Moral ambiguity: how empathy and psychopathy inform moral decision making – a two study approach

James Twose

s4721152

Master Science Thesis

Supervised by:
1. Maaike Cima
2. Fred Hasselman

26th June 2018

Radboud University

Moral ambiguity: how empathy and psychopathy inform moral decision making - a two study approach

James Twose

Radboud University

Word count: 6443

Author Note

James Twose, Behavioural Science Institute, Radboud University, Nijmegen. Student number: s4721152

This research was conducted in fulfilment of the requirements for the completion of SOW-BS28A: Major Research Project, part of the Behavioural Science Research Master programme of Radboud University, Nijmegen.

Heartfelt thanks to Dr. Maaike Cima and Dr. Fred Hasselman for their supervision throughout the development, execution and evaluation of this research. Furthermore, I would like to thank Dr. Marije Keulen-de Vos for her supervision at de Rooyse Wissel forensisch psychiatrisch centrum, Venray. Lastly, I would like to thank Laura van der Helm BSc for her help with data collection throughout the project.

Correspondence concerning this research may be addressed to James Twose, Behavioural Science Institute, Radboud University, Nijmegen. E-mail: j.twose@pwo.ru.nl

Abstract

The link between psychopathy, empathy and morality is currently not well understood, therefore, to elucidate this relationship, a two-study approach was taken. In study 1, participants with psychopathic traits ($N_{\text{Psychopathic offender}} = 24$), non-psychopathic offenders ($N_{\text{Non-psychopathic offender}} = 31$), and a control group ($N_{\text{control}} = 44$) completed the Basic Empathy Scale in Adults (BES-A), and a set of thirty moral dilemmas (Greene et al., 2004). Study 1 tested a moderation model between psychopathy, the scores from the BES-A and the responses to the moral dilemmas and concluded that regardless of the BES-A scores there was no difference in how participants with psychopathic traits respond to moral dilemmas in comparison with a forensic control, and healthy control group. In study 2, the Galvanic Skin Responses (GSR) of a healthy control group ($N = 16$) were recorded when responding to morally pertinent images. The GSR time series were quantified using Recurrence Quantification Analysis (RQA), which is used to identify patterns in lagged time series. Based on the RQAs distinct physiological patterns between individuals when responding to morally pertinent images was uncovered. These two studies add insight to the discussion in psychopathy literature as to possible moderators of how people with psychopathic traits make moral decisions and informs further individualistic analyses into the physiology associated with morality.

Introduction

The idea that empathy is an integral part of moral decision making is the dominant perspective in the current behavioural, neuroscientific, and neuropsychological literature, which argues that intact emotional processes are essential to our moral psychology (Cima, Tonnaer, & Hauser, 2010; Koenigs, Kruepke, Zeier, & Newman, 2011). However, Cima and colleagues (2010) found that psychopaths, who are considered to lack empathy (part of the Psychopathy Checklist-Revised criteria; Blair, 2007), approached moral dilemmas in the same way as controls. Cima and colleagues (2010) concluded that psychopaths understood the moral dilemmas, but they did not care about the consequences. Additionally, research by Lockwood, Bird, Bridge, and Viding (2013) indicated a statistically significant negative correlation between psychopathic traits and performance on an affective empathy response task (r = -0.258, p = 0.007) and a non-significant correlation between psychopathic traits and performance in a cognitive empathy task (r = -0.046, p = 0.634). Moreover, research (experiment 2) by Gleichgerrcht and Young (2013) found a significantly lower affective empathy score in responders who thought personal moral dilemmas as permissible, in comparison to those who didn't ($t_{361} = -4.84$, $p < .001$, $d = 0.51$).

In both the morality and empathy literature there is not a consensus on the definitions used and as such there are multiple definitions in circulation. This presents a problem for researchers because it makes cross study comparisons very difficult in these fields (Batson, 2009). The multiple definitions of empathy are delineated in Batson (2009)'s comprehensive review, however the definition of empathy used in this research encapsulates the most prominent of these concepts and is similar to that outlined by Decety and Cowell (2015). They break empathy down into three components: first, affective empathy, which reflects the natural capacity to become affectively aroused by others' emotions. Second, empathic concern, which corresponds to the motivation of caring for another's welfare. Last, cognitive empathy, which is the ability to consciously put oneself into the mind of another individual and imagine what that person is thinking or feeling.

The empathy questionnaire, the Basic Empathy Scale in Adults (BES-A) which was used in this research has subscales which measures affective empathy and cognitive empathy. The creators of

this scale address this definition and their rationale for not including Empathic concern in their two-factor version of their scale by evaluating the model fit for both a two-factor model (i.e., affective and cognitive empathy), and a three-factor model of empathy (i.e., emotional contagion, emotional disconnection, and cognitive empathy). They found that the questions loaded significantly well on both versions of the model (Carré, Stefaniak, D'ambrosio, Bensalah, & Besche-Richard, 2013). Considering that both versions of the model loaded significantly well, it was decided to use the two-factor structure for the sake of parsimony, and to better compare the results of previous studies to the current study (e.g. Heynen, Van der Helm, Stams, & Korebrits, 2016; Salas-Wright, Olate, & Vaughn, 2013; Litten, Roberts, Ladyshewsky, Castell, & Kane, 2018; Pino et al., 2016).

In philosophy there are many schools of thought on morality e.g. deontology, utilitarianism, virtue theory. However, Greene and colleagues (2004) have distilled moral reasoning into the two different facets of their dual-process theory; socio-emotional moral reasoning and cognitive moral reasoning. They argue that socio-emotional moral reasoning applies to more personal moral dilemmas – moral violations that are 1) directly enacted by the person, which cause 2) physical bodily harm to 3) a tangible person/ people. Whereas, impersonal moral dilemmas require cognitive moral reasoning. Here, 'cognitive' is the set of thought processes that contrast with emotional processes. Impersonal moral violations are those which are missing one or more of the three components seen in personal moral dilemmas, and non-moral dilemmas are those with no component of moral violation at all (see appendix 1). The theory by Greene and colleagues (2004), has been criticised in the morality literature (e.g. Bauman, McGraw, Bartels, & Warren, 2014) because the initial set of moral dilemmas that were used to support this theory (Greene et al., 2001) were not statistically validated. Bauman, McGraw, Bartels and Warren (2014) call into question the whole dual-process theory because they argue that the set of dilemmas are not realistic to moral decision encountered in everyday life. Despite these criticisms it is argued that this theory is still applicable to morality because of the reasoning that Greene and colleagues (2001) use as their starting point. They found a clear difference in the answers given by individuals to the trolley dilemma and the footbridge dilemma (most people in their study would sacrifice the person in the trolley dilemma, but not the footbridge dilemma; c.f. Greene et al., 2001). In both these dilemmas a life is sacrificed to save five lives, however, with the footbridge

dilemma the individual must physically push another person to their death. Whereas, in the trolley dilemma, the person that is sacrificed just happens to be on the other track. Greene and colleagues (2001) argue that the different answers to these dilemmas are due to how personal the footbridge dilemma is. Furthermore, the dilemmas developed by Greene and colleagues (2004) are still often used for research in moral decision making, hence it was decided to use the same dilemmas in the current research to allow cross-comparison between studies.

The psychopathy theory underpinning the current study, and most research into psychopathy, is that of Cleckley (1976) and Hare (1980; 1991; 2003; 2017). Based on his observations of those deemed to have psychopathic personalities Cleckley (1976) identified 16 different characteristic points that were inherent in people with psychopathic traits (appendix 2). Hare (1991; 2003) created a measurement scale called the Psychopathy Checklist(-Revised) to systematically measure and quantify these characteristics for the better diagnosis of psychopathy in a forensic setting. The PCL-R continues to be used for this purpose, however it has gone through some changes since it's conception. The PCL-R currently consists of 20 characteristics marked on a 3-point Likert scale (0, 1, 2) according to the degree to which the individual fits the characteristic. The PCL-R is only valid when administered by a trained psychologist.

