THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Determinants and Correlates of Intra-individual Variability in Reaction Time

Dominika Dykiert

Doctor of Philosophy
The University of Edinburgh
2011

Declaration

This thesis has been composed by me and the work presented in it is my own. The study presented in Chapter 6 has been carried out in collaboration with the University of Edinburgh Wilderness Society, who collected the data during the the Edinburgh Altitude Research Expedition 2008 to the Western Himalayas. My contribution involved preparing and analysing the data and writing of the research report, which has subsequently been published in a peer-reviewed journal, *Neuropsychology* (http://psycnet.apa.org/journals/neu/24/3/391/). The work presented in this thesis has not been submitted for any other degree or professional qualification.

Dominika Dykiert

Abstract

Traditionally, reaction time (RT) was conceived of as an average speed of a number of responses made by an individual, or mean RT. Increasingly, however, intra-individual variability in reaction time (RT IIV) – the consistency of responses by a single person across trials – is used as an additional or even alternative measure. RT IIV is often found to be elevated in a number of conditions that affect the central nervous system functioning, such as traumatic brain injury or neurodegenerative diseases. It can predict change in cognitive performance in ageing, progression from normal ageing to mild cognitive impairment, and even death. Therefore, RT IIV may be of great practical importance. However, RT IIV and mean RT are correlated; therefore it is often problematic to draw conclusions about unique associations between these and other variables. One objective of the work presented in this thesis was to investigate determinants and correlates of simple and choice RT IIV and to test which associations may be accounted for by the individual differences in mean RT. The first investigation was concerned with age differences in RT IIV. Following a systematic review of literature, a series of meta-analyses demonstrated that older individuals (aged 60 years and above) have greater RT IIV than young or middle-aged adults in simple and choice RT tasks. The effects were reduced but still significant when RT IIV was adjusted for mean RT. The next study was a cross-sectional investigation of the associations between age and RT IIV, as well as of sex differences in RT IIV, across the lifespan in participants ranging in age from 4 to 75. Non-linear effects of age were found for RT IIV measures, such that variability decreased with age in children and increased with age in older adults. A novel finding from this study was that sex differences in RT IIV were present among adults but not children, suggesting that there might be an age threshold at which sexes diverge in their RT IIV trajectories. The results also indicated that findings regarding RT IIV may differ depending on the variability measure used (that is, whether and how mean RT is controlled). The second study on the same sample investigated variability on a trial-by-trial basis. Specifically, it tested the hypothesis that sex differences in variability are due to females being disproportionately slower at the first trial which inflates their overall RT IIV. This hypothesis was not supported. Another

investigation used longitudinal data from the West of Scotland Twenty-07 study. Three cohorts of individuals aged approximately 15, 35 and 55, were followed up for 20 years and had RT data collected at four occasions. Analyses confirmed non-linear effects of age on RT IIV found in the earlier cross-sectional investigation. The final study investigated the effect of high altitude on RT IIV. It found that altitude-related increase in RT IIV is fully accounted for by general slowing of RT at high altitude. The overall pattern of results obtained from the investigations suggests that RT IIV increases with age in adults and that not all of the increase is due to general slowing. Moreover, the results show that sex differences in RT IIV are not uniform across the lifespan. Finally, whereas associations of RT IIV with some variables, for example age, are relatively robust to controlling for mean RT, others are fully attenuated by such practice.

Acknowledgements

I am greatly indebted to my supervisors, John Starr, Ian Deary, and Geoff Der, for their expert advice, guidance and encouragement throughouth this project.

Thank you also to David Hall, Nikki van Gemeren, and Richard Benson for organising the Edinburgh Altitude Research Expedition 2008 and for collecting the reaction time data at different altitudes.

I would like to extend my thanks to the numerous Authors who kindly provided their original datasets and additional information required for the systematic review and meta-analysis.

Finally, my sincere gratitude to all those who were there for me during the completion of this thesis (you know who you are)—your support is greatly appreciated.

Table of Contents

List of Tables ix
List of Figures xi

Chapter 1 - Introduction 1

A brief history of reaction time 1
Types of variability 2
Sub-types of intra-individual variability 3
There is stability in variability 5
The importance of RT IIV 6
RT IIV as an index of cognitive functioning 7
Intelligence 7
Specific cognitive abilities 9
RT IIV as a marker of neurological functioning 11
Healthy ageing 11
Brain damage 13
Dementia 16
Mild cognitive impairment 18
Parkinson's Disease 20
Bringing it together 21
Predictive value of RT IIV 21
Methodological issues in the study of IIV 23
Correlation with the RT mean 23
Skewed distribution 25
Task and procedural factors 25
Different measures of intra-individual variability 26
Raw measures of IIV 26
Intra-individual standard deviation and intra-individual variance 26
Inter-quartile range and percentile differences 27
Methods of controlling for mean RT 27
Coefficient of variation 27
Including mean RT as a covariate 28
Regression method and purified residuals 28
Whole distribution analysis 29
Outstanding questions; unresolved issues 29
Age differences in RT IIV 29
Sex differences in RT IIV 30
Disentangling variability from mean 31

