0000ff), yellow(#ffff00) with font size 100. Participants were asked to identify the colour of the word by pressing dedicated buttons on the keyboard; ‘z’ for red, ‘x’ for green, ‘n’ for blue and ‘m’ for yellow. The trials were randomly assessed. Each trial started with a fixation display for 1000 ms and was followed by a congruent or incongruent stimulus for 200 ms (see: Figure. 1). Then a fixation dot appeared and the participants were told to respond “as fast as possible but accurate” and to indicate the correct colour of the word by pressing a button. If the key that they pressed resembled the correct colour, their response was considered correct, yet participants did not receive feedback of their responses. If one of the dedicated buttons was pressed, an ITI randomly chosen between 1000-1500 ms appeared before the start of the next trial. Between blocks, an image of a keyboard appeared and participants were allowed to take a break until they pressed a key to continue. A block consisted of 48 congruent trials (4 words in the same colour repeated 12 times) and 48 incongruent trials (12 combinations repeated 4 times) and stimuli were presented in random order. The experiment consisted of three blocks of 96 trials in total and was designed using Open Sesame software (Mathôt et al., 2012).

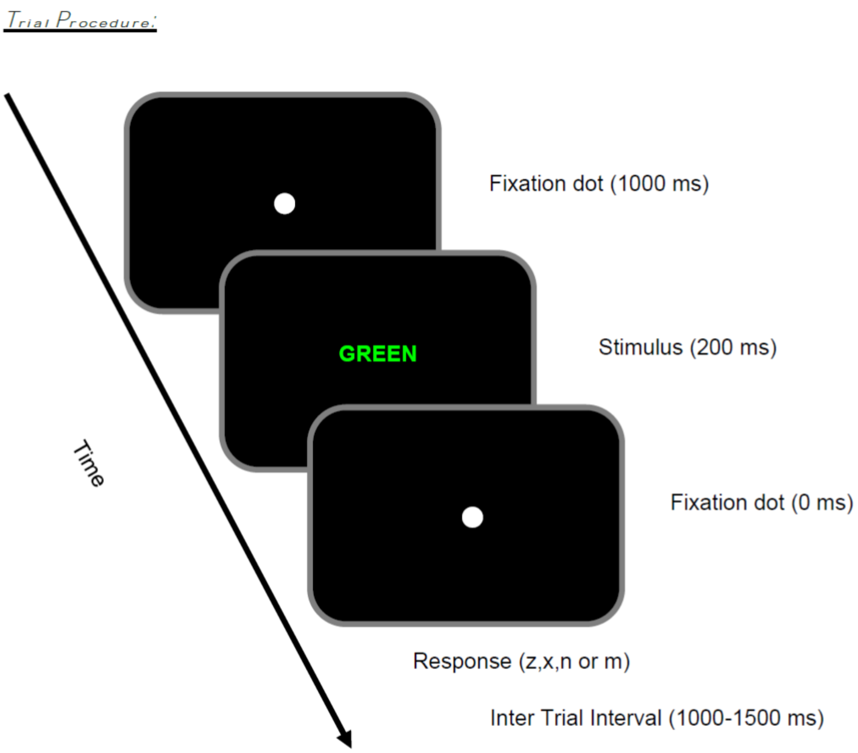


*Figure 1: Test trial Stroop Task*

*Procedure*

Participants were individually tested by being seated in a sound attenuated and dimly lit cabin of approximately 4 square meters and were observed via CCTV. A desk with a computer screen was situated next to a small cabinet where the amplifier was placed on. For efficiency of bachelor's thesis' lab hours, participants performed three other tasks (IAPS, CPT, resting state) irrelevant to this study after performing the Stroop-task. Participants were asked to sit as still as possible whilst performing the Stroop-task.

Before the real experiment started, participants were presented with two practice sessions. In the Colour-response mapping session, participants were trained to remember which keys corresponded to which colours. This was done to prevent confounds of forgetting the correct response key in the real experiment. Stimuli from the Stroop-task were attenuated to show only 'X's instead of words; so *"XXXX*" in green letters corresponds to the response of pressing 'x' and *"XXXXXXXX"* in red letters corresponds to pressing 'z'. This resulted in 4 unique stimuli that were presented each 25 times, 100 trials in total. Participants were given feedback on their response between the trials in this practice session, in order to correct for their incorrect responses. The second practice session was a duplicate of the actual experiment except that it consisted of 24 trial (12 congruent, 12 incongruent). This was done to familiarize participants with the real experiment that followed directly after. Resembling the actual experiment, participants were not given feedback on their responses.

*Neurophysiological measurements*

To measure brain activity during the Stroop-task, EEG signals were recorded from a BioSemi Active-Two amplifier (BioSemi, Amsterdam, The Netherlands) system using 32 active Ag/Cl electrodes positioned according to the international 10-20 system and mounted in an elastic cap. EEG signals were amplified in 100mV and digitized with a sample rate of 512 Hz. During the measurement, the active Common Mode Sense (CMS) located between Cz and C3 served as the online reference and the passive Driven Right Sense (DRL) located between Cz and C4 served as the ground electrode. Two extra electrodes serving as an offline reference were placed on the mastoids. To detect eye movements, four additional EOG electrodes were placed around the eyes. The electrodes above and underneath the left eye in vertical line with the pupil, detected vertical eye movement electrical activity, whilst two electrodes near the outer canthi of both eyes, in horizontal line with the pupil, detected horizontal eye movements. To detect heart rate, participants were instructed to place two ECG electrodes on their torso: one in the middle of the sternum ("The start of your ribcage, then two fingers up") and the other on the left side of the ribcage under the armpit at the same height as the other sternum electrode. ECG signals were not analysed.

