COMT genotype is differentially associated with single trial variability of ERPs as a function of memory type.

Nowparast Rostami, Hadiseh; Saville, Christopher; Klein, Christoph; Ouyang, Guang; Sommer, Werner; Zhou, Changsong; Hildebrandt, Andrea

Biological Psychology

DOI:
10.1016/j.biopsycho.2017.06.002

Published: 01/07/2017

Version created as part of publication process; publisher's layout; not normally made publicly available

Title: COMT genotype is differentially associated with single trial variability of ERPs as a function of memory type.

Authors: Hadiseh Nowparast Rostami, Christopher W.N. Saville, Christoph Klein, Guang Ouyang, Werner Sommer, Changsong Zhou, Andrea Hildebrandt

PII: S0301-0511(17)30116-3
DOI: http://dx.doi.org/doi:10.1016/j.biopsycho.2017.06.002
Reference: BIOPSY 7387

To appear in:

Received date: 7-1-2016
Revised date: 19-6-2017
Accepted date: 19-6-2017

Please cite this article as: Rostami, Hadiseh Nowparast, Saville, Christopher W.N., Klein, Christoph, Ouyang, Guang, Sommer, Werner, Zhou, Changsong, Hildebrandt, Andrea, COMT genotype is differentially associated with single trial variability of ERPs as a function of memory type.Biological Psychology http://dx.doi.org/10.1016/j.biopsycho.2017.06.002

COMT genotype is differentially associated with single trial variability of ERPs as a function of memory type.

Hadiseh Nowparast Rostami[a,b]*, Christopher W. N. Saville[c]*, Christoph Klein[c,d,e], Guang Ouyang[f], Werner Sommer[a], Changsong Zhou[b], Andrea Hildebrandt[f]

* Equal contributions

[a] Department of Psychology, Humboldt-University at Berlin, Germany

[b] Department of Physics, Centre for Nonlinear Studies, Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Kowloon Tong, Hong Kong

[c] School of Psychology, Bangor University, Wales

[d] Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, University of Freiburg, Germany

[e] Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Medical Faculty, University of Cologne, Germany

[f] Department of Psychology, Ernst-Moritz-Arndt-Universität Greifswald, Germany

Corresponding author:

Christopher W. N. Saville

North Wales Clinical Psychology Programme

School of Psychology

Bangor University

Penrallt Road

Bangor, Gwynedd

LL57 3PY

Tel: +441248388740

Email: c.saville@bangor.ac.uk

Highlights

- Intra-subject variability of P300 latency is associated with COMT genotype.
- This association depends on the type of task demands.
- In n-back tasks, Val alleles are associated with lower P300 latency variability.
- In face recognition tasks, the association depends on whether faces were familiar.
- In both tasks, P3b-latencies predicted RTs on a trial-by-trial basis.

Abstract

Previous research on the association between intra-subject variability (ISV) in reaction times (RTs) and the Val^{158}Met polymorphism of the catechol-o-methyltransferase gene (COMT; rs4680) has yielded mixed results. The present study compared the associations between COMT genotype and ISV in P3b latency measured during working and secondary memory tasks using residue iteration decomposition (RIDE) of single trial latencies. We compared the outcome of the present analyses with a previous analysis of the same data (*N*=70, n-back tasks) using an alternative single-trial method. Additionally, we used RIDE to analyse the association between COMT genotype and ISV in an independent sample performing a different task (*N*=91, face-recognition task). Analyses reconfirmed previous results from the n-back tasks, showing that Val alleles are associated with lower ISV. In the face recognition tasks, genotype interacted with task conditions, so Val homozygotes had higher ISV to unfamiliar faces than familiar ones but Met carriers showed no effect of familiarity. Moreover, in both datasets trial-by-trial RTs were predicted by P3b latencies. Therefore, the present data suggests that associations between COMT genotype and ISV depend on the type of cognitive processes, which may explain heterogeneity in previous results.

Keywords: Event-related Potentials; Intra-subject variability; Memory; COMT polymorphism; Single trial analysis

1 Introduction

Intra-subject variability (ISV) in reaction times (RTs), that is, the degree of inconsistency in response speed during cognitive tasks, has long been neglected in favour of measures of average RTs. However, ISV appears not only to be a stable trait (Saville et al., 2011), rather than an error term, but is also elevated in a number of clinically relevant conditions, including ADHD (Klein et al., 2006; Kofler et al., 2013; Saville et al., 2015), schizophrenia (Rentrop et al., 2010), and terminal decline (MacDonald et al., 2008).

Converging evidence further suggests that ISV may be sensitive to the status of the catecholaminergic neuromodulatory system. Psychopharmacological studies in both healthy adults (Nandam et al., 2011; Rammsayer & Stahl, 2006) and psychiatric groups (Spencer et al., 2009) have shown reductions of ISV after administration of catecholaminergic agonists. Increased ISV, relative to healthy controls, has also been found for patients with Parkinson's disease, a disorder associated with loss of catecholaminergic neurons (de Frias et al., 2007).

Alongside psychopharmacological and patient studies, molecular genetics also suggests a link between ISV and catecholamine function. The $Val^{158}Met$ polymorphism of the catechol-*o*-methyltransferase (COMT) gene (rs4680) is a widely studied polymorphism, coding for an enzyme that breaks down catecholamines (dopamine, epinephrine, and norepinephrine) by methylation. This polymorphism is especially important for the deactivation of dopamine in the prefrontal cortex. Val homozygotes show three times greater COMT activity as compared with Met homozygotes, with intermediate levels for heterozygotes (Weinshilboum et al., 1999). Thus, dopamine availability in prefrontal cortex is lowest for Val/Val homozygotes, intermediate for Val/Met heterozygotes and greatest for Met/Met homozygotes.

The COMT polymorphism has been associated with ISV in several studies (e.g., Haraldsson et al., 2010; Stefanis et al., 2005). However, the results are contradictory. For example, relative to the low-activity Met allele, the high-activity Val allele has been associated with increased ISV in a one-back memory task (Stefanis et al., 2005), but with decreased ISV in an antisaccade paradigm (Haraldsson et al., 2010). Therefore, the relationship between catecholamines and ISV may be complex and different task demands may favour different catecholaminergic states (Cools & D'Esposito, 2011).

ISV in RTs can stem from increased latency variability in different stages of cognitive processing – including stimulus-related (perceptual), central (decision related), and response-related (preparation and motor) stages. One approach to identify at what stage of processing ISV arises is to use event-related potentials (ERP). By analysing covert sub-processes that correspond to different stages of processing, we can obtain a more fine-tuned view than using RTs alone.

The study of ERPs, however, has been dominated by the averaging technique. ERPs have low amplitude compared to the background electroencephalographic (EEG) signal and hence are difficult to measure accurately on a single trial level. Using averaging to obtain ERPs makes the simplifying assumption that fluctuations in voltage that occur with consistent timing across trials represent evoked activity, *i.e*. signal, while variation between trials likely represents unrelated background processes, *i.e*. noise. By averaging voltages across trials time-locked to a reference event, evoked activity is revealed because the background EEG is attenuated. This assumption has proven reasonable in a wide range of experimental settings, and the approach has been very productive. The assumption of invariance of the signal across

trials, however, is only an approximation and clearly inappropriate when ISV is the issue.

Instead of average ERPs, electrophysiological research into ISV has made use of single trial approaches. To this end, many researchers use time-frequency methods, and measure variability in phase at particular frequency bands, or other similar quantities (McLoughlin et al., 2014). Other researchers use what might be called a neoclassical approach to ERPs, aiming to identify waveforms similar to classical ERP components in single trials, and directly measure variability in latency or amplitude (Saville et al., 2011; Ouyang et al., 2011; Bender et al., 2015). For studying ISV at the electrophysiological level, and investigating its relationship with behaviour, along with its genetic basis and relations with psychiatric conditions, it is necessary to obtain a better understanding of how results obtained by different single trial approaches to ERPs correspond to one another. Widespread use of the average ERP means that researchers have a common frame of reference; this is not the case for single trial analysis. Furthermore, their relationships with behaviour need to be established in order to validate them and to investigate the commonalities and differences of the hitherto proposed single trial ERP methods.