Skin conductance can be used as a psychophysiological measure during moral decision making (Krosch, Figner, & Weber, 2012; Moretto, Làdavas, Mattioli, & Di Pellegrino, 2010; Figner & Murphy, 2011; Hristova, Kadreva, & Grinberg, 2014; Blair, 1999). For example, Krosch, Figner and Weber (2012) conducted research on the link between moral decision making and skin conductance response. Based on moral decision-making vignettes (let wartime refugees stay - humanitarian response, turn refugees away - military orders), they found strong positive correlations ($r = .55$, $p < .01$) between the skin conductance response and a measure of difficulty associated with the decision for those who chose the humanitarian response. Furthermore, there was also a positive correlation between how strongly participants used Military Role, Self-Interest, and Consequence modes (latent variables based on Structural Equation Modelling conducted a priori; $r = .40$, $p = .02$). Additionally, Moretto, Làdavas, Mattioli, and Di Pellegrino (2010) conducted research on the physiological differences of how people with ventromedial prefrontal cortex make moral decisions, using Greene

and colleagues (2004) moral dilemmas, in comparison with a control group. Moretto, Làdavas, Mattioli, and Di Pellegrino (2010) submitted the mean skin conductance response following the moral dilemmas to a mixed-design ANOVA, with group (ventromedial prefrontal cortex lesion, non-frontal cortex lesion, control) as a between-subject factor, and dilemma (personal, impersonal, nonmoral) and choice (utilitarian, nonutilitarian) as within-subject factors. Both the main factor of choice, $F(1, 30) = 19.4$, $p < .001$, and the interaction between choice and dilemma $F(2, 60) = 4.5$, $p < .01$ were found to be significant.

At present there is no theoretical framework which links these three constructs together, in a forensic population. Blair and Blair (2009) allude to this in their literature review on the subject but, do not present a working theoretical model. In the current research on psychopathy and morality it is assumed that there is a difference between the moral reasoning of controls vs people with psychopathic traits because of a lack of empathy (Koenigs et al., 2011) or normal moral emotional appraisal (Blair, 1995; Blair et al., 1995, Aharoni, Sinnott-Armstrong, & Kiehl, 2012), however, they do not directly measure if the underlying cause is empathy. Aharoni, Sinnott-Armstrong, and Kiehl (2012) address the possible underlying causes between psychopathy and moral reasoning, tentatively reporting that reduced moral categorization accuracy was significantly predicted by affective and antisocial traits. They broke down the PCL-R results into 4 different facets (interpersonal, affective, lifestyle, and antisocial facets) and found that the affective and antisocial facets uniquely predicted reduced moral accuracy, and the lifestyle facet uniquely predicted increased moral accuracy, controlling for the other facets (based on hierarchical regression; $R^2 = .14$, $p < .01$). However, their research consisted of a sample of $N = 109$, incarcerated individuals, with only 6 being diagnosed as people with psychopathic traits.

The assumption that any difference in moral reasoning is caused by empathy appears to stem from the definition used when measuring psychopathy with the PCL-R. The PCL-R breaks the psychopathy diagnosis into two factors; a selfish, callous, and remorseless use of others (Factor 1) and a chronically unstable, antisocial, and socially deviant lifestyle (Factor 2). Since, Factor 1 is so related to empathy it is the opinion of the author that previous research has assumed that any significant

differences in moral reasoning are driven by a difference in empathy. In this research, this assumption was addressed.

Current study

The current research was split into two separate studies and will be dealt with as such in the methodology and results sections. Study 1 was concerned with modelling how empathy moderates the responses of people with psychopathic traits to moral dilemmas. Study 2 revolves around the quantification of patterns in the GSR of a control population whilst they look at morally pertinent images.

To address the link between psychopathy, empathy and moral dilemmas study 1 was conducted, and to address the physiological response to morally pertinent images study 2 was conducted. A moderation model (figure 1) was tested in study 1 to explain the difference in findings between Cima and colleagues (2010); that there was no significant difference in moral dilemma response between people with psychopathic traits and those without psychopathic traits, and Koenigs and colleagues (2011), who did find a significant difference, and to elucidate if there is a moderating effect of empathy on how people with psychopathic traits approach moral dilemmas. This model tested to see if there was an interaction between levels of cognitive and affective empathy, and the outcome of the moral dilemmas in psychopathic and non-psychopathic offenders, and controls. It was predicted by the author that if there was a difference in the permissibility of the moral dilemmas between the groups then this permissibility would be moderated by the levels of affective and cognitive empathy.

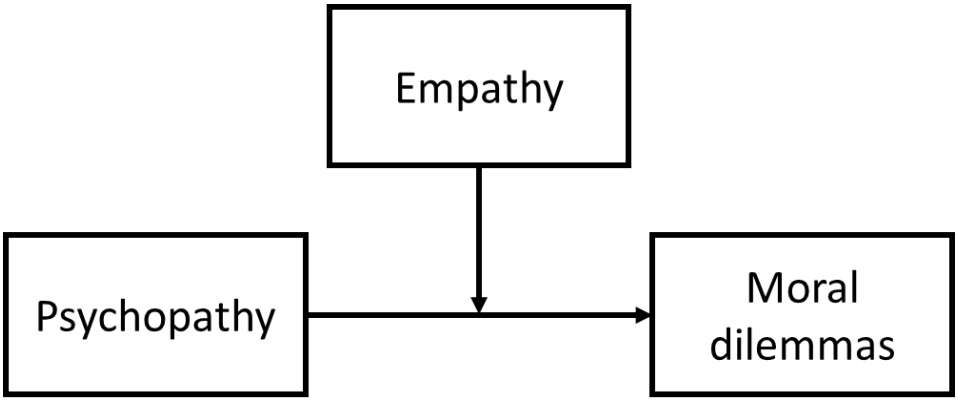

Figure 1. Moderation model of empathy on psychopathy and morality.

In addition to the moderation model conducted in study 1, Galvanic Skin Response (GSR) was quantified in a non-clinical population in study 2, whilst morally pertinent images were presented (Hess & Fischer, 2014). GSR (Levenson & Ruef, 1992) was used as a biomarker to measure physiological arousal as a surrogate for reaction to moral stimuli. Previous studies (Hein, Lamm, Brodbeck, & Singer, 2011; Nešić, Uljarević, Čičević, & Nešić, 2012; Van Rijn, Barendse, van Goozen, & Swaab, 2014; Prguda & Neumann, 2014; Levenson & Ruef, 1992; Forgiarini, Gallucci, & Maravita, 2011) have found significant correlations between GSR and empathy. Hein and colleagues (2011), reported that when participants observed painful stimulation of others (M =0.23, SE =0.05) as compared to non-painful stimulation (M =0.1, SE =0.02) the GSR was significantly higher $t(19)$ =2.78, p =0.006, η_p^2 =0.29. Nešić and colleagues (2012) described a general linear model that evaluated empathy as a predictor of GSR to evocative movie scenes and found a significant interaction $F(28) = 2.202$, $p = 0.001$. Prguda and Neumann (2014) tested whether empathy, as measured by the balanced emotional empathy scale (BEES) was correlated with GSR and found a significant correlation between the BEES scores and overall phasic GSR ($r = 0.24$, $p < 0.05$). The aim of this study was to uncover any individual physiological differences in how individuals approach morality, which may have been overlooked in the moderation model. To account for whether the empathy of the participants was reflected in the skin conductance, the scores from the subscales of the BES-A were correlated with the dynamics measures. The moderation model was based on hierarchical linear regressions which are a group level analysis technique therefore underlying individual differences are not directly accounted for. In the literature, it is argued that the quantification of time series data reveals a lot about the underlying properties of a process (e.g. de Graag et al., 2012; Schiepek et al, 2016). For example, Schiepek and colleagues (2016) currently quantify non-linear time series to model change in different psychopathologies. They argue that high levels of change (or chaos), are linked with a reorganization of the system (the patient) and thus are clinically relevant, due to this reorganization being the best time to administer therapy. Their research has been very clinically promising and hence the quantification of non-linear time series was used in study 2. Reorganisation within the system was thought to play a role in study 2 because prior to the decisions about the morally pertinent images, it was expected that there is an underlying decision-making mechanism which may

be reflected in the reorganisation of the skin conductance response. If this was the case the quantification of patterns in the skin conductance response time series would detect this reorganisation.