*Preprocessing*

Preprocessing analysis was performed with the program Brain Vision Analyzer 2.1. The electroencephalogram was re-referenced offline to the average of the left and right mastoids and bandpass and notch- filtered to eliminate noise and 50Hz interference (0.01-30 Hz, 50Hz notch). Then, continuously recorded EEG signals were segmented into 1100 ms epochs relative to stimulus onset (this included a 100 ms prestimulus baseline). Then the segments were corrected for eye movements using a procedure of Gratton & Coles (1983) since EOG might corrupt true EEG signals. Epochs were rejected if within that epoch at any EEG channel the difference between the most positive and negative value was more than 200 micro–V. Epochs were averaged separately per condition and baseline corrected (-100 to 0 ms).

*Statistical analysis*

The statistical analysis was focused on differences in reaction times, error rates and the difference between amplitudes of the ERP components identified in the Stroop task. To calculate RT and error rate, response times and percentage of correct responses in relation to total responses were averaged over all participants. To calculate the components, ERP data over all electrodes between 325 and 465 ms (for N450) and separately between 600 and 900 ms (LPC) was exported to SPSS. These components were at a topographical maximum at the Pz, PO3, PO4 electrodes, which coincides with earlier findings of N450 and LPC at left centroparietal- and centroparietal-electrodes respectively (Coderre et al., 2011). Therefore, an average over those electrodes in both conditions was separately calculated for N450 and LPC data. A Repeated Measures ANOVA followed by a Repeated Measures ANCOVA was performed in SPSS version 2018. A RM ANOVA instead of independent sample t-test was chosen since a covariate (‘EQ score’) can be easily added to an ANOVA making it an ANCOVA.

## Results

Two out of sixty participants were excluded from analysis since they did not have any EEG data. The assumptions of the Repeated Measures ANCOVA, independence, normality and sphericity and linearity were not violated. The participants were randomly assessed, assuming independence of measures. Concerning normality, five out of seven variables followed an abnormal distribution, what would have violated the assumption of normal distribution if the sample size would have been under thirty (Blanca et al., 2017). Mauchly's test was not significant, assuming sphericity and the correlation between inhibition and empathy has shown to be linear (Yan et al., 2020).

*Behavioral data*

Response times (RT) and error rates for the Stroop task are presented in Figure 2. Error rate and RT data in both incongruent ($r= 0.218$, $p=.0097$) and congruent ($r=0.249$, $p=.057$) conditions were uncorrelated, indicating that speed-accuracy trade-off did not influence participant performance. A repeated measures ANOVA with one factor "congruency" (congruent vs incongruent), showed a significant Stroop RT-interference effect, $F(1,57)=21.43$, $p<.001$. Likewise, Stroop error rate interference was found in statistically higher error rates in incongruent conditions, $F(1,57)=29.93$, $p<.001$.

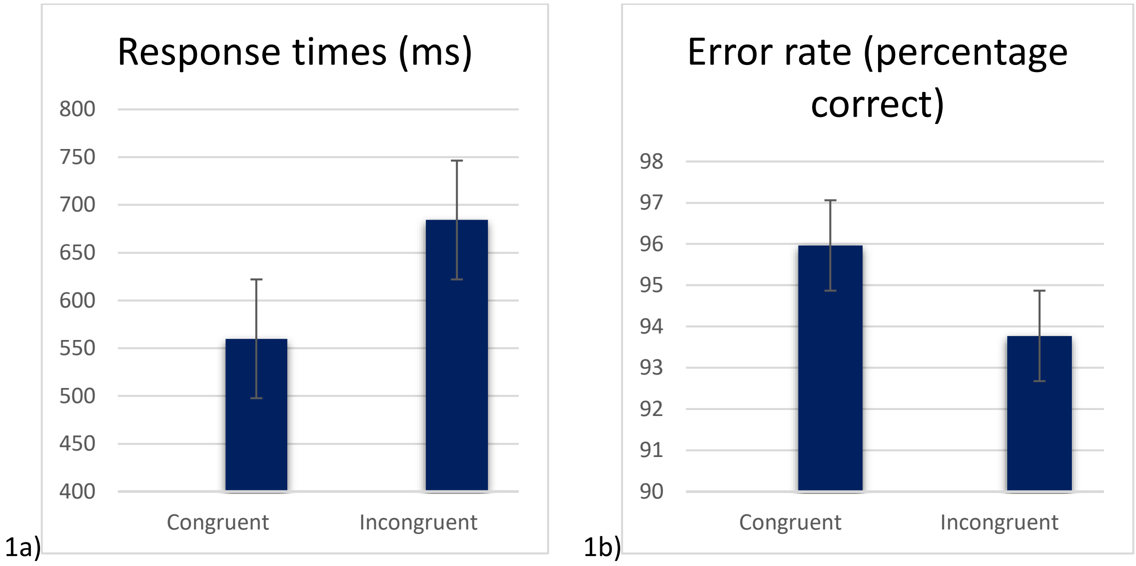


*Figure 2a. Response times (Milliseconds). Mean (SD); Congruent:559.87 (206.46), Incongruent: 684.21(247.32)*

*Figure 2b. Error rates (Percentage correct). Mean (SD); Congruent: 95.97 (3.16), Incongruent: 93.77 (4.50). Smaller error rate indicates more mistakes.*

*Empathy Quotient data*

Firstly, responses on the EQ-questions were scored by hand according to the scoring sheet where answers are coded with a 1, 2 or 0 depending on range of agreement and the question being inverse or a filler question. Total scores showed an average of 46.40 (SD=11.11), with a minimum of 19 and a maximum score of 56. There were no outliers present.