Using the single trial approach of Saville et al. (2011), a method of identifying single trial ERP peaks using data aggregation across electrodes and peak picking, Saville et al. (2014) found Val genotypes to be associated with lower behavioural and electroencephalographic ISV, similar to the findings of Haraldsson et al. (2010), but different from those of Stefanis et al. (2005). The present study explored the replicability and generalizability of the findings of Saville et al. (2014) in two ways. Firstly, we reanalysed Saville et al.'s (2014) data using an alternative method of single trial analysis that has been evaluated multiple times in recent work. By

applying the Residue Iteration Decomposition (RIDE; Ouyang et al., 2011, Ouyang et al, 2015a) method, we will assess the generalizability of these findings across these two analysis approaches. The RIDE method was developed to decompose ERPs into component clusters that can be distinguished by their latencies being timelocked to stimulus onset (S component cluster) or response time (R component cluster) or by not being timelocked to either of those (C component cluster). In a RIDE framework, the most obvious measure of ISV is the latency variability of the C component cluster. Depending on the task, the C components cluster may capture different components but in many tasks the dominant component is the P3b or late positive complex.

Moreover, as discussed above, previous findings regarding the association between COMT Val^{158}Met genotype and ISV at the level of behaviour have been inconsistent, possibly due to differences in experimental paradigms or tasks. We therefore reanalysed a second dataset from a face recognition task, reported previously in Kaltwasser et al. (2014), but for an independent research question as compared with the present one. For the present analysis we applied RIDE method to the electrophysiological data and related the obtained ISV measures to the COMT genotypes, asking whether the associations found in Saville et al. (2014) would generalize to a different task requiring familiarity decisions about faces within a repetition priming paradigm.

Hence, Objective 1 of this study is to assess the relationship between ISV at the electrophysiological and behavioural level and to determine whether this relationship is moderated by tasks and COMT genotypes.

Secondly, Objective 2 of the present work is to assess the association between COMT genotypes and ISV, measured as trial-to-trial variability in the latency of the C component in two data sets: The original Bangor data (Saville et al., 2014) and a

second dataset collected in Berlin (Kaltwasser et al., 2014). Replication of Saville *et al*'s (2014) results in the same data with a new method will assess to what extent RIDE and Saville *et al*'s neoclassical approach give converging results. Extension in the Berlin data, which were obtained in a long-term memory retrieval task rather than in a short-term memory task, will assess the robustness and generalizability of the results with regard to the COMT polymorphism and ISV.

2 Materials and Methods

In this study we reanalysed two different datasets. One dataset was collected in Bangor to study the specific processes underlying the association between ISV and the COMT Val^{158}Met polymorphism and was reported in Saville et al. (2014) An analysis of the association between ISV and the ZNF804A gene was also conducted with this dataset (Saville et al., 2015).

The other dataset was collected in Berlin to study the neural mechanisms underlying individual differences in face cognition and was previously reported by Kaltwasser et al. (2014).

2.1 Participants

2.1.1 Bangor study

The Bangor sample consisted of 70 healthy young Caucasian adults, predominantly students from Bangor University. All participants had normal or corrected-to-normal visual acuity. The sample includes 18 (25.7 % of participants) Val/Val carriers (age 21.5 ±2.9, 41 % female, 88 % right-handed), 36 (51.4 % of participants) Val/Met carriers (age 21.3 ±2.8, 61 % female, 94 % right-handed), and 16 (22.8 % of participants) Met/Met carriers (age 20.1±1.8, 63 % female, 81% right-handed). Hardy-Weinberg equilibrium was assessed using a chi-squared test, and did

not appear to be violated (χ^2 = .06, p = .82). An additional three participants' data were excluded, one because genotyping failed and two because they reported ≤ 4 hours sleep the previous night.

2.1.2 Berlin study

The Berlin sample consisted of 91 healthy young Caucasian adults with heterogeneous occupational and educational backgrounds (38.2% with high school degree, 26.4% with university degree, 46.4% student, 34.5% employed, 13.6% unemployed). All participants had normal or corrected-to-normal visual acuity. The sample includes 20 (21.9 % of participants) Val/Val carriers (age 26.50 ± 4.98, 45.0% female, 90.0 % right-handed), 44 (48.3 % of participants) Val/Met carriers (age 27.04 ± 5.08, 45.4 % female, 81.8 % right-handed), and 27 (29.6 % of participants) Met/Met carriers (age 26.29 ± 4.40, 44.4 % female, 85.1 % right-handed. Hardy-Weinberg equilibrium assumptions did not also appear to be violated (χ^2 = .067, p = .79). An additional nineteen participants' data were excluded, five because genotyping failed, two for poor EEG data quality (assessed using visual inspection of averaged ERPs), two for error rates >30%, and ten for technical problems during the EEG recording.

2.2 Apparatus and materials

2.2.1 Bangor study

Direct-current EEG was recorded with 63 Ag/AgCl electrodes (Falk Minow, Germany) in 10-10 configuration (American Electroencephalographic Society, 1991) and two infra-orbital electrodes. Recording reference and ground electrodes were Cz and FPz, respectively. Impedance was reduced to ≤5 kΩ at all electrodes prior to data collection using Abralyt high-chloride gel (Easycap, Brain Products, Germany). Data

were differentially amplified using two BrainAmp DC amplifiers (.1 μV resolution, 1 kHz sampling rate, 250 Hz low-pass filter), and recorded on Brain Vision Recorder (both Brain Products, Germany). Participants were tested in a sound-attenuated Faraday cage, and stimuli were delivered on a 17" LCD monitor with an electrically shielded power source.

Saliva-derived genomic DNA samples were genotyped using Illumina BeadXPress Golden Gate assay to verify 96 SNPs, including COMT Val^{158}Met. Nucleic acid quality and concentration were evaluated using PicoGreen assay. Genotyping was conducted according to manufacturer's protocols. Genotype calling and annotation were performed using GenomeStudio assay with default settings (all Illumina, San Diego, USA).

2.2.2 Berlin study

EEG was recorded from 40 Ag/AgCl electrodes (Easycap, Brain Products, Germany), mounted on an elastic cap (Easycap, Brain Products, Germany) according to the extended 10-20 system (Pivik et al., 1993), as well as two electrodes positioned at the outer canthi of the eyes, and one infra-orbital electrode on the right. The ground electrode was AFz and the reference electrode was on the left mastoid. Impedance on each electrode was reduced to ≤5 kΩ prior to data collection. EEG data were amplified using BrainAmp DC amplifier (.1 μV resolution, 1 kHz sampling rate; Brain Products, Germany), and recorded on Brain Vision Recorder (software filters: 5 s low cutoff, 125 Hz high cutoff, 50 Hz notch; Brain Products, Germany).

DNA was extracted from buccal cells and purified using a commercial extraction kit (MagNA Pure LC DNA Isolation Kit; Roche Diagnostics, Mannheim, Germany). Genotyping of the COMT Val^{158}Met polymorphism (rs4680) was carried

out by real-time polymerase chain reaction using fluorescence melting curve detection by the Light Cycler System (Roche Diagnostics, Mannheim, Germany).

2.3 Stimuli and procedure

2.3.1 Bangor study

Participants performed three n-back tasks: a zero-back task (0BT), a one-back task (1BT), and a two-back task (2BT). In each task, participants responded to series of letters depending on whether each letter was an oddball or not. In the different tasks, oddballs were defined as follows: In the 0BT, the letter 'E' was the oddball, in the 1BT, letters matching the previous letter were oddballs, and in the 2BT, letters that matched the previous-but-one letter were oddballs. Oddball probability was 25% in all tasks. In all other trials, that is, standards, participants responded with the other hand (75% of trials). Participants were asked to respond as quickly and accurately as possible. They performed one block of each task, which was then repeated in the same order. There were five-minute breaks after the second and after the fourth block. Task order was counterbalanced across participants.

Stimuli were white Arial letters (visual angle: ~3°) on a black background. Stimulus duration was 1 s and stimulus-onset asynchrony jittered uniformly between 1950 and 2050 ms in 25 ms steps, averaging 2000 ms. Each block contained 280 trials. Stimuli were delivered using E-Prime V1.2 (Psychology Software Tools, PA, USA).

2.3.2 Berlin study

All stimuli were black and white portrait photographs of previously unfamiliar faces, fitted into a vertical ellipse of 259 × 388 pixels (7.0 × 10.2 cm). Photographs

were taken from two databases (Endl et al., 1998; Lundqvist et al., 1998), as well as a database collected by the authors and used in a series of large-scale studies, after ratings of stimuli distinctiveness (e.g., Hildebrandt et al., 2010). All portraits had the same gaze direction and were matched in terms of luminance. They all had neutral (un-emotional) expressions and no specific features like glasses or scars. Half of the portraits in each set were from females.