Methodology

Study 1

Sample. Participants (all male adults) were asked to give informed consent in accordance with the policies of the Ethical Commission of the Faculty of Social Sciences, Radboud University, The Netherlands. Healthy controls were recruited from the south of the Netherlands. The psychopathic and non-psychopathic offenders were recruited from the Forensic Psychiatric Centre de Rooyse Wissel (FPCdRW) in Venray, the Netherlands. To correctly compare the control group to the experimental group, age (M_{FPCdRW} = 37.88 years, SD = 10.31; M_{control} = 34.07, SD = 11.55) and gender were matched.

A power analysis using the Gpower computer program (Faul & Erdfelder, 1998) indicated that a total sample size of 159 participants would be needed to detect small effect sizes (f^2 =.05; because interactions often only account for a small percentage of variance) with 80% power using a Linear multiple regression: Fixed model, single regression coefficient testing the effect of 2 main predictors (psychopathy and empathy) and their 3 interactions (psychopathy-cognitive empathy, psychopathy-affective empathy, psychopathy-cognitive and affective empathy) with an alpha of .05. However, due to the sample population (incarcerated persons with psychopathy) it was not feasible to recruit such a high number. Furthermore, based on the high power found in previous research (Cima et al., 2010; Koenigs et al., 2011), it would not be unreasonable to calculate power for the interaction based on a medium effect size (f^2 =.15) which would give a total sample size of 55. The final sample size collected was 101 participants (N_{FPCdRW} = 57, M = 37.88 years, SD = 10.31; N_{control} = 44, M = 34.07, SD = 11.55), which, based on the power analyses, was deemed a sufficient population size for the moderation model. The groups that were entered into the moderation model were psychopathic offenders ($N_{\text{Psychopathic offender}}$ = 24), non-psychopathic offenders ($N_{\text{Non-psychopathic offender}}$ = 31), and the control group (N_{control} = 44).

The eligibility requirements of the participants were an IQ greater than 70 and no history of psychosis or bipolar disorder. Males only were recruited in this study because the FPCdRW clinic is a single sex institution which only houses incarcerated males.

Measures. FPCdRW participants had the Basic Empathy Scale for Adults (BES-A; Carré, Stefaniak, D'ambrosio, Bensalah, & Besche-Richard, 2013), and the set of moral dilemmas (see appendix 1) developed by Greene and colleagues (2004) administered at the FPCdRW. The control participants followed the same procedure at the Behavioural Science Institute, Nijmegen. For the identification of psychopathy in the FPCdRW the PCL-R was used, a score of 26 or higher was defined as a psychopathic offender. The PCL-R is the most well-validated scale that is used to evaluate psychopathy (Hare, 1991; 2003; 2017). Unfortunately, because the PCL-R score is a clinical measure, psychopathy scores were not available for the control group.

Empathy. To ascertain the empathy levels of the participants the BES-A (Carré et al., 2013) was used. This questionnaire was chosen instead of other more commonly used questionnaires because the BES-A reports to measure the multidimensional nature of empathy; cognitive and affective empathy. The creators of the BES-A (Carré et al., 2013) conducted confirmatory factor analysis (N = 370) and found high factor loadings of the questions on the two-factor structure of the latent variables cognitive (i.e. I can understand my friend's happiness when she/he does well at something) and affective empathy (i.e. after being with a friend who is sad about something, I usually feel sad), in comparison with other models applied to the same data (single factor structure; empathy, and the three-factor structure; emotional contagion, cognitive, and affective empathy). There is high internal validity of the two-factor structure of the BES-A based on the goodness of fit indices, χ^2 (169) = 510.65, $p < .001$, RMSEA = .074 (90% CI [.067, .81]), GFI = .95 reported by Carré and colleagues (2013).

Moral dilemmas. Study 1 used a set of 30 moral dilemmas taken from the moral dilemmas developed by Greene and colleagues (2004). The moral dilemmas were grouped as "Personal" (N =

14), "Impersonal" (N = 10) and "Non-moral" (N = 6; see appendix 1 for examples). Greene and Haidt (2002) and Greene and colleagues (2001) provided evidence for the use of the chosen moral dilemmas and the distinction between the types of moral dilemmas in the form of fMRI studies which outlined higher activation of the medial frontal gyrus, posterior cingulate gyrus, and bilateral superior temporal sulcus in participants experiencing personal moral dilemmas, and higher activation of areas associated with working memory: dorsolateral prefrontal and parietal areas in those experiencing impersonal moral dilemmas. This provides support for the personal and impersonal distinction between the moral dilemmas that were used in the moderation model in study 1. Furthermore, Moore, Lee, Clark, and Conway (2011) validated the use of the moral dilemmas using a between item ANOVA where the dependent variable was the across-subjects proportion of moral approval for each dilemma and the cases were individual dilemmas. They found that when killing one to save many, there was significantly less moral approval for personal compared to impersonal killing, $F(1, 46) = 5.45$, $p = .02$, $\eta_p^2 = .11$.

Study 2

Sample. Participants (all male adults; N = 16) were asked to give informed consent in accordance with the policies of the Ethical Commission of the Faculty of Social Sciences, Radboud University, The Netherlands. Healthy controls were recruited from Nijmegen, the Netherlands.

Materials. In study 2 the materials of interest are the morally pertinent images (see supplemental materials), the BES-A (Carré et al., 2013), and the Galvanic Skin Response. The morally pertinent images were chosen by the author based on the perceived moral relevance. Control images were included to allow for comparison, however they were not considered due to timing issues with equipment. The images were presented at 2000 millisecond intervals as part of a computer task (figure 2) followed by a question as to whether the image was morally pertinent (question 1), how morally impermissible the act in the image was (question 2), and a question as to what extent the participant was affected by the image (question 3) via a python script ran in Psychopy 2. The participants were required to answer the aforementioned questions by indicating yes or no (question 1) and 1 – 7 (question 2 and 3) on the keyboard of the laptop. The Python script ran on a Dell Inspiron 15 5000

series with the Windows 10 operating system. The images in the script were counterbalanced, with two different orders of images, to discount any order effects of the images affecting the results.

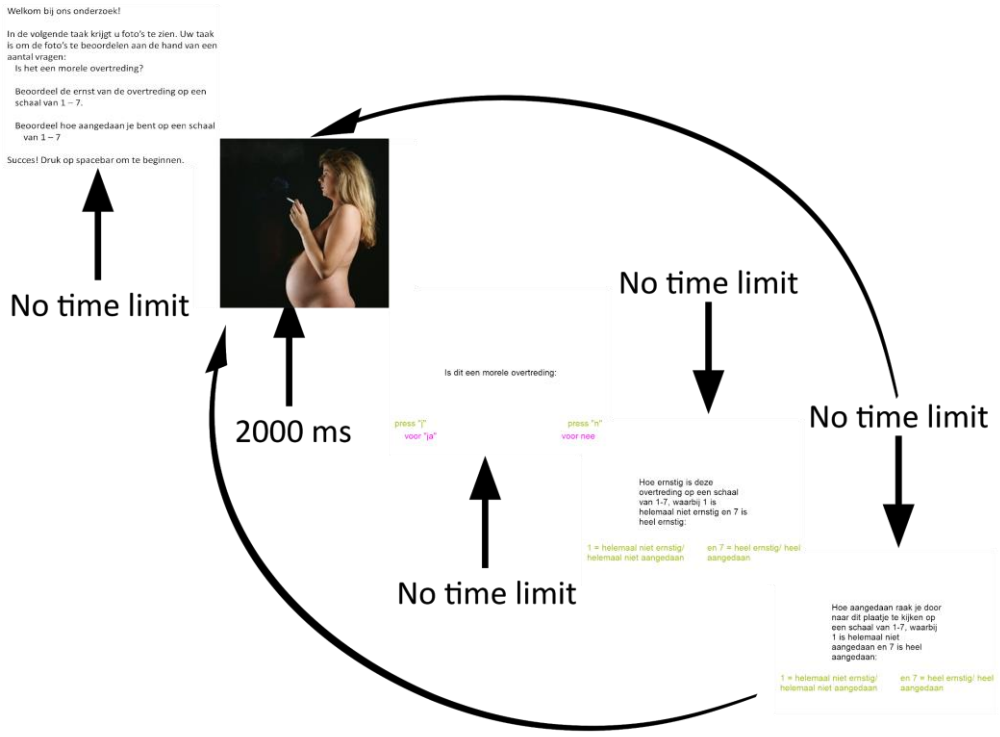

Figure 2. Study 2 computer task setup. Instructions were presented at the beginning of the task then the last 4 screens were iterated 30 times.