For illustration purposes (see Figure 3), the scores were centred and divided into three groups; around the mean ("normal"), 2 standard deviations below the mean ("low empathy")and 2 standard deviations above the mean ("high empathy"). The low- and high empathy group were used for the calculation and illustration of the ERP differences between conditions.

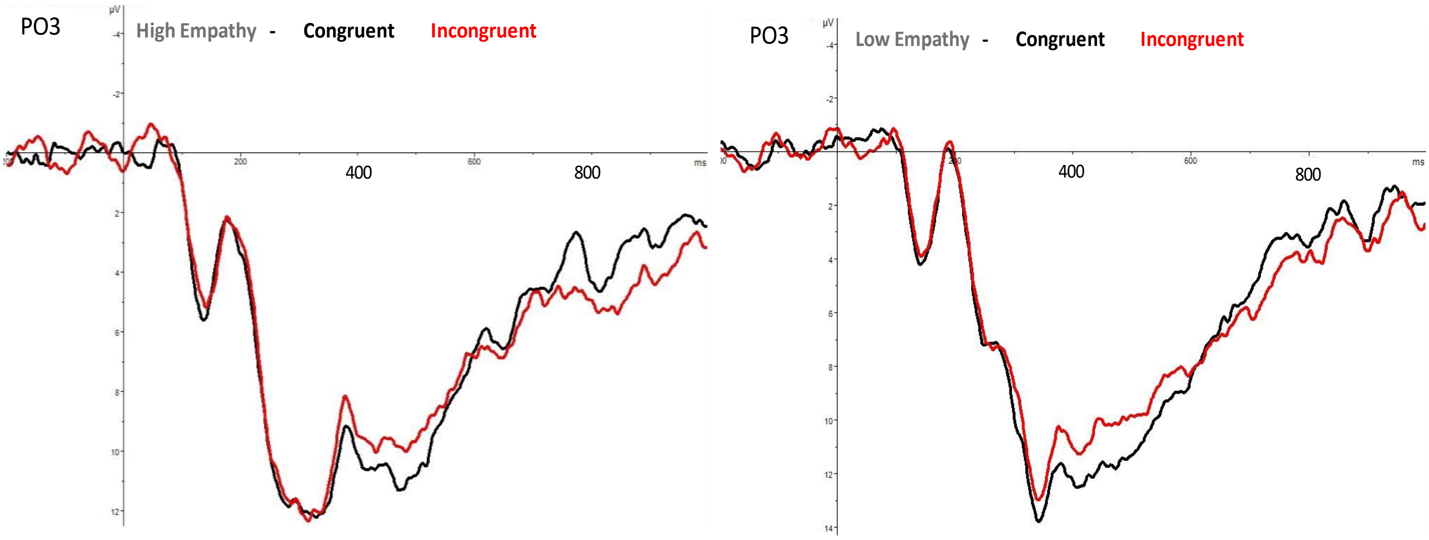


*Figure 3. Stimulus locked ERP waves on the PO3 electrode for (left) 'High' (+2sd) and (right) 'Low' (-2sd) empathy between the congruent (black) and incongruent (red) condition.*

*EEG data*

The number of trials per participant included in the Grand Average contained an average of 131.96 (sd= 15.52) for congruent trials, and an average of 132.67 (sd=14.95) for incongruent trials. Grand Average ERP waveforms and voltage maps for congruent and incongruent trials reflecting the N450 and LPC in the PO3 electrode are shown in Figure 4. Stimulus-locked ERPs averaged across centro-parietal electrodes (PO3, PO4 and Pz) showed a significant negativity of the incongruent condition around 400 ms and a significant positivity of the incongruent condition between 600 and 900 ms; the occurrence of the N450 and LPC. The amplitude difference between conditions was significant for both the N450, $F(1,57)= 12.09$, $p<.001$ and the LPC, $F(1,57)=5.085$, $p=.028$.