After a learning phase, the face recognition task was conducted. This task included four conditions: Primed-familiar (PF), primed-unfamiliar (PUF), unprimed-familiar (UPF), unprimed-unfamiliar (UPUF). Participants responded with a right-handed key press if a probe stimulus was one of ten familiar faces and a left-handed key press if the stimulus was a new, unfamiliar face. Probes were proceeded by a prime, which was either the same as the probe face (PF and PUF trials) or different. In the UPUF condition, primes were familiar faces, and in the UPF condition, primes were unfamiliar faces. In all trials there was always an unfamiliar face as a perceptual mask between the prime and target face. All faces in any given trial were the same sex, but the sex of faces was balanced and randomised across trials.

Trials began with a black fixation cross for 200 ms, followed by a prime face for 500 ms, then a mask of an unfamiliar face for 500 ms. The mask was then replaced by a fixation circle for 800 ms and a blank screen for 500 ms, before the probe appeared for 1500 ms. There was then an inter-trial interval of 1000 ms. The task was split into eight blocks of ten trials in each of the four conditions (PF, UPF, PUF, and UPUF), for a total of 80 trials per condition overall.

2.4 Pre-processing of EEG data

2.4.1 Bangor study

Continuous EEG data were pre-processed using Brain Vision Analyzer 2.0 (BVA; Brain Products, Germany). Any 200 ms window of data where amplitude varied by >1500 μV or <.5 μV was excluded to remove large artefacts. Infomax independent components analysis (ICA) was then run on 180 s of data starting at 120 s into each task for each participant, and factor loadings were applied to the rest of their dataset. Components representing oculomotor, cardiac, or electromyographic artefacts were identified via visual inspection of independent component topographies and timecourses and removed before remaining components were back-projected.

These data were average-referenced and .05–50.00 Hz zero-phase Butterworth filtered (24 dB/octave roll-offs), before a more stringent automatic artefact rejection procedure was run; rejecting data ranging >150 μV in any 200 ms window, as well as data 200 ms before and after this section. Stimulus-locked segments, ranging from 600 ms pre-stimulus until 1450 ms post-stimulus, were obtained from oddball trials with no rejected data and a correct response between 120 and 1400 ms post-stimulus. The period 600–400 ms pre-stimulus was used to baseline correct the segments. Segmented data were then exported to Matlab (MathWorks Inc., Natick, MA, USA), and epochs were shortened to -200–1450 ms, before RIDE was run on each participant's data.

2.4.2 Berlin study

Continuous EEG data were pre-processed using Brain Vision Analyzer 2.0 (BVA; Brain Products, Germany). After down-sampling data to 250 Hz (offline), ICA (ICA algorithm: Infomax [Gradient, Restricted]) was used to remove blinks, eye movements, and electrocardiac artefacts from the data, again identified via visual inspection of independent component topographies and timecourses. Data were then

filtered by a low-pass butterworth zero phase filter (30 Hz cutoff, 12 dB/oct roll-offs). After removing the electrooculography (EOG) channels (HO1, HO2, VO2), data were recalculated to common average reference, then segmented from 200 ms pre-probe onset to 2000 ms post-probe onset; only trials with a correct response between 100 and 2000 ms post-probe were kept.

Segments where amplitudes exceeded ±150 μV, with voltage steps >50 μV/ms, or amplitude ranges >200 μV within any 200 ms window, were rejected to eliminate residual artefacts. Segments were baseline-corrected using the 200 ms before probe onset.

2.5 Data Analysis

2.5.1 Single trial analysis

We applied RIDE, using the RIDE toolbox (Ouyang et al., 2015b) in Matlab (MathWorks Inc., Natick, MA, USA), to all pre-processed EEG trials, separately for each participant and condition, to extract the latencies of the single trial C component cluster, which mainly captures the P300 component. RIDE decomposes ERPs to stimulus-locked (S), intermediate (C), and response-locked (R) component clusters and also estimates the single trial latencies of C components using the following stepwise procedure: **Step 1** initially estimates the single trial C latencies by picking the peak latency of the cross-correlation curve between standard ERP as template and each single trial waveform (i.e., Woody filtering, Woody, 1967), where the correlation curves are averaged over all electrodes; the initial estimation of all RIDE components is set to zero (S = C = R = 0). **Step 2** calculates RIDE components in an inner loop. Each iteration calculates (1) the *S component cluster* waveform by subtracting C and R from EEG in each trial and synchronizing the residuals to

stimulus onset and calculating the median waveform, that is, the median values of the residuals at each time point; (2) *C component cluster* waveform by subtracting S and R from EEG in each trial, synchronizing the residuals to the C latencies, and calculating the median waveform; (3) *R component cluster* waveform by subtracting S and C from EEG in each trial and synchronizing the residuals to response onsets (reaction times) and calculating the median waveform. Iterations in this loop are continued until all RIDE components converge to stable waveforms. **Step 3** re-estimates C latency in each trial by calculating the peak latency of the cross-correlation between the converged C component in previous step as template and single trials after removal of converged S and R calculated in previous step. **Step 4** repeats the Step 2 and Step 3 iterations in several outer loop iterations using the information calculated in previous iteration until convergence is achieved for both single trial C latencies and RIDE component waveforms. Median waveforms are used to prevent low-frequency distortions (Ouyang et al., 2015a); in the final outer iteration median waveforms are restored to mean waveform.

The single trial latencies of the C component represent the lag of the C component in each trial, relative to the averaged C component across trials obtained after synchronizing all single trials to single trial latencies of C. A single-valued C latency for each trial is obtained by computing cross-correlation curves for all electrodes and averaging across electrodes.

In the RIDE toolbox (Ouyang et al., 2015b), it is necessary to define the time window for searching each RIDE component. As the objective of the present analysis was to characterise the variance of the C component latencies, fairly broad time windows were used to constrain RIDE. The S component cluster was assumed to occur 0–800 ms post-stimulus. The C component cluster was assumed to occur within

100–1200 ms post-stimulus. The R component was assumed to occur -400–400 ms relative to RT in each trial.

2.5.2 Inferential statistics

To assess the relationship between C and RT on a trial-by-trial basis (Objective 1) and the association between COMT genotypes and variability in C component latency (Objective 2), we fitted linear mixed effects models using the lme4 (Bates et al., 2015) package in the R environment (R Core Development Team, 2014).

To select the best model for each sample and objective, we started with a full model and removed terms until model fit became significantly worse. In each model the random effects structure was the maximal structure supported by data (*i.e.* where the model converged) and experimental design (Barr et al., 2013). After selecting the random effects structure, the non-significant fixed effects were dropped one by one and the models (sharing the same random effect structure) were compared using likelihood ratio test (using the `anova()` function in R) to assess whether dropping a fixed effect led to a significant difference in model fit. We also reported Bayesian Information Criteria (BIC; decrease in value is equivalent to increase in goodness of fit). The BIC values correct the log-likelihood statistic for the number of estimated parameters and also for the number of observations (BIC = -2logLik+nparam logNobservation; Schwarz, 1978). Models were fitted using maximum likelihood.

To assess the relationship between C latency and RT (Objective 1) in the Bangor data, our full model predicted RT using C-LATENCY, GENOTYPE (coded as number of Val alleles), LOAD (coded as a numeric variable), and the GENOTYPE*LOAD interaction term as fixed effects, with random intercepts and

LOAD slopes for each participant. The model for the Berlin study was identical but replaced LOAD with FAMILIARITY and PRIMING terms. The use of mixed effects models is particularly important here because such models account for the nesting of single trials within participant and trial type. Ignoring this nesting, would violate the assumption of independence that is made by other statistical methods

Regarding the association between COMT genotype and variability in C component latency (Objective 2), we computed the standard deviation of C component latencies (C variability) for each participant in each condition. The full model for the Bangor study predicted C variability using GENOTYPE (coded as number of Val alleles), LOAD (coded as a numeric variable), and the GENOTYPE*LOAD interaction term as fixed effects and random intercepts for each participant as well as random slopes for LOAD for each participant.

The full model for the Berlin study was identical but instead of LOAD it included FAMILIARITY (*familiar* [F] and *unfamiliar* [UF] faces); and PRIMING (*primed* [P] trials with the same face for prime and target and *unprimed* [UP] trials with different face for prime and target), as well as interactions between GENOTYPE and both FAMILIARITY and PRIMING. Random slopes for FAMILIARITY and PRIMING for each participant were also included in the models.

To assess whether any associations with genotype penetrated to the behavioural level in the Berlin study, similar models as for C latency variability were fitted to mean RT and standard deviation of the RT. Behavioural associations between COMT genotype and behavioural data in the Bangor sample can be found in Saville et al. (2014); in brief, Val/Val carriers were significantly less variable, faster, and more accurate than Met/Met carriers.