Chapter 2 - Age effects on intra-individual variability in simple and choice reaction time: A systematic review 34

Magnitude of age differences in RT IIV and the effects of procedural factors 35
Comparing the old with the young and the middle-aged 37
Different measures of RT IIV 37
Methods 39
Search strategy 39

Study selection 40
General inclusion criteria 40
Ageing-specific inclusion criteria 42
Multiple publications 42
Quality assessment 42
Data extraction 43
Results 45
SRT IIV 46
Raw SRT IIV: old versus young 47
Raw SRT IIV: old versus middle-aged 49
Mean-adjusted SRT IIV: old versus young 49
Mean-adjusted SRT IIV: old versus middle-aged 50
Attenuation of age group difference effect sizes by adjusting SRT IIV for SRT mean 50
Evidence from studies not included in meta-analyses of SRT IIV 51
CRT IIV 52
Raw CRT IIV: old versus young 53
Raw CRT IIV: old versus middle-aged 55
Mean-adjusted CRT IIV: old versus young 55
Mean-adjusted CRT IIV: old versus middle-aged 56
Attenuation of age group difference effect sizes by adjusting CRT IIV for CRT mean 57
The number of possible choices in CRT 57
Evidence from CRT IIV studies not included in meta- analyses 58
Assessing publication bias 59
Discussion 60
Explaining heterogeneity 63
Publication bias 66
Strengths and limitations 66
Summary 67

Chapter 3 - Sex differences in reaction time mean and intra-individual variability across the lifespan 95

Reaction time and age 96
Reaction time intra-individual variability and age 97
Sex differences in reaction time 98
Sex differences in reaction time through the lifespan 99
The relationship between mean reaction time and intra-individual variability 102
The present study 103
Method 104
Participants 104
Reaction time tests and procedures 105
Data preparation 107
Statistical analyses 108
Results 109
Age and sex effects on reaction time through the lifespan 110
Age and sex effects in childhood and adolescence 111
Age and sex effects in adulthood 112
Indices of Multiple Deprivation and reaction time measures 113

Discussion 113
Sex differences in reaction time 113
Age effects on reaction time 116
Reaction time intra-individual variability controlling for reaction time mean 119
Indices of Multiple Deprivation and reaction time measures 121
Methodological issues 122
Strengths and limitations 123
Conclusion 124

Chapter 4 - Age and sex effects on reaction time variability: Insights from a longitudinal Twenty-07 study 141

Method 144
Participants 144
Reaction time measures 145
Data preparation 146
Statistical analysis 147
Results 147
SRT IIV 147
CRT IIV 148
Discussion 150
Cross-sectional and longitudinal effects of age 150
Sex differences 152
Raw versus mean-adjusted measures of RT variability 153
Strengths and limitations 154
Conclusion 156

Chapter 5 - Sex differences in reaction time trial by trial 161

Method 163
Participants 163
Reaction time measures 163
Statistical analysis 164
Results 164
Discussion 166
Main findings 166
Other findings 170
Concluding remarks 171

Chapter 6 - The effects of high altitude on mean and intra-individual variability in choice reaction time 176

Method 180
Participants 180
Measures 181
Reaction time 181
Vital signs 182
Lake Louise score 182
Altitude 182
Time 183

Procedure 183
Statistical analyses 184
Results 185
Correlations between the variables 185
Mean reaction time 186
The time dimension 186
Altitude and other variables 187
Reaction time variability 188
The time dimension 188
Altitude and other variables 189
Discussion 190
Correlations between the variables 191
Reaction time mean 192
Reaction time intra-individual variability 193
Strengths and limitations 195
Future research 197
Conclusion 197

Chapter 7 - General Discussion 208

Summary of findings 208
Contributions to the existing knowledge 211
Age and RT IIV 211
Sex and RT IIV 212
Controlling for mean RT 213
Task complexity and task characteristics 214
Practice effects 216
Future directions and recommendations 218
Final summary 219

References 221

Appendices 233
Appendix A. Search terms used for each database 234
Appendix B. Data extraction sheet 236
Appendix C. Details of data extraction from studies for which sufficient information were not reported in the paper nor obtainable from the authors. ... 239
Appendix D. Excluded studies - general search 242
Appendix E. Retained studies from the general search considered for the healthy ageing review 273
Appendix F. Considered studies from the update search 282

List of Tables

Table 2.1. Sample characteristics of studies included in the review of simple and choice reaction time intra-individual variability studies.....69

Table 2.2. Summary of task characteristics and data available from simple reaction time intra-individual variability studies.....72