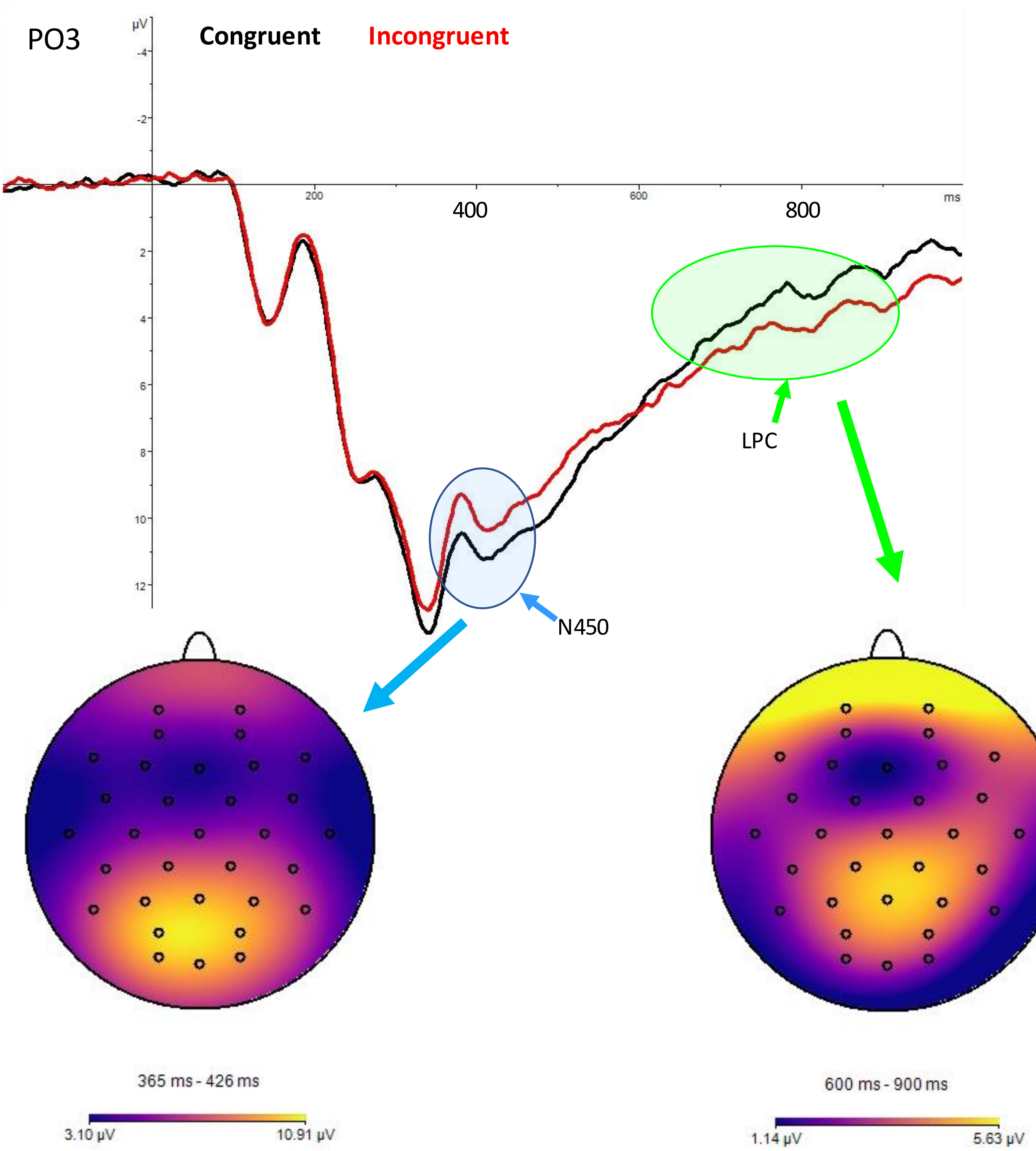


*Figure 4: Stimulus locked ERP waves between congruent (black) and incongruent (red) conditions with voltage maps for N450 and LPC. The N450 (325-465 ms) shows greater negativity in the incongruent condition, the LPC (600-900 ms) is identified with more positivity in the incongruent condition.*

Zero-order Pearson correlations between all variables showed no significant correlations (see correlation Table 5) except between the congruencies. A repeated measures ANCOVA with within-subject factor ‘congruency’ (congruent vs incongruent) and covariate EQ revealed no significant effects. In all four analyses of RT, error rate and both ERP components, adding the covariate EQ total score made the main effects of congruency as covered above insignificant (F(1, 56)=.000, p=1.00; F(1, 56)=.000, p=1.00; F(1, 56)=.077, p=.738 for N450; F(1,56)=3.103, p=.084 for LPC).

EQ main effect turned out to be insignificant as well for all variables RT (F(1,56)=.98, p=.33), Error rate (F(1, 56)=.000 p<.99) amplitude for N450 (F(1, 56)=.54, p=.47) and for LPC (F(1, 56)=.057, p=.81). Interactions between the variables RT, error rate and amplitude of N450 and LPC and the covariate EQ were not significant (F(1, 56)=1.19, p=.28; F(1, 56)=.077, p=.738; F(1, 56)=1.139, p=.290; F(1,56)=1.629, p=.207 respectively) See Table 6 for a summary of the ANOVA and ANCOVA results.