2.5.3 Consistency with Saville et al. (2014)

In addition to assessing whether the 'headline' results of Saville *et al*. (2014) could be replicated with RIDE, we wanted to assess the consistency between the C component latency variability measured here and the P3b latency variability reported in Saville et al. (2014). We thus computed Pearson's correlations between the EEG ISV parameters produced by the two different methods. For comparison, we also computed Pearson's correlations between the three tasks for each analysis method separately, to give a lower-bound indication of reliability.

3 Results

3.1 Objective 1: Relationship between C latency and RT as a function of task and genotype

In this section we will assess the relationship between RT and C latency, starting with the data from the Bangor study, followed by the Berlin study. Of special interest is the modulation of this relationship by tasks and genotype.

3.1.1 Bangor study

C latencies appear to have a Gaussian distribution, with greater dispersal as load increases and greater dispersal in the Met/Met group. Figure 1 presents probability density plots of single trial C latencies for each group on each task.

To assess the relationship between ISV in RT and ISV in C latency, the full model of RT included LOAD, C-LATENCY, GENOTYPE, C-LATENCY*GENOTYPE, LOAD*GENOTYPE and LOAD*C-LATENCY as fixed effects. In addition, random intercepts and slopes of LOAD were included for each participant (models including C-LATENCY as a random slope did not converge). In this model the LOAD*GENOTYPE, and LOAD*C-LATENCY interactions were not significant (p =

.81 and .40 respectively) and dropping these terms did not lead to a significant decline in goodness of fit (p = .69, BIC = 280808 in full model vs. 280789 in reduced model). Individually dropping C-LATENCY*GENOTYPE and LOAD decreased the goodness of fit in both cases (p < .01, BIC = 280841 and p < .01, BIC = 280882, respectively). Thus, the best-fitting model contained LOAD, C-LATENCY, GENOTYPE, and the C-LATENCY*GENOTYPE interaction (BIC = 280789). The significant interaction C-LATENCY*GENOTYPE (p < .01) suggests that C latency is not equally predictive of RT in all genotypes. C latency is less predictive of RT for participants with fewer Val alleles of the COMT polymorphism (Fig. 1). Table 1 shows the details of the best-fitting model.

-------- Insert Figure 1 about here --------

-------- Insert Table 1 about here --------

-------- Insert Figure 2 about here --------

3.1.2 Berlin study

The full model predicting single trial RT, consisted of random intercept for each participant and by-subject random slopes of FAMILIARITY and PRIMING, and fixed effects of FAMILIARITY, PRIMING, C-LATENCY, GENOTYPE, and all possible two-way interactions among them (models including random slopes of C-LATENCY did not converge). In this model, the main effect of GENOTYPE (p = .66), the interaction effects involving GENOTYPE (p [C-LATENCY*GENOTYPE] = .21; p [FAMILIARITY*GENOTYPE] = .86; p [PRIMING*GENOTYPE] = .53), and the C-LATENCY*PRIMING interaction (p = .14) were not significant, while the interaction of C-LATENCY*FAMILIARITY (p < .01) was significant. Dropping all non-significant terms did not lead to a significant decline in goodness of fit (p = .50; BIC = 345898 in

reduced model vs. 345945 in full model). Individually dropping C-LATENCY*FAMILIARITY and PRIMING diminished the goodness of fit in both cases ($p < .01$, BIC = 345920 and $p < .01$, BIC = 345902, respectively).

The best-fitting model included C-LATENCY, FAMILIARITY, PRIMING, and the interaction C-LATENCY*FAMILIARITY as fixed effects, and is shown in Table 2.

-------- Insert Table 2 about here --------

Figure 3a illustrates the fixed effect of C-LATENCY and its interaction with FAMILIARITY factor. The significant interaction C-LATENCY* FAMILIARITY ($p < .05$) suggests that C latency is more predictive of RT in familiar than in unfamiliar condition.

In addition, Figure 3b illustrates the lack of interaction between C latency and genotype. C latency is equally predictive of RT in all genotypes in the Berlin sample.

-------- Insert Figure 3 about here --------

3.2 Objective 2: Association between COMT Genotype and Variability in C Latency

3.2.1 Bangor study

Figure 4 presents C variability for all three genotype groups on all three tasks. Again, C variability appears to increase with increasing load and decrease with greater number of Val alleles.

-------- Insert Figure 4 about here --------

The best-fitting model predicting C latency variability included random intercepts and slopes of LOAD for each participant, included fixed effects of LOAD and GENOTYPE, but not their interaction (BIC = 1843.98), which outperformed the full model (BIC = 1849.14; $p = .66$). Dropping either LOAD (BIC = 1926.58; $p < .01$) or GENOTYPE (BIC = 1847.48; $p < .01$) led to significant drops in the goodness of fit.

The best model found that an increase in LOAD increased C latency variability and an increase in the number of Val alleles was associated with a decrease in C variability. Table 3 shows the details of the best-fitting model.

-------- Insert Table 3 about here --------

3.2.2 Berlin study

The full model for C latency variability included fixed effects of PRIMING, FAMILIARITY, GENOTYPE, PRIMING *FAMILIARITY, GENOTYPE*FAMILIARITY, and GENOTYPE* PRIMING and random intercepts for each participant. In this model, neither the main effect of GENOTYPE and PRIMING nor their interaction was significant. Thus, dropping the non-significant effects increased the goodness of fit (BIC = 2893.0 vs. 2898.8, $p = .76$). However, the main effect of FAMILIARITY, and the interaction effect of GENOTYPE*FAMILIARITY, and PRIMING *FAMILIARITY were significant (see Table 4) and dropping each term from the decreased model led to a significant change in model fit (BIC = 2934.5 ($p < .01$), BIC = 2893.1 ($p = .01$), BIC = 2897.7 ($p < .01$), respectively); therefore, the best-fitting model was the one including the main effects of FAMILIARITY, PRIMING, and GENOTYPE and the interactions of PRIMING *FAMILIARITY, and GENOTYPE*FAMILIARITY as fixed effects as well as random intercepts for each participant. The best model found that C latency is more variable for unfamiliar than familiar faces and that an increase in the number of Val alleles was associated with an increase in C variability only in unfamiliar condition. Table 4 shows the details of the best-fitting model.

-------- Insert Table 4 about here --------

Figure 5 illustrates the main effect and the interaction between genotype and FAMILIARITY for C latency variability.

-------- Insert Figure 5 about here --------

Table A1 and A2 in Appendix shows the details of the best-fitting model on mean RT and RT variability.

3.3 Consistency with Saville et al. (2014)

Correlation coefficients between C latency variability values reported here, and the standard deviation of P3b latency values reported in Saville et al. (2014), as well as within-method cross-task correlations can be seen in Table 5. While the within-method cross-task coefficients suggest promising psychometric reliability, as shown elsewhere for measures of ISV (Saville et al., 2011), the cross-method correlations are low, even when comparing the same raw data.

-------- Insert Table 5 about here --------

4 Discussion

The present study aimed to assess (1) the relationship between C latency and RT in individual trials and (2) the association between COMT Val^{158}Met genotype and ISV in C latency. A further aim was to investigate the degree of consistency between RIDE and the single trial method used by Saville et al. (2014).

It should be noted that, unlike in Saville et al. (2014), our focus on a component with a P3b-like topography was data-driven. We focused on this component because C is intuitively the most plausible correlate of ISV, but the topography of C emerged naturally from the data. Given COMT's known site of action in the prefrontal cortex (PFC), it is perhaps counter-intuitive that the COMT genotype is primarily associated with a component having a parietal topography. The

Bangor sample performed working memory tasks, the neural correlates of which are mainly linked to the PFC. The Berlin sample performed a face recognition task; recent studies showed that dopamine modulates the neural activity also in the fusiform gyrus (FFG) during face recognition (e.g., Rypma et al., 2015) and dopamine availability in the central nervous system has been shown to be affected by the COMT genotype (Egan et al., 2001). Thus, although we identify an association between COMT genotype and a P3b-like component with a largely parietal topography, this association may be a downstream consequence of differences in PFC activity.

4.1 Relationship between ISV on the Electrophysiological and Behavioural Levels

In Objective 1 we assessed the relationship between RT and the C component latency at the single trial level. We were especially interested in whether trial-to-trial variability in RT can be predicted by trial-to-trial variability the C component cluster.