For the duration of this computer task the GSR of the participants was recorded using a GSR100C module (BIOPAC MP150) employing a .5 V constant voltage technique, as recommended by Fowles and colleagues (1981). The GSR100C module was connected to a Dell XPS L501X computer running the Windows 7 operating system. The GSR was recorded from the distal phalanges of the second and third fingers of the nondominant hand. The two computers were connected with a buttonbox (Technical Support Group - Radboud University) which recorded a binary button press at the beginning of the computer task to synchronize the images with the GSR time series.

Analysis

Study 1

The statistical approaches used for the model, in study 1, between psychopathy, type of empathy, and the two types of moral dilemmas was a moderation model conducted using the lavaan R package (Rosseel, 2012; Rstudio team, 2015). The lavaan R package was used to specify the moderation model because of the multiple imputation methods that are available to account for missingness in the data, and for the robust maximum likelihood estimator which adjusts for skewness and kurtosis in the data. Furthermore, the lavaan R package works in conjunction with the semPlot R package (Epskamp, 2015) which was used to display the different weights and significances for paths in the model.

Moderation model analyses. Groups (based on FPCdRW and controls) and score on the BES-A Cognitive Empathy and Affective Empathy subscales were used to predict outcomes on the moral dilemmas; Impersonal and Personal dilemmas. Data were checked for outliers and the regression assumptions and no violations were found. Data were centered and scaled prior to analysis. The lavaan R package was used to conduct the multiple regressions and to analyse the interactions that make up the moderation model under investigation. The overall moderation model (figure 3) fitted the data significantly $F(13, 101) = 52.488$, $p < 0.001$, $R^2_{personal} = 0.035$, $R^2_{impersonal} = 0.032$, however the main effects and interactions were not significant $p > 0.163$ (for a breakdown of all the estimates and p-values see Appendix 3). What can be seen from the model (figure 3) is that the only significant relations in the model were the covariance between the Personal and Impersonal moral dilemmas (seen by the bold green line).

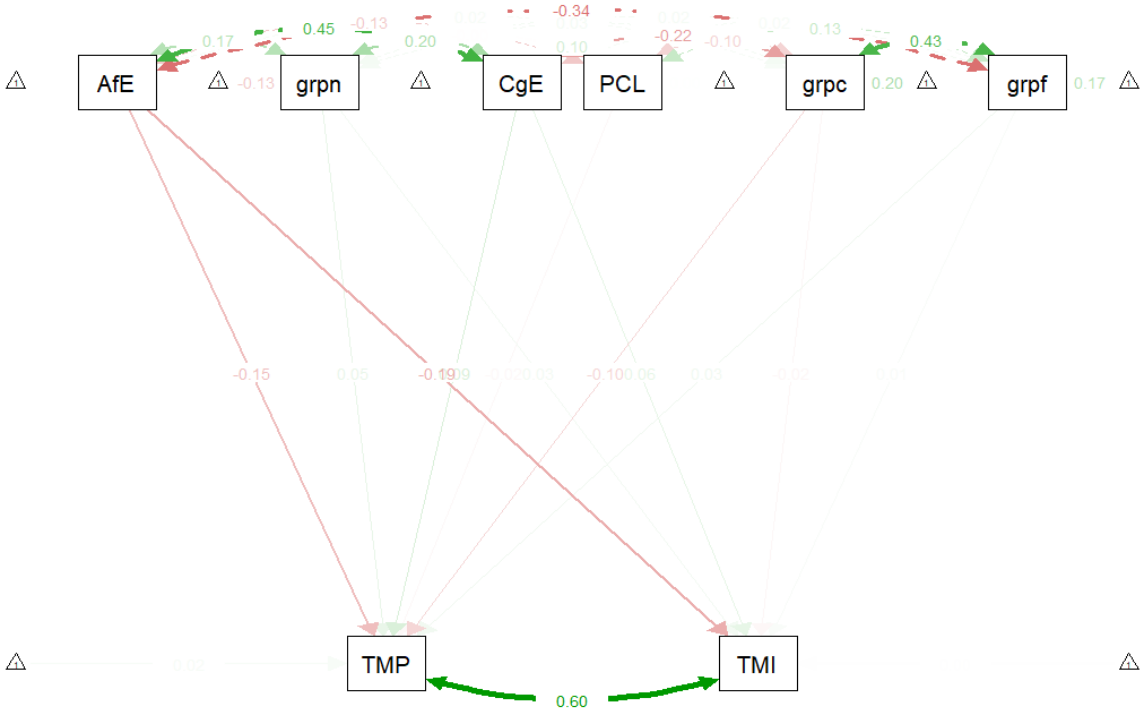

Figure 3. Overall moderation model conducted using lavaan in R. AfE = Affective Empathy, grpn = Group, CgE = Cognitive Empathy, PCL = PCL-R score, grpc = the interaction between group and cognitive empathy score, grpf = the interaction between group and affective empathy score, TMP = Personal Moral Dilemmas, TMI = Impersonal Moral Dilemmas. Bold arrows amount to significance, translucent arrows are not significant. Dashed lines indicate fixed parameter estimates. Green indicates a positive effect, red indicates a negative effect.

Study 2

The dynamics of GSR were quantified via Recurrence Quantification Analysis (RQA; see Marwan, Romano, Thiel, & Kurths, 2007), using the methods available in the casnet R package (Hasselman, 2017). RQA is a data analysis technique taken from the dynamics toolbox which is based on three main principles of individuality: jaggedness principle - people are all different, hence it is not possible to apply one-dimensional thinking to understand something as complex as human behaviour; Context principle – this works on the IF – THEN paradigm, in that IF a person is in a certain situation

THEN they will react in a certain way; Pathways principle – there are multiple different physiological, psychological, neurological pathways to research human behaviour (Rose, 2016; Rose, Rouhani, & Fischer, 2013). These principles of individuality contrast from the regular linear model that is mainly used in this type of research because they infer the method of analysing before aggregating, rather than aggregating then analysing. Therefore, patterns are identified in individuals prior to attempting to find groups differences in the patterns. These principles of individuality give dynamics an edge when it comes to this type of research because often complex differences between the behaviour of people are often missed in the regular linear model because it is such a broad and sweeping analysis technique. Furthermore, the analyses associated with dynamics are all non-linear, thus no assumptions are made about the data prior to analysing it. It is the aim of the RQA in study 2 to undercover patterns within the GSR data that may have been missed in previous research due to the aggregation across groups which is necessary in the general linear model.

Recurrence Quantification Analysis. In study 2 RQA was conducted on a healthy/ non-clinical population (N = 16). The data used in this analysis was the GSR of this group whilst the aforementioned computer task was completed. The measures of interest that are calculated from RQA are the recurrence rate, determinism, laminarity, entropy (vertical and diagonal), and the times between recurrent patterns(T1 (diagonal) and T2(vertical)) (see appendix 4 for formulas). The recurrence rate is defined as the percentage of data points that share a common area in phase space, dependent on a defined radius (the mean Euclidean distance separating data points in reconstructed phase space). The determinism is defined as the percentage of recurrent points that constitute line segments - recurrent patterns- parallel to the diagonal identity line in a recurrence plot. The laminarity is defined as the percentage of recurrent points at a 45-degree angle to the diagonal identity line (Wijnants et al., 2009). Practically speaking the determinism and laminarity values is a quantification of how deterministic (i.e. predictable) the time series is. The entropy is based on the Shannon entropy of the probability distribution of the diagonal and vertical line lengths (Marwan, Romano, Thiel, & Kurths, 2007). The entropy measures can be understood as a measure of how stable the recurring patterns of (i.e. random) the time series are - the higher the entropy the less stable the recurring patterns are. In order to

compare across individuals it was decided that an embedding lag of 33.5 (the median of all the embedding lags chosen based on the mutual information method; Cao, 1997) and an embedding dimension of 9 (maximum embedding dimension of the time series chosen based on the false nearest neighbours algorithm; Kennel, Brown, & Abarbanel, 1992) would be used to define the appropriate radii (size of which there was a recurrence rate of 0.01) of the different time series (Van Orden, 2005). The embedding lag, embedding dimension, and the radii were used when calculating the RQA measures. Different line lengths that would be considered recurrent points were ran in order to find where the individuals diverged most, in order to best disentangle the different patterns. The line length of 10 was chosen because it was at this line length that there was the biggest divergence (see supplementary material for graphs to support this choice).