*Table 5: Zero-order correlations between measures of empathy and behavioral and Event-related Potentials*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1. EQ_total score | 1 | | | | | | | | |
| 2. RT_Congruent | .058 | 1 | | | | | | | |
| 3. RT_Incongruent | .167 | .606** | 1 | | | | | | |
| 4. ER_Congruent | .020 | .241 | .211 | 1 | | | | | |
| 5. ER_Incongruent | -.011 | .203 | .213 | .736** | 1 | | | | |
| 6. N450 incongruent | -.072 | .101 | .068 | -.015 | .131 | 1 | | | |
| 7. N450 congruent | .013 | .153 | .080 | -.066 | .011 | .870** | 1 | | |
| 8. LPC incongruent | -.122 | -.134 | -.045 | -.164 | -.073 | x | x | 1 | |
| 9. LPC congruent | -.072 | -.124 | -.065 | -.181 | -.114 | x | x | .952** | 1 |

*Note:* EQ= Empathy Quotient; RT= Reaction times; ER= Error rate; LPC= Late Positive Component. ** p< 0.01 level (2-tailed)

*Table 6: ANOVA and ANCOVA results*

| | Congruency | | F | p- value |
|---|---|---|---|---|
| ANOVA | | RT | 21.43 | <.001 |
| | | Error rate | 29.93 | <.001 |
| | | N450 | 12.09 | <.001 |
| | | LPC | 5.09 | 0.03 |
| ANCOVA | RT | Main effect RT | 0.00 | 1.00 |
| | | Main effect EQ | 0.98 | 0.33 |
| | | Interaction | 1.19 | 0.28 |
| | ER | Main effect ER | 0.00 | 1.00 |
| | | Main effect EQ | 0.00 | 0.99 |
| | | Interaction | 0.08 | 0.78 |
| | N450 | Main effect N450 | 0.06 | 0.81 |
| | | Main effect EQ | 0.54 | 0.47 |
| | | Interaction | 1.14 | 0.29 |
| | LPC | Main effect LPC | 3.10 | 0.84 |
| | | Main effect EQ | 0.06 | 0.81 |
| | | Interaction | 1.63 | 0.21 |

*Note:* EQ= Empathy Quotient; RT= Reaction times; ER= Error rate; LPC= Late Positive Component. Congruency (within-subjects factor), EQ (covariate).

## Discussion

The purpose of this study was to further examine the relationship between executive functions and empathy. This study tried to contribute to literature by demonstrating a correlational effect between empathy and inhibition with the use of EEG measurements, across healthy adults. Inconsistent with our hypotheses, no effect between empathy and reaction times, error rate and ERP components turned out to be significant. The association between empathy and executive functioning has been well established (Yan et al., 2020). This study can still be relevant as it shows a null result to an accepted relationship.

The first hypothesis expected a Stroop-task interference effect, meaning that reaction times and error rates in incongruent trials are higher than in congruent trials. Data showed that this indeed was the case, indicating that the desirable inhibition of reading over colour naming was present. Additionally, response times and error rate did not correlate, indicating that no trade-off between speed and the wrong response was present. Stroop interference has been mostly measured with behavioural data like reaction speed and error rate. Yet, Scarpina & Tagini (2017) bring to attention that in the original Stroop, error rates and speed were not independently calculated, by computing error rates in relation to mean reaction time per item. Why this is important goes beyond the scope of this thesis, but it explains that only error rates - and not the reaction times - can be used as an index of inhibitory control and researchers should be careful in directly interpreting RT as Stroop interference. Additionally, another common use of the Stroop task, in neuropsychological assessment, is the measurement of more executive functions since it has been associated with working memory and cognitive flexibility as well (Periáñez et al., 2021).

The second hypothesis included a relation between the EQ total score and the Stroop interference effect. However, the interference effect of both response times and error rates disappeared after including EQ scores in the analysis. These results reflect that EQ scores explain a part of the variance in speed and error rates over conditions, yet this part is too insignificant to be classified as an effect. EQ main effects for both error rate and RT were insignificant as well, meaning that EQ has no influence on RT or error rates over the conditions. Interaction effects were not significant as well, meaning that performance in either conditions was not dependent on different levels of EQ. These results imply that no effect between empathy and inhibition has been found, which contradicts existing literature. Zeng et al., (2021) did find that more inhibition correlated with higher empathy, yet they used the Go/NoGo task instead of the SCWT, and the Vignette story telling task instead of an empathy questionnaire.

Morooka et al. (2012) found that the Stroop and Go/NoGo task tap different aspects of response inhibition since the Stroop task involves more target cues (i.e. four keys for the correct response compared to one). This would require more executive processes than the Go/NoGo task. However, with that idea, this thesis should have found even stronger relations between Stroop interference and empathy than Zeng et al. (2021) found with Go/NoGo interference. Zeng et al. (2012) also used the Vignette story task - where participants have to react to stories how they felt and how they think the people in the story would feel - as a more active measurement of empathy than a questionnaire, requiring elaborate explanation of feelings towards oneself and another mostly involving cognitive empathy. Since inhibition is associated more closely to cognitive than affective empathy (Yan et al., 2020) it could be that the results of Zeng et al. (2012) could not be replicated by this thesis since it did not make that distinction.

The third hypothesis expected higher amplitudes in the N450 and LPC component of the response locked ERPs in incongruent conditions. As the results and the difference wave showed, amplitudes in incongruent trials were indeed significantly more negative in incongruent trials between 325 and 465 ms and 600 to 900 ms. These results show that the Stroop interference effect is significantly visible in the EEG data and that inhibition is marked by these components. These results are in line with existing literature about the components of inhibition and the Stroop Colour and Word Task. Even though the N450 has been identified as a marker of response conflict (Ahumada-Méndez et al., 2022) and additionally rather conflict detection than conflict resolution (Coderre et al., 2011), Heidlmayr et al. (2020) showed that specifically for the Stroop task the N450 is interpreted to reflect inhibitory processes and suppression of interference. However, it must be noted that Heidlmayr et al. (2020) considered the N400 and the N450 in the Stroop task as the same component while the N400 has been identified in literature in general as a separate component associated with conflict processing of semantics measured with tasks involving sentence manipulations and high-level expectancy

(Ahumada-Méndez et al., 2020). The LPC has been linked to conflict resolution or response selection as well or it could also involve semantic processing since it is prominent around Wernicke's area (Coderre et al., 2011).