The best-fitting models differed in the two studies, although C latency showed substantial predictive power in both samples. Consistent with these results, several studies have shown a positive correlation between P3b latency and RT in single trials, indicating that P3b latency and RT are sensitive to many of the same factors, such as stimulus classification demands (Kutas et al., 1977; McCarthy & Donchin, 1981; Polich, 2007). In addition, Ouyang et al. (2015a) showed that the latencies of the cognitive processes in each trial extracted by the RIDE method are highly correlated with RTs in various experimental conditions in most subjects.

In the Bangor sample, the significant interaction between COMT genotype and C latency showed that the latency of central cognitive processes was a stronger predictor of RT in Val homozygotes than Met allele carriers (Fig. 2a). In the Berlin sample, the interaction of C latency with the familiarity factor was significant, showing that C latency in familiar conditions,

which is less variable than in unfamiliar conditions, is a better predictor of RT than in the unfamiliar conditions. However, no interaction of COMT genotype and C latency in predicting RT was found.

The results from both data sets suggest that greater variability in RT likely represents increased variability in several sub-processes, such as perception, structural stimulus analysis, memory retrieval, response preparation, and execution, only some of which are indexed by C latency. Thus when ISV is high, the predictive validity of any given sub-component of the RT process (such as C latency) may be reduced.

4.2 Association between COMT Genotype and ISV on the Electrophysiological Level

We aimed to investigate whether Saville et al.'s (2014) finding of reduced inter-trial variability in P3b latency with more copies of the Val allele on the COMT Val^{158}Met polymorphism held when using an alternative single trial analysis approach (RIDE). We also aimed to assess whether or not these results generalised to a different task.

We found the same pattern of results in the Bangor data as reported by Saville et al. (2014) when using RIDE. There was a main effect of COMT genotype, with increased P3b/C component variability as the number of Met alleles increased. This shows that the findings about single trial P3b latency variability are robust across different analysis techniques. However we found low cross-method correlations for measures of ISV, suggesting that this may not always be the case.

In the Berlin dataset, which featured a different task and participants, the findings were more complicated. Although the main effect of the COMT genotype was not significant, we found a significant interaction between the COMT genotype and the face familiarity condition; thus the relationship between COMT genotype and ISV in this task depends on the

familiarity of the faces. In Val allele carriers, ISV was higher for unfamiliar faces than for familiar faces. Met/Met carriers, in contrast, showed no significant difference in their variability between familiar and unfamiliar faces.

Associations between COMT genotype and personality traits (see Montag, Jurkiewicz, & Reuter, 2012, for review) might explain why ISV is affected by familiarity for Val/Val carriers, but is not for Met/Met carriers. It has been shown that Val/Val genotype is associated with novelty seeking (e.g., Reuter & Henning, 2005; Tsai, Hong, Yu, & Chen, 2004). This personality trait is associated with avoidance of frustration and with exploratory activity in response to novel stimulation (Cloninger, Svrakic, and Przybeck, 1993). Hence, it can be speculated that when engaging in an exhaustive memory search, as required in the unfamiliar face cognition task, Val/Val carriers are more prone to frustration than Met carriers.

Further, the evidence from the Berlin sample showing the relationship between COMT genotype and ISV being dependent on the specific task demands is a possible explanation of the heterogeneity of results observed for ISV and COMT association in previous studies (Stefanis *et al.*, 2005; Haraldsson *et al.*, 2010).

As the COMT Val^{158}Met polymorphism has been shown to affect dopamine neurotransmission in the synaptic cleft, its influence on memory has long been of research interest (for an overview see Dickinson & Elvevåg, 2009). It has been shown in several studies that Met/Met carriers often perform better in cognitive tasks as compared with Val/Val, or even Val/Met carriers (e.g., Meyer-Lindenberg & Weinberger, 2006). However, the superiority of the Met/Met carriers in their cognitive performance still remains a controversial issue. Moreover, associations between COMT genotype and memory performance were mostly investigated in imaging studies, rather than neurophysiological and behavioral studies (Bishop, Fossella, Croucher, & Duncan, 2008).

The association between COMT genotype and ISV in n-back tasks and the absence of this association in the face recognition task might be explained by the different cognitive demands of the two paradigms. The n-back task, as applied in the present study, is a recognition working memory task. However, the applied face recognition task in the Berlin study is a specific form of a secondary memory task. Thus, the two tasks indeed measure different memory construct (e.g., Wilhelm et al, 2013), which may be differentially associated with the COMT gene.

In our n-back tasks, Val carriers outperform Met/Met carriers, with greater levels of accuracy, in contrast to some other studies (*e.g.* Egan et al., 2001). This might suggest that, in this task, dopamine levels in Met/Met carriers exceeded an optimal level, leading to poorer working memory performance (Barnett et al., 2008). Based on the theory of an inverted-U shaped dopamine actions on human working memory, the poorer cognitive functioning is associated with both too high and too low level of dopamine availability (e.g., Cools & Esposito, 2011).

4.3 Comparability of single trial analysis methods

Although we found the same association between COMT and ISV with RIDE as was found by Saville *et al*., (2014), the cross-method correlations between ISV measures were low (.01 < *r*s < .24). This is unlikely to be due to a general problem with reliability, as cross-task, within-method correlations, were substantially higher (RIDE: *r*s = .45 to .66, Saville *et al.*: *r*s = .67 to .75).

The two methods measure single trial latencies of somewhat different set of sub-processes, even if both forms of variability are similarly sensitive to differences with respect to COMT genotype. The C component cluster extracted by RIDE includes only sub-processes that are neither locked to the stimuli (e.g., perception processes) nor locked to the responses (e.g., motor processes). However, Saville et al.'s (2014) method measures 'raw' P3b components,

which are likely sensitive to stimulus-locked, response-locked and latency-variable sub-processes. Therefore, ISV estimated using RIDE might include less contribution from perceptual and motor sources, compared to Saville et al.'s method. Ouyang et al. (2017), comparing eight different algorithms for measuring ISV on ERP, showed that the latency-invariant ERP component cluster biases the measurement of the C component latency, which may explain the low correlation between these two methods. Moreover, the different distributions of C latencies, which appear largely Gaussian, and single trial P3b latencies estimated with Saville et al.'s (2014) method, which have the positive skew that is characteristic of RT distributions, also suggest that the two methods are likely sensitive to different aspects of single trial ERPs.

The low cross-method correlations suggest that different techniques could plausibly yield very different results and this is an area that needs a great deal more work. For a formal treatment and comparison of different single trial methods to estimate P3 latency please see Ouyang et al. (2017).

4.4 Limitations

The present study extends our knowledge about COMT genotype associations with ISV on the behavioural and ERP level and broadens our understanding on commonalities and differences of two recently suggested single trial ERP estimation methods. However, three limitations should be mentioned.

First, although the use of different paradigms applied to the two samples is relevant in the context of generalizability of COMT genotype's association with ISV, it weakens the possibility of replication. Differences in results between samples could be attributable to differences in the task paradigm or they could represent a simple failure to replicate. Along with differences in task demands, different paradigms elicit different ERP components, which require different EEG pre-processing, further hampering comparability. To advance the field,

further replication studies are needed that use identical or equivalent task paradigms as applied in previous studies in different labs, but with new samples of participants.

Second, the two labs used different EEG hardware and different data collection procedures. For example, the Bangor data were collected in a direct current recording with no analogue high pass filter, while the Berlin data were collected with such a filter. This then necessitated different offline digital filter settings in the two data analysis protocols. Furthermore, slightly different electrode montages and recording reference electrodes were used (although in both datasets, data were re-referenced offline to the common average reference). Such differences may seem to limit the comparability of the two datasets. That said, given the present study's focus on the C component generated using RIDE, we are confident that this is not a major problem here. Because the C component is usually dominated by activity in the 0-4 Hz range, data are filtered with a 4 Hz cut-off prior to Woody filtering, as part of the standard RIDE algorithm. Data are also linearly detrended. This is likely to minimize the impact of differences between the two datasets due to online versus offline filtering. In terms of the differences in recording montage and online referencing, we do not claim that the C component represents exactly the same component in the two datasets, due to the differences in task demands mentioned above, so the differences in recording protocols is less of an issue than it would be for a direct replication.

While one might argue that results that are not robust to variability in such procedures are questionable, these factors do represent alternative explanations for differences between findings in studies where other aspects (*e.g.*, task demands) also vary.

Third, both datasets share the modest statistical power common to the imaging genetics approach. The measurement of EEG phenotypes is time-consuming, and consequently studies often have lower sample sizes than one would wish for. Our attempt to integrate data across studies compensates somewhat, but the limited power of the studies

must be acknowledged as a shortcoming. We emphasize that the aim of the present study was to further investigate a previously reported association by using another single-trial analyses method and an additional dataset. Larger samples are sought for future research on gene-brain-behaviour variability studies.