As this research is very exploratory, with there being a dearth of dynamical data analysis techniques in morality, empathy and psychopathy research, it was decided that the analysis would be conducted on the overall GSR time series, rather than sub setting the time series based on image. The implications of this for the results are that it is not possible to directly link the image to the specific section of the GSR time series, which would provide more insight into what image elicits what physiological pattern.

Whilst acknowledging the drawbacks associated with study 2, two groups clearly emerge from the data based on the recurrence matrix plots. A group with very fast changing GSR related to all the pictures that are shown, and a group with much slower GSRs (see figures 4 and 5 and supplemental material). This can be seen based on the square structures that have formed in figure 4, any quick change from red to blue indicates a fast change in GSR. A subset of 10 individuals from this control groups emerge with similar square structures seen in their recurrence plots. Contrasted to the square structures seen in figure 4, figure 5 consists of three different individual recurrence plots which show very slow shifts from blue to white to red, which is indicative of very slow changes in GSR. Each point that is in the recurrence matrix plots is linked with the recurrence rate. Hence, the longer the vertical lines that form from these points, the higher the laminarity of the system. Similarly, the longer the diagonal lines that form from these points, the higher the determinism of the system. Therefore, the

recurrence matrix plots with the quick changes (figure 4) have a higher laminarity than those with the slow changes (figure 5).

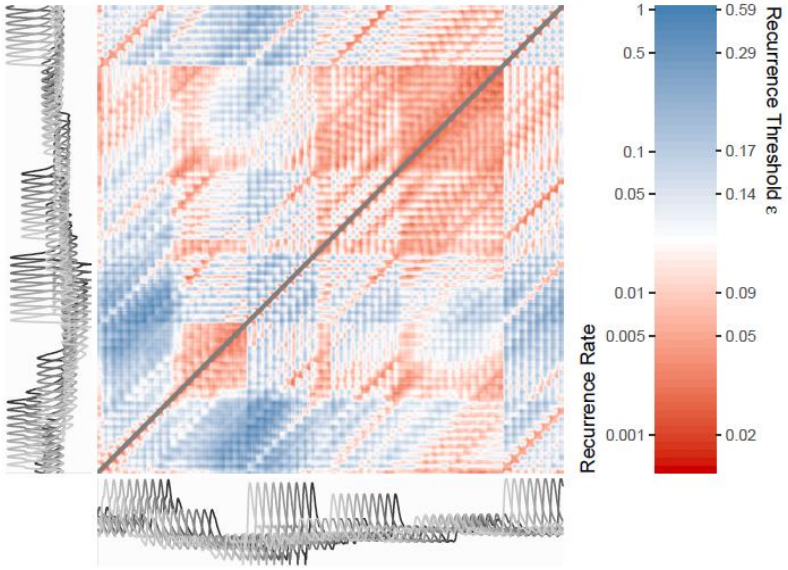

Figure 4. Unthresholded recurrence matrix plot of the fast changing GSR (group of 10 similar to this plot). The higher the recurrence rate, the bluer the point in the matrix. The embedded time series can be seen mapped to the bottom and side of the recurrence matrix plot.

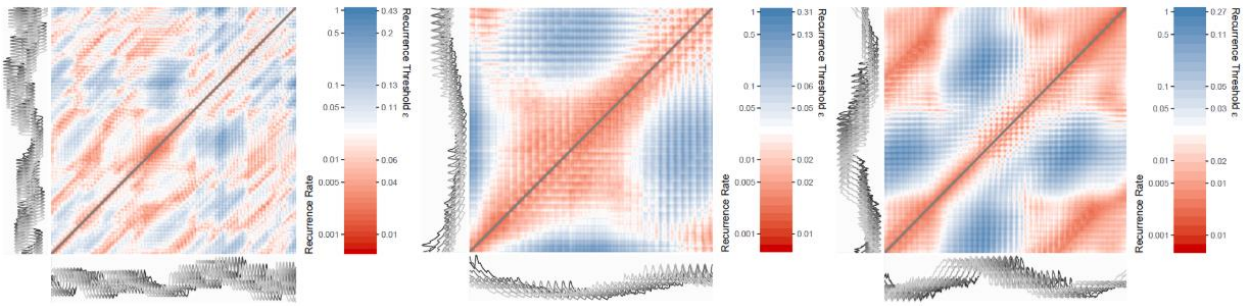

Figure 5. Three unthresholded recurrence matrix plots of the slow changing GSR group (group of 6 similar to this plot). The higher the recurrence rate, the bluer the point in the matrix. The embedded time series can be seen mapped to the bottom and side of the recurrence matrix plot.

From the large differences that can be seen between figures 4 and 5, it is discernible that two groups of people that emerge from study 2. A group which have very quick physiological reactions to the morally pertinent images (figure 4), and a group with a much slower physiological response to the images (figure 5).

To try to better understand the group difference seen in figures 4 and 5, and to relate the RQA measures (vertical entropy, diagonal entropy, determinism, laminarity, T1 and T2) to the empathy subscales taken from the BES-A, a difference score was calculated. These RQA measures and difference score are plotted against each other in figure 6. Based on the interquartile range, any score on the empathy subscales that was in the lower 25% quartile was grouped as low, and any score that was above the 75% quartile was grouped as high. This grouping method was chosen with the main aim of finding the biggest difference in groups. This was necessary because the population were healthy controls and so all fell within a comparatively small range on the subscales. The main point of interest from these graphs is that individuals with a higher affective empathy score (N = 3; blue shaded area) have less changeable RQA measures and thus a more stable physiological response to the morally pertinent images at the group level in comparison to those who have a higher cognitive empathy (N = 13; red shaded area). This can be seen from the relatively flat LOESS line. The group with a higher cognitive empathy score have a much higher chance of variability and thus a less stable physiological response at the group level, although there does appear to be a pattern, but the pattern itself is non-linear, possibly cubic. This pattern occurs across all the different RQA measures, so this finding seems to be stable.