Fourthly, it was hypothesized that a higher amplitude difference in the N450 and LPC component would correlate with higher EQ scores. However, similarly to the behavioural data, the main effect of amplitude in the N450 difference between conditions disappeared after including EQ scores in the analysis, indicating that EQ scores explain a part of the variance in N450 interference over conditions, yet too insignificant to be a statistical effect. The same can be said for LPC amplitudes, where the results were similar. EQ main effects were insignificant for both N450 and LPC, indicating that EQ scores did not influence the amplitude of N450 nor LPC over conditions. Lastly, no interaction effects between N450 and EQ or LPC and EQ were significant. These result imply that no effect of empathy on inhibition has been found, which contradicts existing literature. No previous studies were identified that examined the relation between empathy and inhibitory control via ERP components of inhibition. Yet, regarding the established relationship between the N450/LPC and inhibition in the Stroop task and the relationship between inhibitory control and empathy measures (see: Introduction) , our findings should have been in line with existing literature that assumes a relationship between inhibitory processes and empathy. Besides the use of other tasks for empathy and inhibition (as mentioned above with H2), age could also be a factor in explaining why no relationship has been found here. Zorza et al. (2019) found that inhibitory control could not predict empathy in adolescents and connected this to evidence that shows that regulation of emotion in childhood depends on behavioural inhibition while in adolescents it depends on far more complex processes like reappraisal of a situation. If this would be true, it might explain why this thesis could not replicate the relationship established by for example Zeng et al. (2021), since that study (and most literature with a healthy sample) was focused on the development of empathy and the relationship with executive functions in childhood.

Additionally to specific differences in method as discussed above, more general possible explanations for the null results in this study will be discussed here. Firstly, these null results might be due to a lack of variance in empathy scores. Even though data and visible analysis of scatter plots did reveal this sample variance as reliable, the values in the population might fluctuate more, as this sample included participants that were WEIRD (Clancy & Davis, 2019), around the same age and all psychology students. Manipulation of empathy via empathic stories like the Vignette story telling task might be more effective in producing significant results.

Secondly, it should be noted that much literature about executive functions in combination with empathy is concerned with neuroclinical populations; often with psychopathy, AD(H)D, Autism Spectrum Disorder and Schizophrenia. Although the importance of empathic development in these groups is much higher, it is mostly concerned with absence or a dysfunction in empathy. In the current participant group, most empathy scores were considered to be between average and higher than average. If the influence of inhibitory processes only becomes significant in low levels of empathy, it might explain why this thesis found no correlation.

Thirdly, whilst performing literature searches, familiarity came with a phenomenon also described by Ahumada-Méndez et al. (2022); crisis in replicability of ERP results. Lack of specification of software or procedure, inconsistency in component names (see: discussion of H3) and no standards for comparison of results, makes it difficult to find uniformity in the results of ERP analysis. For example, the "LPC" has various names (Late Positive Component (Coderre et al., 2011; Late Positive Complex (Liotti et al., 2000); Conflict Slow/Sustained Potential (Larson et al., 2009) in literature and is identified within different time windows (i.e.

variating from '> 350 ms' to 'between 500 and 1000 ms') depending on task and electrodes and should not be confused with the 'Late Slow Wave' associated with belief recognition and left frontal activation (Jiang et al., 2016; Bricker, 2020), what makes it difficult to draw consistent lines throughout literature.

Lastly, it could be that this study has unknown differences that moderate effect sizes or that this study produces a false negative. These possibilities are common in psychology and are reasons for the current replication crisis (Open science collaboration, 2015). Nevertheless, publication bias, the favouring of significant effects being published over null results, might be why these null results are inconsistent with published literature.

Further limitations of this study apply to method and experiment conditions. Firstly, participants were asked to fill in the EQ questionnaire in combination with seven other questionnaires, possibly resulting in fatigue and boredom. Because of this, it could be that the EQ scores did not 100% accurately reflect empathy levels in participants. Secondly, the Empathy Quotient is a self-report, prone to response bias and dependent on the level of self-reflection of the participant. However, the most extreme scores were between 2 and 3 standard deviations from the mean, showing that there were no outliers. Furthermore, this study was solely correlational and thus no causal conclusions can be drawn. Strengths of this study include the examination of young adults rather than children or neuropsychological patients. It is also the first known study that combines ERP analysis with inhibition and empathy in this way. Another strength is the use of the EQ scale, which is easy to assess and proven to be reliable over various populations and cultures. Since this study used international students as well, this was important.