4.5 Conclusion

To conclude, both studies showed that larger ISV in RT represents increased variability in several neurocognitive sub-processes in ERPs. In addition, in a reanalysis of Saville et al.'s (2014) data, we found that different analysis techniques both yielded lower ISV for carriers of the Val allele of the COMT Val^{158}Met polymorphism. However, despite similar 'headline' results, inter-technique correlations were low and this could not be explained low reliability. In a second dataset, using a face recognition task, we found no significant main effect association between COMT genotype and ISV, but an interaction between COMT genotype and familiarity suggested that the association between COMT and ISV may vary based on task demands. Multiple single trial ERP techniques appear to be sensitive to different aspects of ISV, but these different aspects may share common associations with catecholaminergic system. The use of complementary techniques may be a useful tool to better understand these associations in future work.

Acknowledgements: This work was partially supported by Hong Kong Baptist University (HKBU) Strategic Development Fund and Hong Kong RGC 12302914, partly supported by a grant from the Deutsche Forschungsgemeinschaft (HI 1780/2-1 & SO 177/26-1) to AH and WS and by a Research Group Linkage Project funded by the Alexander von Humboldt Foundation to CZ, AH, and WS. We are grateful to our anonymous reviewers for their constructive criticism.

5 References

American Electroencephalographic Society. (1991). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *2*, 200–202.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Mixed-effects modeling with R.

Bender, S., Banaschewski, T., Roessner, V., Klein, C., Rietschel, M., Feige, B., … Laucht, M. (2015). Variability of single trial brain activation predicts fluctuations in reaction time. *Biological Psychology*, *106*, 50–60. http://doi.org/10.1016/j.biopsycho.2015.01.013

Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, *69*(12), e113-25. http://doi.org/10.1016/j.biopsych.2011.03.028

de Frias, C. M., Dixon, R. a, Fisher, N., & Camicioli, R. (2007). Intraindividual variability in neurocognitive speed: a comparison of Parkinson's disease and normal older adults. *Neuropsychologia*, *45*(11), 2499–507. http://doi.org/10.1016/j.neuropsychologia.2007.03.022

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., … Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(12), 6917–22. http://doi.org/10.1073/pnas.111134598

Haraldsson, H. M., Ettinger, U., Magnusdottir, B. B., Sigmundsson, T., Sigurdsson, E., Ingason, A., & Petursson, H. (2010). Catechol-o-methyltransferase val158met polymorphism and antisaccade eye movements in schizophrenia. *Schizophrenia Bulletin*, *36*(1), 157–164. http://doi.org/10.1093/schbul/sbn064

Kaltwasser, L., Hildebrandt, A., Recio, G., Wilhelm, O., & Sommer, W. (2014). Neurocognitive

mechanisms of individual differences in face cognition: a replication and extension. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 861–78. http://doi.org/10.3758/s13415-013-0234-y

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry*, *60*(10), 1088–97. http://doi.org/10.1016/j.biopsych.2006.04.003

Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S. a, Friedman, L. M., & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: a meta-analytic review of 319 studies. *Clinical Psychology Review*, *33*(6), 795–811. http://doi.org/10.1016/j.cpr.2013.06.001

Macdonald, S. W. S., Hultsch, D. F., & Dixon, R. a. (2008). Predicting impending death: inconsistency in speed is a selective and early marker. *Psychology and Aging*, *23*(3), 595–607. http://doi.org/10.1037/0882-7974.23.3.595

McLoughlin, G., Palmer, J. a, Rijsdijk, F., & Makeig, S. (2014). Genetic overlap between evoked frontocentral theta-band phase variability, reaction time variability, and attention-deficit/hyperactivity disorder symptoms in a twin study. *Biological Psychiatry*, *75*(3), 238–47. http://doi.org/10.1016/j.biopsych.2013.07.020

Nandam, L. S., Hester, R., Wagner, J., Cummins, T. D. R., Garner, K., Dean, A. J., … Bellgrove, M. a. (2011). Methylphenidate but not atomoxetine or citalopram modulates inhibitory control and response time variability. *Biological Psychiatry*, *69*(9), 902–904. http://doi.org/10.1016/j.biopsych.2010.11.014

Ouyang, G., Herzmann, G., Zhou, C., & Sommer, W. (2011). Residue iteration decomposition (RIDE): A new method to separate ERP components on the basis of latency variability in single trials. *Psychophysiology*, *48*(12), 1631–47. http://doi.org/10.1111/j.1469-

8986.2011.01269.x

R Core Development Team. (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rammsayer, T., & Stahl, J. (2006). Sensorimotor effects of pergolide, a dopamine agonist, in healthy subjects: a lateralized readiness potential study. *Psychopharmacology*, *187*(1), 36–46. http://doi.org/10.1007/s00213-006-0400-9

Rentrop, M., Rodewald, K., Roth, A., Simon, J., Walther, S., Fiedler, P., … Kaiser, S. (2010). Intra-individual variability in high-functioning patients with schizophrenia. *Psychiatry Research*, *178*(1), 27–32. http://doi.org/10.1016/j.psychres.2010.04.009

Saville, C. W. N., Feige, B., Kluckert, C., Bender, S., Biscaldi, M., Berger, A., … Klein, C. (2015). Increased reaction time variability in attention-deficit hyperactivity disorder as a response-related phenomenon : evidence from single-trial event-related potentials. *Journal of Child Psychology and Psychiatry*, *56*(7), 801–813. http://doi.org/10.1111/jcpp.12348

Saville, C. W. N., Lancaster, T. M., Davies, T. J., Toumaian, M., Pappa, E., Fish, S., … Klein, C. (2015). Elevated P3b latency variability in carriers of ZNF804A risk allele for psychosis. *NeuroImage*, *116*, 207–213. http://doi.org/10.1016/j.neuroimage.2015.04.024

Saville, C. W. N., Lancaster, T. M., Stefanou, M. E., Salunkhe, G., Lourmpa, I., Nadkarni, a, … Klein, C. (2014). COMT Val158Met genotype is associated with fluctuations in working memory performance: converging evidence from behavioural and single-trial P3b measures. *NeuroImage*, *100*, 489–97. http://doi.org/10.1016/j.neuroimage.2014.06.006

Saville, C. W. N., Pawling, R., Trullinger, M., Daley, D., Intriligator, J., & Klein, C. (2011). On the stability of instability: Optimising the reliability of intra-subject variability of

reaction times. *Personality and Individual Differences*, *51*(2), 148–153. http://doi.org/10.1016/j.paid.2011.03.034

Spencer, S. V., Hawk, L. W., Richards, J. B., Shiels, K., Pelham, W. E., & Waxmonsky, J. G. (2009). Stimulant treatment reduces lapses in attention among children with ADHD: The effects of methylphenidate on intra-individual response time distributions. *Journal of Abnormal Child Psychology*, *37*(6), 805–816. http://doi.org/10.1007/s10802-009-9316-2

Stefanis, N. C., Van Os, J., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., & Stefanis, C. N. (2005). Effect of COMT Val158Met polymorphism on the continuous performance test, identical pairs version: Tuning rather than improving performance. *American Journal of Psychiatry*, *162*(9), 1752–1754. http://doi.org/10.1176/appi.ajp.162.9.1752