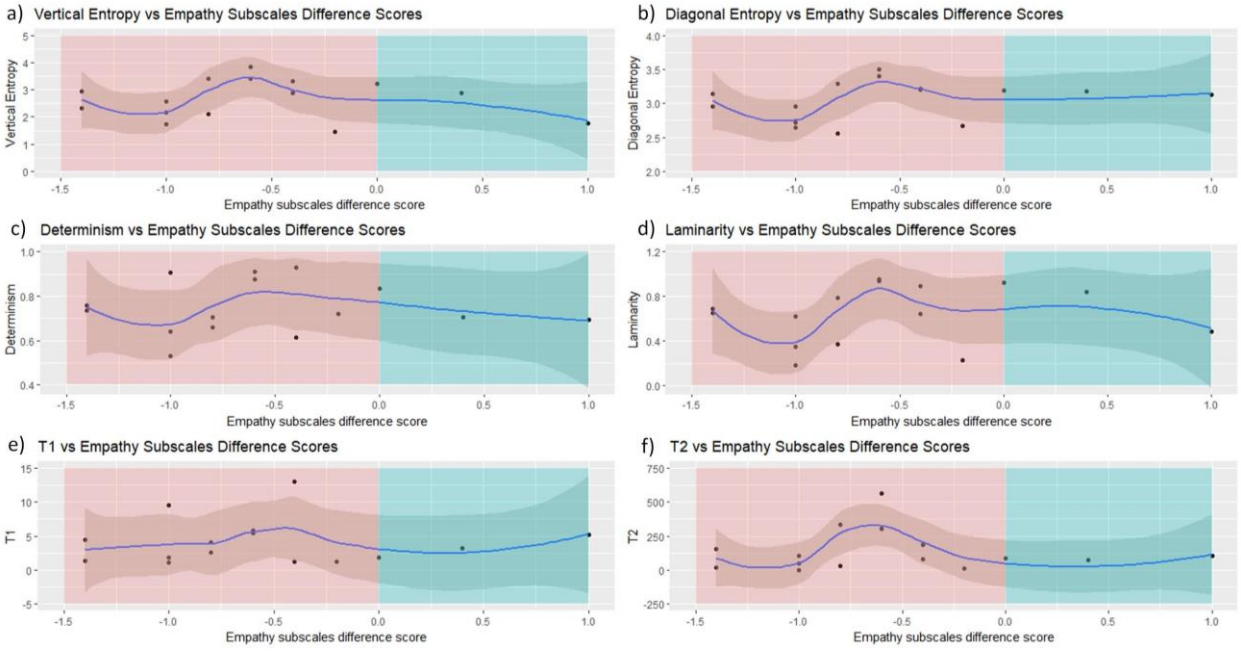

Figure 6. A comparison of how a) vertical entropy, b) diagonal entropy, c) determinism, d) laminarity, e) T1, and f) T2 change in relation to empathy subscale difference score. The blue section indicates a higher affective empathy BES-A subscale score and the red section indicates a higher cognitive empathy BES-A subscale score.

To further investigate the fast and slow physiological groups which appeared in figures 4 and 5 in relation to the empathy subscales, figure 7 displays the two groups on the same scale as in figure 6. Based on where these groups lie on the empathy subscales difference scores it does not seem that the physiological group difference is a result of the difference in the empathy score.

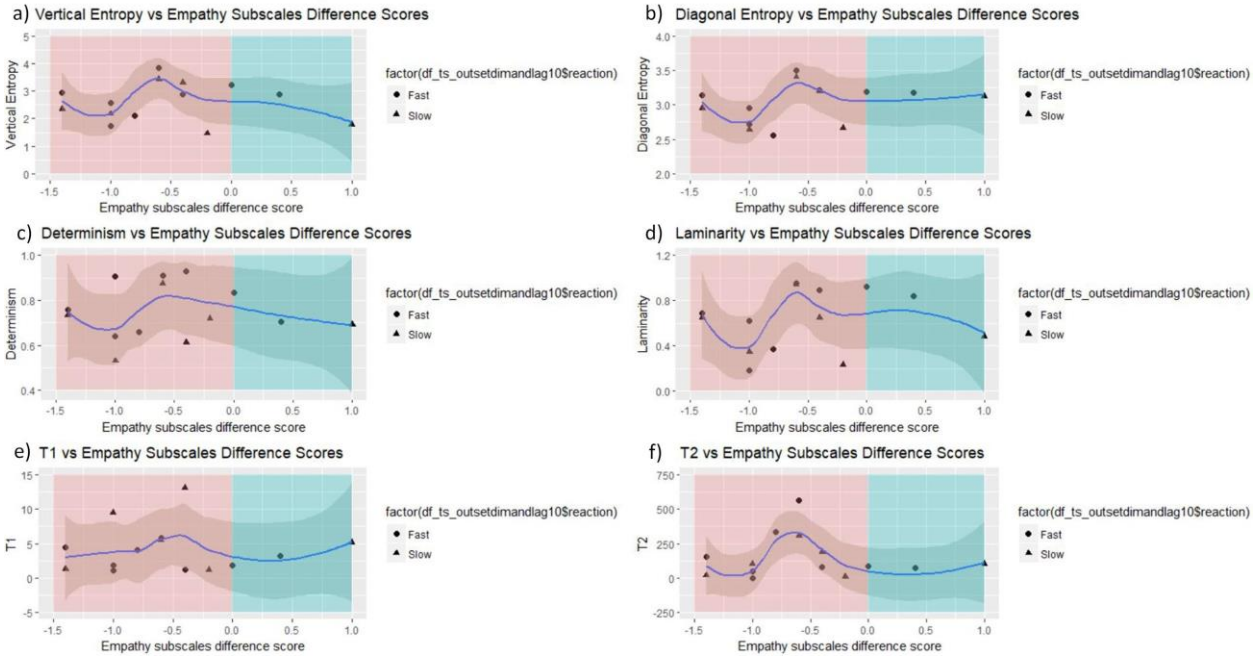

Figure 7. A comparison of how a) vertical entropy, b) diagonal entropy, c) determinism, d) laminarity, e) T1, and f) T2 change in relation to empathy subscale difference score. The blue section indicates a higher affective empathy BES-A subscale score and the red section indicates a higher cognitive empathy BES-A subscale score. Participants are grouped based on the speed of their physiological reaction to the morally pertinent images.

Discussion

Based on the results of study 1, the moderation model appears to show support for the Cima and colleagues (2010) conclusion, in that there were no significant main effects or interactions of

psychopathy, nor of the empathy subscales, on the responses to the moral dilemmas. Therefore, when it comes to responding to both the Personal and Impersonal moral dilemmas the results suggest that people with psychopathic traits know what the correct response is, no matter what their level of cognitive and affective empathy are. However, from the results of study 2, there does appear to be an underlying physiological difference between people within a control population when they are asked to rate moral and non-moral images. This can be seen from the distinct patterns garnered from the RQA analyses.

Study 1 failed to replicate the findings of Koenigs and colleagues (2011) who found a significant difference between people with psychopathic traits and a control population in the way that the answered Greene and colleagues' (2004) moral dilemmas. A possible reason for this is that the statistical model of multiple t-tests that Koenigs and colleagues (2011) used was much simpler than the moderation model in study 1. The advantages of using a moderation model over multiple t-tests is that it is possible to control for the different (co)variances between all the variables, and to measure interactions between variables which may account for variance seen between groups. Furthermore, the sample size used by Koenigs and colleagues (2011) was 48 ($N_{\text{Low anxious psychopathic offenders}} = 12$; $N_{\text{High anxious psychopathic offenders}} = 12$; $N_{\text{control}} = 24$), in comparison with the sample size of 101 ($N_{\text{Psychopathic offender}} = 24$; $N_{\text{Non-psychopathic offender}} = 31$; $N_{\text{control}} = 44$) used in study 1.

The fact that the only significant part of the moderation model in study 1 was the covariance between the Personal and Impersonal moral dilemmas perhaps warrants further attention in a follow up study. Considering there was so much overlap in the variance of the two scores, this could be possible evidence for the argument that the Greene and colleagues (2001) distinction between Personal and Impersonal moral dilemmas may be suboptimal, or inaccurate. Therefore, follow up research should be conducted to provide further support for this distinction in Greene and colleagues (2001) dual process theory.

The findings of study 2 are consistent with those found by Moretto, Làdavas, Mattioli, and Di Pellegrino (2010). They found significant within group differences between the types of moral dilemma (personal, impersonal, nonmoral) and the choice made by their participants (utilitarian, nonutilitarian). Whilst study 2 uses a different paradigm to investigate moral decision making, there is

clearly an individual difference between people when it comes to interacting with morally pertinent stimuli, which can be seen in both study 2 and the study of Moretto, Làdavas, Mattioli, and Di Pellegrino (2010). From the results of study 2 it can be concluded that the link between skin conductance and moral decision making is complex and as such should be investigated from a dynamics of complex systems perspective in future research. Whether, this individual difference can be found when participants answer Green and colleagues' (2004) moral dilemmas remains to be seen, however based on study 2, it is promising. To investigate if there is indeed a difference in the physiological response of people to personal as compared to impersonal moral dilemmas at the individual level, research with a similar approach to study 2 should be conducted, but instead of the morally pertinent images, the dilemmas should be used.

The implications for clinical practice based on these two studies are that a more individualistic approach to disorders of moral judgement should be used. Whilst study 2 did not consist of a clinical group, there were still considerable differences in the skin conductance response of people on an individual level. If clinical practice were to move away from a group level approach of validating therapies, conceivably treatment would be better tailored to patients.

Limitations and future directions

The results found in study 2 are very exploratory and further research should be done, before it is certain that an intrinsic physiological difference has been quantified in the population in relation to morally pertinent images. It could be argued that in study 2, it is not clear whether arousal to morally pertinent images is being measured or rather that it is arousal to stimuli in general. To control for this, future research should compare the GSR of people responding to morally pertinent images and the GSR of people responding to other physiologically arousing images (e.g. painful images).