The results show no significant relation between empathy and executive functioning. Yet a part of variance in the variables is possibly explained by empathy scores. However, no conclusions can be drawn from this, since this was not included in the analysis. Additionally, causal relations could be explored by an onsite manipulation of empathy. Furthermore, next studies could distinguish between affective and cognitive empathy in testing the relationship with executive functions, especially since cognitive empathy has been associated more with inhibition than affective empathy. This study should be replicated with a broader sample to see whether null results are replicable. In order to establish whether these null results in a healthy sample are of significant value against the publication bias, it should be investigated whether age is an important factor in the relationship between empathy and inhibition like Zorza et al. (2019) hypothesized. Future studies should also include gender effects on empathy. Lastly, it could be interesting to investigate whether the study psychology has an influence on levels of empathy. As there is certain evidence that medical students might decrease empathy in their education trajectory since medicine involves a distinction of oneself to the patient (Quince et al., 2016), psychology students might develop more empathy in their education trajectory that involves understanding other people.

Implications of this study expand beyond adding knowledge to the existing literature. This study was performed with healthy young adults and showed null results. Since the relationship between empathy and executive functioning is often established in abnormal psychology, null results in healthy samples might indicate that caution should be held when generalizing results to the healthy population. This study also came across important gaps in literature concerning disagreement on ERP analysis with the LPC component specifically. Additionally, these results could also show that it is important in assessing certain functions to empathy, that consideration of measuring cognitive and affective empathy independently should be taken seriously. Importantly, if these null results would be replicated, they should be likely appreciated as significant results in order to establish relations between cognitive functions as accurately as possible.

In conclusion, these results contradict the hypothesis that empathy is related to inhibitory control in healthy adults. Specifically, lower RT, lower error rates and higher N450 and LPC amplitudes have no significant relation to higher nor lower EQ scores. These results allow us to better understand the relationship between empathy and inhibition since they expand beyond existing literature with neurophysiological insights, and show that caution should be held in generalizing relationships from neuroclinical settings to healthy populations. They also emphasize the importance of agreement and reproducibility in the field of psychological experiments, specifically with an EEG setup.

## References

Ahumada-Méndez, F., Lucero, B., Avenanti, A., Saracini, C., Muñoz-Quezada, M. T., Cortés-Rivera, C., & Canales-Johnson, A. (2022). Affective modulation of cognitive control: A systematic review of EEG studies. *Physiology &Amp; Behavior*, *249*, 113743. https://doi.org/10.1016/j.physbeh.2022.113743

Allison, C., Baron-Cohen, S., Wheelwright, S., Stone, M., & Muncer, S. (2011, November). Psychometric analysis of the Empathy Quotient (EQ). *Personality and Individual Differences*, *51*(7), 829–835. https://doi.org/10.1016/j.paid.2011.07.005

Ardila, A. (2008). On the evolutionary origins of executive functions. *Brain and Cognition*, *68*(1), 92–99. https://doi.org/10.1016/j.bandc.2008.03.003

Baron-Cohen, S., & Wheelwright, S. (2004, April). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. https://doi.org/10.1023/b:jadd.0000022607.19833.00

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017b). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Clancy, K. B., & Davis, J. L. (2019). Soylent Is People, and WEIRD Is White: Biological Anthropology, Whiteness, and the Limits of the WEIRD. *Annual Review of Anthropology*, *48*(1), 169–186. https://doi.org/10.1146/annurev-anthro-102218-011133

Coderre, E., Conklin, K., & van Heuven, W. J. (2011). Electrophysiological measures of conflict detection and resolution in the Stroop task. *Brain Research*, *1413*, 51–59. https://doi.org/10.1016/j.brainres.2011.07.017

Decety, J. (2015). The neural pathways, development and functions of empathy. *Current Opinion in Behavioral Sciences, 3*, 1–6. https://doi.org/10.1016/j.cobeha.2014.12.001

Decety, J., Bartal, I. B. A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150077. https://doi.org/10.1098/rstb.2015.0077

Decety, J., & Howard, L. H. (2013). The Role of Affect in the Neurodevelopment of Morality. *Child Development Perspectives*, *7*(1), 49–54. https://doi.org/10.1111/cdep.12020

De Corte, K., Uzieblo, K., Buysse, A., & Crombez, G. (2006). Authorized Dutch translation of the empathy quotient. *Unpublished manuscript, Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium*.

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135.

Egner, T., & Hirsch, J. (2005). The neural correlates and functional integration of cognitive control in a Stroop task. *NeuroImage*, *24*(2), 539–547. https://doi.org/10.1016/j.neuroimage.2004.09.007

Estimating the reproducibility of psychological science. (2015). *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Firat, R. B. (2019). Opening the “Black Box”: Functions of the Frontal Lobes and Their Implications for Sociology. *Frontiers in Sociology, 4*. https://doi.org/10.3389/fsoc.2019.00003

Gratton, G., Coles, M. G., & Donchin, E. (1983, April). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Groen, Y., Fuermaier, A. B. M., Den Heijer, A. E., Tucha, O., & Althaus, M. (2015, April 25). The Empathy and Systemizing Quotient: The Psychometric Properties of the Dutch

Version and a Review of the Cross-Cultural Stability. *Journal of Autism and Developmental Disorders*, *45*(9), 2848–2864. https://doi.org/10.1007/s10803-015-2448-z

Guo, X., Li, Y., Zhang, Q., Cui, L., & Wei, P. (2018). Weakening of recognition memory caused by task-irrelevant emotional encoding context can be modulated by individuals' inhibitory control. *Personality and Individual Differences*, *134*, 201–209. https://doi.org/10.1016/j.paid.2018.06.012