Figure Caption

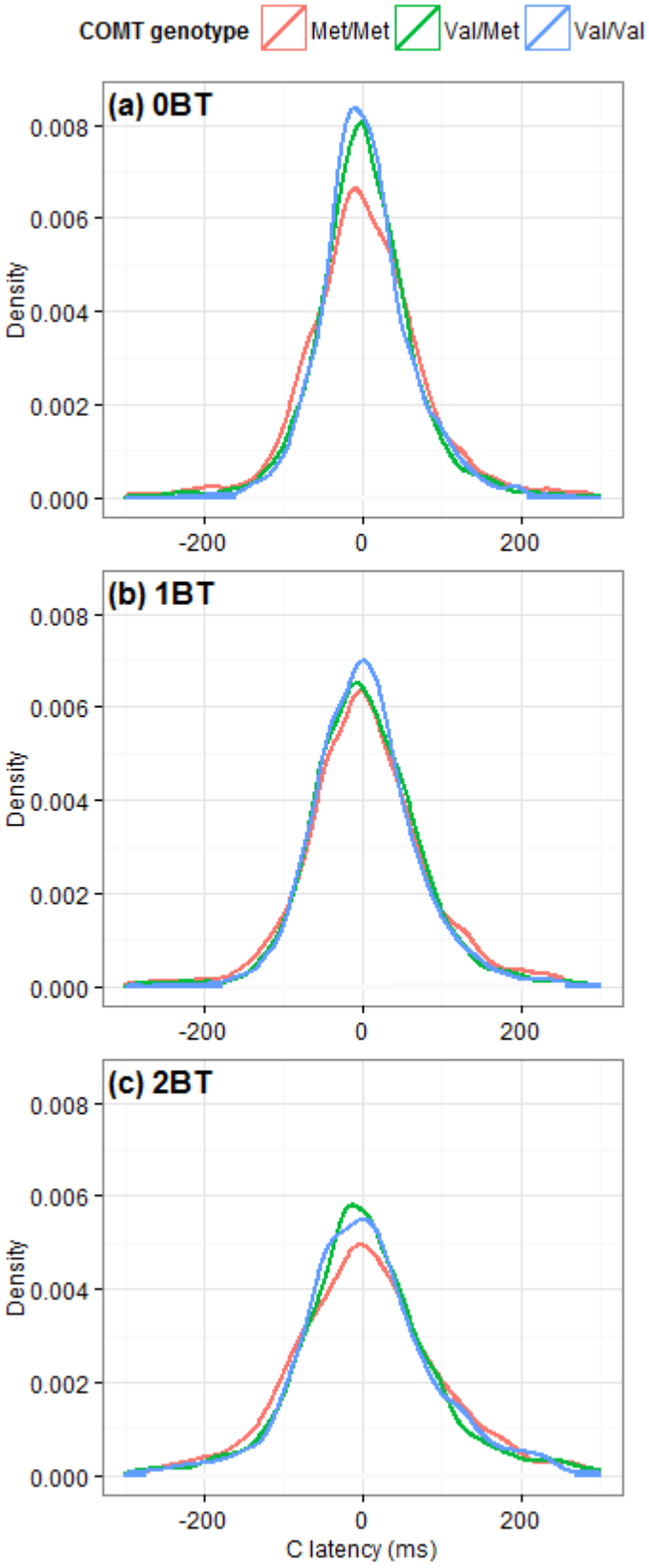
Figure 1. Bangor study; Probability densities of single trial C latencies for the three genotype groups in each task.

Figure 2. Bangor study. Scatter plot and the mean of C variability (standard deviation of C latencies) for all three genotype groups on all three tasks.

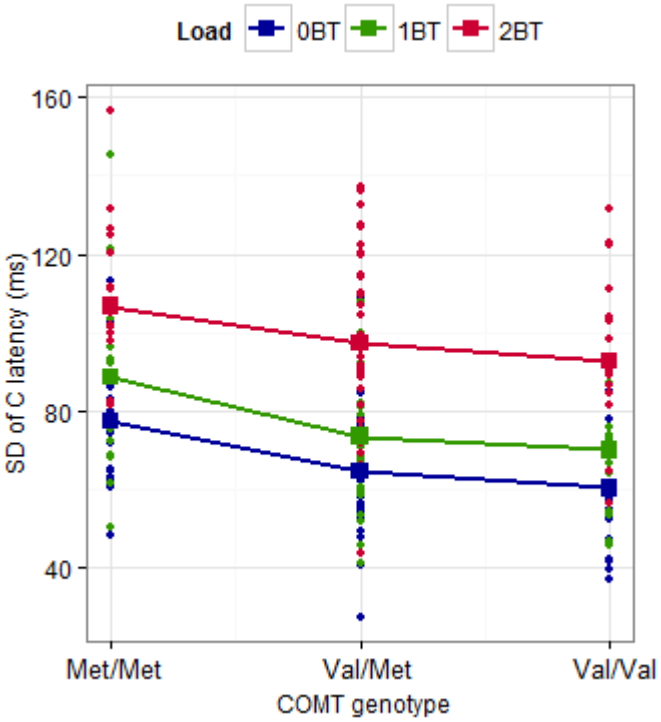
Figure 3. Berlin study. Scatter plot and the mean of C variability (standard deviation of C latencies) for all three genotype groups and different conditions.

Figure 4. Bangor study. Main effects of COMT genotype (a) and load (b) on the prediction of RT by C latency.

Figure 5. Berlin study. Effect of COMT genotype (a), familiarity (b), and priming (c) on the prediction of RT by C latency.