Study 2 should be viewed as a starting point for future studies assessing the moral cognition of different groups of people and informs more individualistic analyses into the physiology associated with morality. Based on the results of both study 1 and study 2, it would suggest that both psychopathic and non-psychopathic people understand what is morally right and morally wrong (study 1), however between people there is a distinct physiological group difference between people (study

2). This group difference may not be due to a difference in empathy levels, however, future research should provide evidence for what may be driving this physiological difference. Additionally, further research should build on study 2 by comparing the GSR of different groups of people (e.g. people with psychopathic traits, people on the autistic spectrum) when asked to rate morally pertinent images, to better elucidate the link between psychopathy and morality.

Conclusions

In conclusion, study 1 demonstrated that regardless of level of cognitive or affective empathy there is no difference in how people with psychopathy respond to both personal and impersonal moral dilemmas in comparison with the FPCdRW and healthy control population.

Study 2 demonstrated a distinct physiological difference between individuals from a healthy control population when responding to morally pertinent images. Whilst this distinction was not a result of the affective and empathy subscales, there does appear to be an underlying difference in people with regards to how they physiologically react to morally pertinent images. Therefore, study 2 is a promising indicator that there are underlying individual differences, in relation to morality, that have been missed in previous studies due to their use of group level statistical methods. In conjunction, these studies greatly add to the empathy, morality and psychopathy literature, from both a group and individual level perspective.

References

Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. A. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *Journal of abnormal psychology*, *121*(2), 484.

Baron-Cohen, S., & Wheelwright, S. (2004). 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.

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *Social neuroscience. The social neuroscience of empathy* (pp. 3-15). Cambridge, MA, US: MIT Press.

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, *8*(9), 536-554.

Blair, R. J. R. (1999). Psychophysiological responsiveness to the distress of others in children with autism. *Personality and Individual Differences*, *26*(3), 477-485.

Blair, R. J. R. (2007). Empathic dysfunction in psychopathic individuals. *Empathy in mental illness*, 3-16.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, *57*, 1–29. doi:10.1016/0010-0277(95)00676-P

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1995). Is the psychopath 'morally insane'? Personality and Individual Differences, 19, 741–752. doi:10.1016/0191-8869(95)00087-M

Blair, R. J. R., & Blair, K. S. (2009). Empathy, morality, and social convention: Evidence from the study of psychopathy and other psychiatric disorders. *The social neuroscience of empathy*, 139-152.

Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, *110*(1-2), 43-50.

Carré, A., Stefaniak, N., D'ambrosio, F., Bensalah, L., & Besche-Richard, C. (2013). The Basic Empathy Scale in Adults (BES-A): Factor structure of a revised form. *Psychological assessment*, *25*(3), 679.

Cima, M., Tonnaer, F., & Hauser, M. D. (2010). Psychopaths know right from wrong but don't care. *Social cognitive and affective neuroscience*, *5*(1), 59-67.

Cleckley, H. (1976). The mask of sanity: An attempt to clarify some issues about the so-called psychopathic personality . St. Louis, MO: Mosby. *Inc.(Original work published 1941)*.

Decety, J., & Cowell, J. M. (2015). Empathy, justice, and moral behavior. *AJOB neuroscience*, *6*(3), 3-14.

de Graag, J. A., Cox, R. F., Hasselman, F., Jansen, J., & de Weerth, C. (2012). Functioning within a relationship: Mother–infant synchrony and infant sleep. *Infant Behavior and Development*, *35*(2), 252-263.

Epskamp, S. (2015). semPlot: Unified visualizations of structural equation models. *Structural Equation Modeling: a multidisciplinary journal*, *22*(3), 474-483.

Eres, R., Decety, J., Louis, W. R., & Molenberghs, P. (2015). Individual differences in local gray matter density are associated with differences in affective and cognitive empathy. *Neuroimage, 117*, 305-310.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. *A handbook of process tracing methods for decision research*, 163-184.

Forgiarini, M., Gallucci, M., & Maravita, A. (2011). Racism and the empathy for pain on our skin. *Frontiers in psychology, 2*, 108.

Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., & Venables, P. H. (1981). Committee report: Publication recommendations for electrodermal measurements. *Psychophysiology, 18*, 232–239.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work?. *Trends in cognitive sciences, 6*(12), 517-523.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*(2), 389-400.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fmri investigation of emotional engagement in moral judgment. *Science, 293*, 2105–2108.

Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PloS one*, *8*(4), e60418.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and individual differences*, *1*(2), 111-119.

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised.* Toronto, Ontario, Canada: Multi-Health Systems.

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems

Hare, R. D. (2017, June 20). *Key References Related to Psychopathy and the Hare Scales* Retrieved from http://www.hare.org/references/

Hasselman, F. (2017). casnet: A toolbox for studying Complex Adaptive Systems and NETworks URL https://github.com/FredHasselman/casnet.

Hein, G., Lamm, C., Brodbeck, C., & Singer, T. (2011). Skin conductance response to the pain of others predicts later costly helping. *PloS one*, *6*(8), e22759.

Hess, U., & Fischer, A. (2014). Emotional mimicry: Why and when we mimic emotions. *Social and Personality Psychology Compass*, *8*(2), 45-57.

Heynen, E. J. E., Van der Helm, G. H. P., Stams, G. J. J. M., & Korebrits, A. M. (2016). Measuring Empathy in a German Youth Prison: A Validation of the German Version of the Basic Empathy Scale (BES) in a Sample of Incarcerated Juvenile Offenders. *Journal of Forensic Psychology Practice*, *16*(5), 336-346.

Hogan, R. (1969). Development of an empathy scale. *Journal of consulting and clinical psychology*, *33*(3), 307

Hristova, E., Kadreva, V., & Grinberg, M. (2014). Moral Judgments and Emotions: Exploring the Role of 'Inevitability of death' and 'Instrumentality of harm'. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).

Kennel, M. B., Brown, R., & Abarbanel, H. D. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, *45*(6), 3403.

Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2011). Utilitarian moral judgment in psychopathy. *Social cognitive and affective neuroscience*, *7*(6), 708-714.

Krosch, A., Figner, B., & Weber, E. U. (2012). Choice processes and their post-decisional consequences in morally conflicting decisions.

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of personality and social psychology*, *63*(2), 234.

Litten, V., Roberts, L. D., Ladyshewsky, R. K., Castell, E., & Kane, R. (2018). The influence of academic discipline on empathy and psychopathic personality traits in undergraduate students. *Personality and Individual Differences*, *123*, 145-150.

Lockwood, P. L., Bird, G., Bridge, M., & Viding, E. (2013). Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Frontiers in Human Neuroscience*, *7*, 760. http://doi.org/10.3389/fnhum.2013.00760

Malterer, M. B., Lilienfeld, S. O., Neumann, C. S., & Newman, J. P. (2010). Concurrent Validity of the Psychopathic Personality Inventory with Offender and Community Samples. *Assessment, 17*(1), 3–15. http://doi.org/10.1177/1073191109349743

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics reports*, *438*(5-6), 237-329.

Moore, A. B., Lee, N. L., Clark, B. A., & Conway, A. R. (2011). In defense of the personal/impersonal distinction in moral psychology research: Cross-cultural validation of the dual process model of moral judgment. *Judgment and Decision Making*, *6*(3), 186.

Moretto, G., Làdavas, E., Mattioli, F., & Di Pellegrino, G. (2010). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *Journal of cognitive neuroscience*, *22*(8), 1888-1899.

Mehrabian, A. (2000). *Manual for the Balanced Emotional Empathy Scale (BEES)*. Unpublished manuscript.

Nešić, M., Uljarević, M., Čičević, S., & Nešić, V. (2012). 36. Empathy as a predictor of skin conductance response to evocative movie scenes. *Clinical Neurophysiology*, *123*(7), e85.

Pino, M. C., De Berardis, D., Mariano, M., Vellante, F., Serroni, N., Valchera, A., Valenti, M., & Mazza, M. (2016). Two systems for empathy in obsessive-compulsive disorder: mentalizing and experience sharing. *Revista Brasileira de Psiquiatria*, *38*(4), 307-313.