Heidlmayr, K., Kihlstedt, M., & Isel, F. (2020, December). A review on the electroencephalography markers of Stroop executive control processes. *Brain and Cognition*, *146*, 105637. https://doi.org/10.1016/j.bandc.2020.105637

Heyes, C. (2018). Empathy is not in our genes. *Neuroscience &Amp; Biobehavioral Reviews*, *95*, 499–507. https://doi.org/10.1016/j.neubiorev.2018.11.001

Jurado, M. B., & Rosselli, M. (2007). The Elusive Nature of Executive Functions: A Review of our Current Understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Larson, M. J., Kaufman, D. A., & Perlstein, W. M. (2009). Neural time course of conflict adaptation effects on the Stroop task. *Neuropsychologia*, *47*(3), 663–670. https://doi.org/10.1016/j.neuropsychologia.2008.11.013

Lima, F. F. D., & Osório, F. D. L. (2021). Empathy: Assessment Instruments and Psychometric Quality – A Systematic Literature Review With a Meta-Analysis of the Past Ten Years. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.781346

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, *38*(5), 701–711. https://doi.org/10.1016/s0028-3932(99)00106-2

Liu, H., Rossi, S., Zhou, H., & Chen, B. (2014). Electrophysiological Evidence for Domain-General Inhibitory Control during Bilingual Language Switching. *PLoS ONE, 9*(10), e110887. https://doi.org/10.1371/journal.pone.0110887

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). *Open Sesame* [Software].

Melchers, M. C., Li, M., Haas, B. W., Reuter, M., Bischoff, L., & Montag, C. (2016). Similar Personality Patterns Are Associated with Empathy in Four Different Countries. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00290

Morooka, T., Ogino, T., Takeuchi, A., Hanafusa, K., Oka, M., & Ohtsuka, Y. (2012). Relationships between the color-word matching Stroop task and the Go/NoGo task: toward multifaceted assessment of attention and inhibition abilities of children. *Acta Medica Okayama, 66*(5), 377–386. https://doi.org/10.18926/amo/49385

Muncer, S. J., & Ling, J. (2006). Psychometric analysis of the empathy quotient (EQ) scale. *Personality and Individual Differences, 40*(6), 1111–1119. https://doi.org/10.1016/j.paid.2005.09.020

Periáñez, J. A., Lubrini, G., García-Gutiérrez, A., & Ríos-Lago, M. (2020). Construct Validity of the Stroop Color-Word Test: Influence of Speed of Visual Search, Verbal Fluency, Working Memory, Cognitive Flexibility, and Conflict Monitoring. *Archives of Clinical Neuropsychology, 36*(1), 99–111. https://doi.org/10.1093/arclin/acaa034

Quince, T., Thiemann, P., Benson, J., & Hyde, S. (2016). Undergraduate medical students; empathy: current perspectives. *Advances in Medical Education and Practice, Volume 7*, 443–455. https://doi.org/10.2147/amep.s76800

Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development, 34*(3), 447–458. https://doi.org/10.1016/j.infbeh.2011.04.007

Sahar, N. S., Safri, N. M., & Zakaria, N. A. (2022). Use of EEG Technique in a Cognitive Process Study-A Review. *ELEKTRIKA-Journal of Electrical Engineering, 21*(2), 24-32

Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00557

Seitz, R. J., Nickel, J., & Azari, N. P. (2006). Functional modularity of the medial prefrontal cortex: Involvement in human empathy. *Neuropsychology*, *20*(6), 743–751. https://doi.org/10.1037/0894-4105.20.6.743

Shamay-Tsoory, S. G., Tomer, R., Goldsher, D., Berger, B. D., & Aharon-Peretz, J. (2004). Impairment in Cognitive and Affective Empathy in Patients with Brain Lesions: Anatomical and Cognitive Correlates. *Journal of Clinical and Experimental Neuropsychology*, *26*(8), 1113–1127. https://doi.org/10.1080/13803390490515531

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

Sucec, J., Herzog, M., Van Diest, I., Van den Bergh, O., & von Leupoldt, A. (2018). The impairing effect of dyspnea on response inhibition. *International Journal of Psychophysiology*, *133*, 41–49. https://doi.org/10.1016/j.ijpsycho.2018.08.012

Yan, Z., Hong, S., Liu, F., & Su, Y. (2019). A meta-analysis of the relationship between empathy and executive function. *PsyCh Journal*, *9*(1), 34–43. https://doi.org/10.1002/pchj.311

Yan, Z., Zeng, X., Su, J., & Zhang, X. (2021). The dark side of empathy: Meta-analysis evidence of the relationship between empathy and depression. *PsyCh Journal*, *10*(5), 794–804. https://doi.org/10.1002/pchj.482

Zaki, J. (2014). Empathy: A motivated account. *Psychological Bulletin*, *140*(6), 1608–1647. https://doi.org/10.1037/a0037679

Zeng, X., Yang, L., Zhu, X., Zhou, K., Zhang, J., & Yan, Z. (2021). Preschoolers' executive function boots the development of empathy: one-year cross-lagged panel analysis. *Early Child Development and Care*, *192*(13), 2175–2183. https://doi.org/10.1080/03004430.2021.1990908