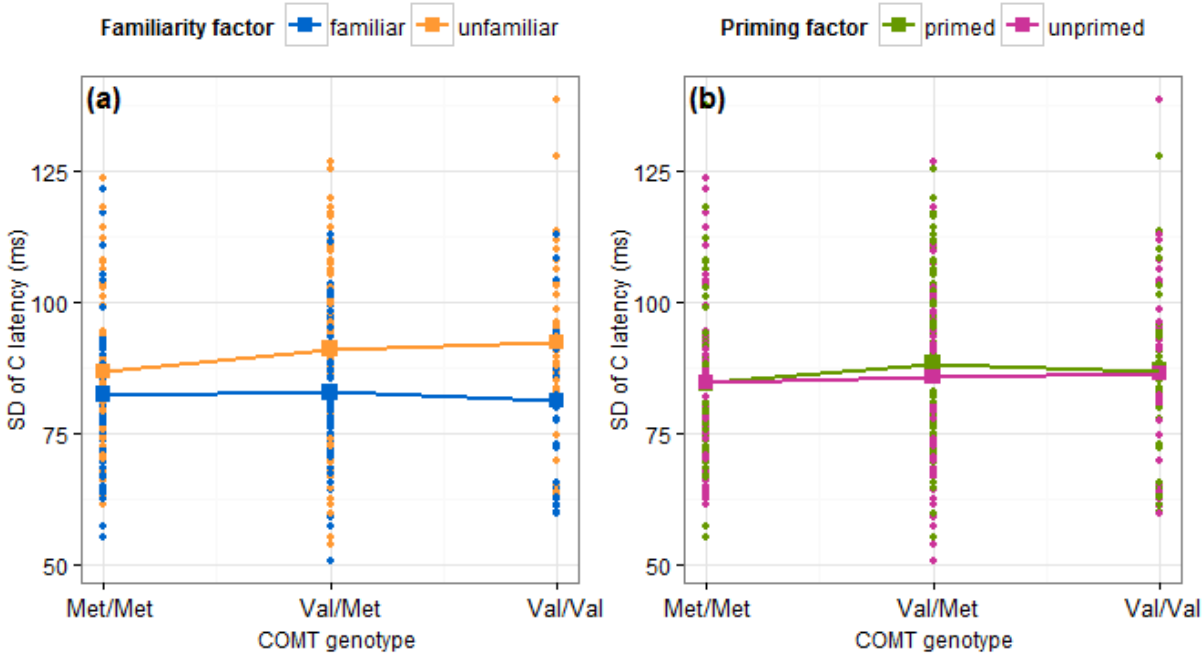


Figr-1

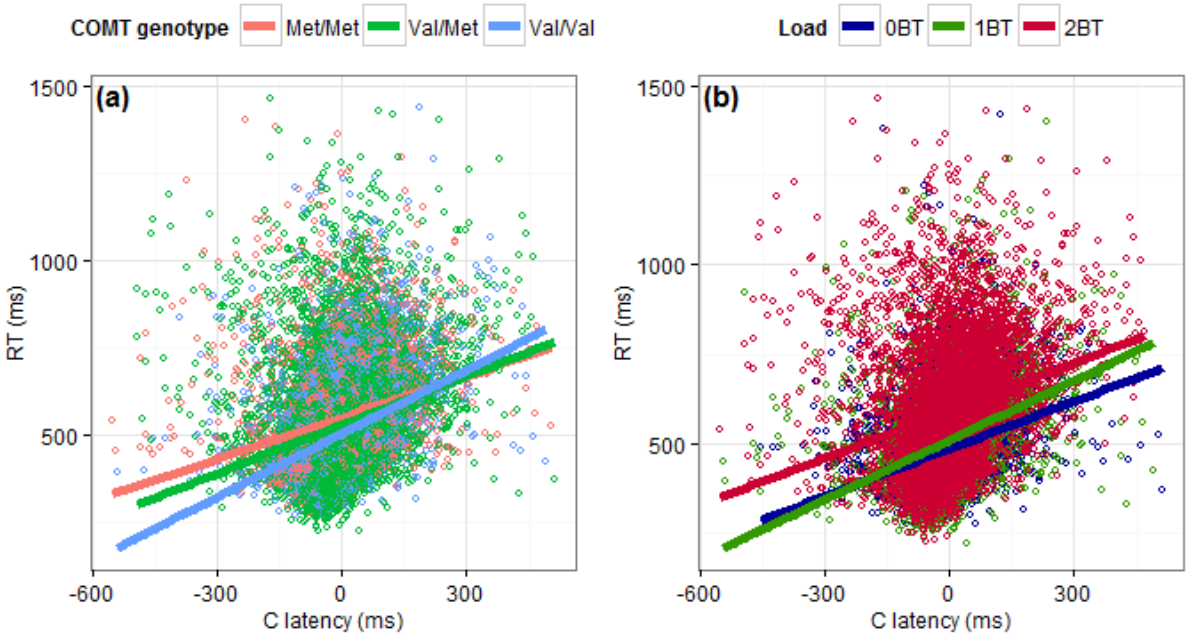


Figr-2

Figr-3



Figr-4



Figr-5

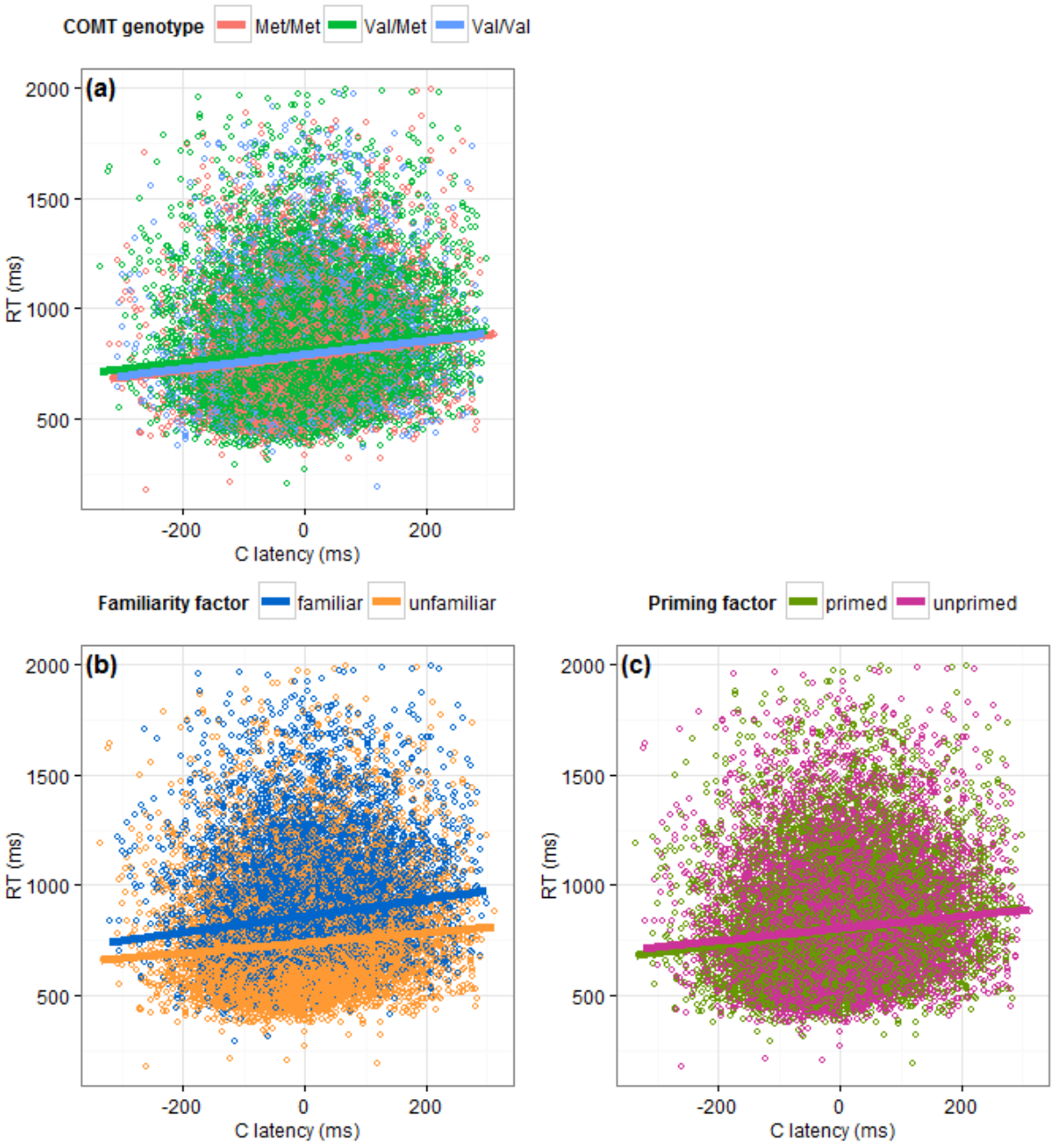


Table 1. Bangor study. Linear mixed effects models statistics for C latency variability measures as dependent variables

Model: SD_C ~ GENOTYPE+LOAD+(1+LOAD|subjects)

Fixed effects	**Estimate**	**Std. Er.**	***t*-value**
Intercept (F, P)	72.95	3.48	20.92***
COMT Genotype (slope)	-8.34	2.71	-3.07**
Load	15.91	1.23	12.92***
Random effects	**Variance**	**Std. Dev.**	
Subject (intercept)	136.35	11.67	
Load (slope)	4.43	2.1	
Residual	203.58	14.26	

***$p < .0001$, ** $p < .001$, * $p < .01$

Table 2. Berlin study. Linear mixed effects models statistics for C variability measures as dependent variables

Model (main effects are included):

SD_C ~ GENOTYPE*FAMILIARITY+PRIMING*FAMILIARITY+(1|subjects)

Fixed effects	**Estimate**	**Std. Er.**	***t*-value**
Intercept (F, P)	81.89	2.41	32.93***
COMT Genotype (slope)	-.53	1.97	-.27
Familiarity (UF)	7.90	1.89	4.18***
Priming (UP)	1.88	1.39	1.35
COMT Genotype: Familiarity (UF)	3.37	1.38	2.44*
Familiarity (UF): Priming (UP)	-6.49	1.97	-3.28**
Random effects	**Variance**	**Std. Dev.**	
Subject (intercept)	137.52	11.72	
Residual	88.87	9.42	

***$p < .0001$, ** $p < .001$, * $p < .01$

Table 3. Bangor study. Linear mixed effects models statistics for RT measures as dependent variables

Model (main effects are included):

RT ~ LOAD+GENOTYPE*C-LATENCY+(1+LOAD|subjects)

Fixed effects	**Estimate**	**Std. Er.**	***t*-value**
Intercept (Load 0; MM group)	507.60	12.62	40.23***
C Latency	.41	.01	23.99***
Load	54.25	3.54	15.32***
COMT Genotype (VM)	-33.17	15.12	-2.19*
COMT Genotype (VV)	-53.47	17.30	-3.09**
C Latency: COMT Genotype(VM)	.05	.02	2.39*
C Latency: COMT Genotype(VV)	.20	.02	8.00***
Random effects	**Variance**	**Std. Dev.**	
Subject (intercept)	2496	49.96	
Load (slope)	821	28.65	
Residual	11510	107.29	

***$p < .0001$, ** $p < .001$, * $p < .01$

Table 4. Berlin study. Linear mixed effects models statistics for RT measures as dependent variable

Model (main effects are included):

RT ~ C-LATENCY*FAMILIARITY+PRIMING+(1+(FAMILIARITY+PRIMING)|subjects)

Fixed effects	**Estimate**	**Std. Error**	***t*-value**
Intercept (F, P)	857.80	13.12	65.37***
C latency	.37	.01	19.02***
Familiarity UF	-120.30	2.32	-51.64***
Priming UP	8.98	2.32	3.86***
C latency: Familiarity UF	-.14	.02	-5.42***
Random effects	**Variance**	**Std. Dev.**	
Subject (intercept)	15292	123.70	
Residual	35209	187.60	

*** $p < .0001$, ** $p < .001$, * $p < .01$

Table 5. Cross- and within-method correlations

Cross-method correlation

		RIDE method		
		0BT	**1BT**	**2BT**
Saville et al.	**0BT**	.19	.18	.06
(2014)	**1BT**	.17	.24	.01
Method	**2BT**	.08	.09	.09

Within-method correlation

Above diagonal: Saville et al. method; below diagonal: RIDE

	0BT	**1BT**	**2BT**
0BT	1	.75	.67
1BT	.66	1	.69
2BT	.52	.45	1