Prguda, E., & Neumann, D. L. (2014). Inter-human and animal-directed empathy: A test for evolutionary biases in empathetic responding. *Behavioural processes*, *108*, 80-86.

Rose, L. T. (2016). *The end of average: How to succeed in a world that values sameness.* Penguin UK.

Rose, L. T., Rouhani, P., & Fischer, K. W. (2013). The science of the individual. *Mind, Brain, and Education*, 7(3), 152-158.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1-36.

Salas-Wright, C. P., Olate, R., & Vaughn, M. G. (2013). Assessing empathy in Salvadoran high-risk and gang-involved adolescents and young adults: a Spanish validation of the Basic Empathy Scale. *International journal of offender therapy and comparative criminology*, *57*(11), 1393-1416.

Schiepek, G., Aichhorn, W., Gruber, M., Strunk, G., Bachler, E., & Aas, B. (2016). Real-time monitoring of psychotherapeutic processes: concept and compliance. *Frontiers in psychology*, *7*, 604.

Team, R. (2015). RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www. rstudio. com*.

Van Orden, G. C. (Ed.). (2005). *Tutorials in contemporary nonlinear methods for the behavioral sciences*. National Science Foundation.

Van Rijn, S., Barendse, M., van Goozen, S., & Swaab, H. (2014). Social Attention, Affective Arousal and Empathy in Men with Klinefelter Syndrome (47,XXY): Evidence from Eyetracking and Skin Conductance. *PLoS ONE*, *9*(1), e84721. http://doi.org/10.1371/journal.pone.0084721

Vyas, K., Jameel, L., Bellesi, G., Crawford, S., & Channon, S. (2017). Derailing the trolley: Everyday utilitarian judgments in groups high versus low in psychopathic traits or autistic traits. *Psychiatry Research*, *250*, 84-91.

Wijnants, M., Cox, R., Hasselman, F., Bosman, A., & Van Orden, G. (2013). Does sample rate introduce an artifact in spectral analysis of continuous processes?. *Frontiers in physiology*, *3*, 495.

Appendix 1

Examples of dilemmas from Greene and colleagues' (2004) moral dilemmas

Non-moral

Reversed Turnips

You are a farm worker driving a turnip-harvesting machine. You are approaching two diverging paths. By choosing the path on the left you will harvest thirty bushels of turnips. By choosing the path on the right you will harvest fifteen bushels of turnips. If you do nothing your turnip-picking machine will turn to the left.

Is it appropriate for you to turn your turnip-harvesting machine to the right in order to harvest fifteen bushels of turnips instead of thirty?

Impersonal

Lost Wallet

You are walking down the street when you come across a wallet lying on the ground. You open the wallet and find that it contains several hundred dollars in cash as well the owner's driver's license. From the credit cards and other items in the wallet it's very clear that the wallet's owner is wealthy. You, on the other hand, have been hit by hard times recently and could really use some extra money. You consider sending the wallet back to the owner without the cash, keeping the cash for yourself.

Is it appropriate for you to keep the money you found in the wallet in order to have more money for yourself?

Personal

Architect

You are a young architect visiting one of your construction sites with your boss. Your boss is a despicable individual who makes everyone around him miserable including you. It occurs to

you that if you were to push him off of the building you are inspecting he would fall to his death and everyone would think it was an accident.

Is it appropriate for you to push your boss off of the building in order to get him out of your life?

Appendix 2

Cleckley's (1976) 16 characteristics of psychopathy

1. Superficial charm and good "intelligence"
2. Absence of delusions and other signs of irrational thinking
3. Absence of "nervousness" or psychoneurotic manifestations
4. Unreliability
5. Untruthfulness and insincerity
6. Lack of remorse or shame
7. Inadequately motivated antisocial behavior
8. Poor judgment and failure to learn by experience
9. Pathologic egocentricity and incapacity for love
10. General poverty in major affective reactions
11. Specific loss of insight
12. Unresponsiveness in general interpersonal relations
13. Fantastic and uninviting behavior with drink and sometimes without
14. Suicide rarely carried out
15. Sex life impersonal, trivial, and poorly integrated
16. Failure to follow any life plan

Appendix 3

All output from the moderation model in study 1

Table 1. Regression estimates, standard errors, z-values, p-values, and lower and upper confidence intervals of the moderation model in study 1.

	Estimate	Std.Err	z-value	P(>|z|)	ci.lower	ci.upper
Personal Moral dilemmas ~						
Affective Empathy	-0.153	0.136	-1.123	0.261	-0.419	0.114
Group	0.053	0.105	0.503	0.615	-0.152	0.257
Cognitive Empathy	0.087	0.127	0.689	0.491	-0.161	0.335
PCLR	-0.017	0.133	-0.127	0.899	-0.277	0.243
Impersonal Moral dilemmas ~						
Affective Empathy	-0.19	0.136	-1.396	0.163	-0.457	0.077
Group	0.029	0.105	0.278	0.781	-0.176	0.234
Cognitive Empathy	0.056	0.127	0.442	0.659	-0.192	0.304
PCLR	0.001	0.13	0.004	0.997	-0.254	0.256
Personal Moral dilemmas ~						
Group Cognitive interaction	-0.1	0.127	-0.788	0.431	-0.348	0.148
Group Affective interaction	0.035	0.136	0.254	0.8	-0.232	0.301
Impersonal Moral dilemmas ~						

Group Cognitive interaction	-0.02	0.127	-0.155	0.877	-0.268	0.229
Group Affective interaction	0.015	0.136	0.11	0.912	-0.252	0.282

Table 2. Covariances between the different types of moral dilemmas from the moderation model in study 1.

	Estimate	Std.Err	z-value	P(>|z|)	ci.lower	ci.upper
.Personal Moral dilemmas ~~						
.Impersonal Moral dilemmas	0.574	0.116	4.946	0	0.347	0.801

Table 3. The intercepts associated with the different types of moral dilemmas from the moderation model in study 1.

	Estimate	Std.Err	z-value	P(>|z|)	ci.lower	ci.upper
.Personal Moral dilemmas	0.019	0.105	0.185	0.853	-0.186	0.224
.Impersonal Moral dilemmas	0.003	0.105	0.033	0.974	-0.202	0.209

Table 4. The variances of the different types of moral dilemmas from the moderation model in study 1.

	Estimate	Std.Err	z-value	P(>|z|)	ci.lower	ci.upper
.Personal Moral dilemmas	0.952	0.14	6.781	0	0.677	1.227
.Impersonal Moral dilemmas	0.955	0.141	6.782	0	0.679	1.232

Table 5. The R^2 values associated with the different types of moral dilemmas from the moderation model in study 1.

	Estimate
Personal Moral dilemmas	0.035
Impersonal Moral dilemmas	0.032

Appendix 4

RQA formulas

Recurrence rate RR	The percentage of recurrence points in an RP: $$RR = \frac{1}{N^2} sum_{i,j=1}^{N} R_{i,j}$$ Corresponds to the correlation sum·
Determinism DET	The percentage of recurrence points which form diagonal lines: $$DET = \frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}$$ $P(l)$ is the histogram of the lengths l of the diagonal lines·
Laminarity LAM	The percentage of recurrence points which form vertical lines: $$LAM = \frac{\sum_{v=v_{\min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)}$$ $P(v)$ is the histogram of the lengths v of the vertical lines·
Entropy $ENTR$	The Shannon entropy of the probability distribution of the diagonal line lengths $p(l)$: $$ENTR = -\sum_{l=l_{\min}}^{N} p(l) \ln p(l)$$

$$\{T_k^{(1)} = j_{k+1} - j_k\}_{k \in \mathbb{N}}$$

$$\{T_k^{(2)} = j'_{k+1} - j'_k\}_{k \in \mathbb{N}}$$

- N – number of points on the phase space trajectory
- N_l – number of diagonal lines in the recurrence plot
- N_v – number of vertical lines in the recurrence plot
- $P(l), P(v)$ – histogram of the line lengths of diagonal/ vertical lines
- $\tilde{N}$ – maximal number of diagonals parallel to the LOI which will be considered for the calculation of $TREND$