Table 2.3. Summary of subgroup analysis results for simple reaction time intra-individual variability studies.....75

Table 2.4. Summary of task characteristics and data available from choice reaction time intra-individual variability studies.....76

Table 2.5. Summary of subgroup analysis results for choice reaction time intra-individual variability studies.....81

Table 3.1. Sample size by age groups.....126

Table 3.2. Transformation parameters (λ) used for Box-Cox transformation..... 127

Table 3.3. Means and standard deviations (in seconds) of reaction time measures for both genders in three broad age groups.....127

Table 3.4. Summary statistics of reaction time models for the whole sample.....128

Table 3.5. Summary statistics of reaction time models for ages below 18 years.....130

Table 3.6. Summary statistics of reaction time models for ages 18 and above..... 131

Table 3.7. Correlations between reaction time measures and Index of Multiple Deprivation scores (overall and for each domain) by country.....132

Table 4.1. Means and standard deviations of reaction time (in miliseconds) for males and females, different cohorts and at different waves.....158

Table 4.2. Effect sizes (Cohen's d) for wave-to-wave differences in reaction time intra-individual variability measures..... 158

Table 5.1. Group means and standard deviations of choice reaction time measures (in seconds) for males and females and different age groups.....172

Table 5.2. Summary statistics of the final linear mixed model..... 173

Table 6.1. Correlations between the variables..... 198

Table 6.2. Summary statistics of fixed effects in models of mean reaction time.....199

Table 6.3. Summary statistics of fixed effects in models of reaction time intra-individual standard deviation..... 199

List of Figures

Figure 1.1. A hypothetical trajectory of performance across measurement occasions of individuals with equal mean performance but different variability..... 32

Figure 1.2. Three-dimensional data box (based on Cattell, 1966; Hultsch & MacDonald, 2004), illustrating different types of variability.....33

Figure 2.1. Flow diagram of the review process.....83

Figure 2.2. Effect sizes for old versus young comparisons in raw simple reaction time intra-individual variability studies..... 84

Figure 2.3. Scatterplots of effect sizes for old-young differences in raw and mean-adjusted simple reaction time intra-individual variability (SRT IIV) and the three covariates: old group age mean, number of trials and the length of preparatory interval..... 85

Figure 2.4. Effect sizes for old versus middle-aged comparisons in raw simple reaction time intra-individual variability studies..... 86

Figure 2.5. Effect sizes for old versus young comparisons in mean-adjusted simple reaction time intra-individual variability studies..... 87

Figure 2.6. Effect sizes for old versus middle-aged comparisons in mean-adjusted simple reaction time intra-individual variability studies..... 88

Figure 2.7. Effect sizes for old versus young comparisons in raw choice reaction time intra-individual variability studies..... 89

Figure 2.8. Scatterplots of effect sizes for old-young differences in raw and mean-adjusted choice reaction time intra-individual variability (CRT IIV) and the three covariates: old group age mean, number of trials and the length of preparatory interval.....90

Figure 2.9. Effect sizes for old versus middle-aged comparisons in raw choice reaction time intra-individual variability studies..... 91

Figure 2.10. Effect sizes for old versus young comparisons in mean-adjusted choice reaction time intra-individual variability studies.....92

Figure 2.11. Effect sizes for old versus middle-aged comparisons in mean-adjusted choice reaction time intra-individual variability studies..... 93

Figure 2.12. Funnel plots showing little evidence of publication bias of raw or mean-adjusted simple reaction time intra-individual variability studies, with some bias apparent among studies of raw and mean-adjusted choice reaction time intra-individual variability..... 94

Figure 3.1. Number of representatives from each quintile of Index of Multiple Deprivation by country and sex..... 133

Figure 3.2. Reaction time apparatus.....134

Figure 3.3. Means of reaction time measures across age (grouped in 1-year bands up to 14 and 5-year bands thereafter)..... 135

Figure 3.4. Predicted values and confidence bands for reaction time measures by age and sex.....136

Figure 3.5. Observed and predicted vales of reaction time measures for males and females aged under 18.....137

Figure 3.6. Predicted values and their corresponding 95% confidence bands of simple and choice reaction time coefficients of variability (SRT CV, CRT CV) and intra-individual standard deviation from models controlling for mean reaction time for participants aged under 18..... 138

Figure 3.7. Observed and predicted values of reaction time measures for males and females aged 18 or over..... 139

Figure 3.8. Predicted values and their corresponding 95% confidence bands of simple and choice reaction time coefficients of variability (SRT CV, CRT CV) and intra-individual standard deviation from models controlling for mean reaction time for participants aged 18 and over..... 140

Figure 4.1. Distribution of ages in each cohort at the first wave of testing..... 159

Figure 4.2. Means of reaction time intra-individual variability measures for males and females..... 160

Figure 5.1. Mean choice reaction times (CRT) in seconds at each test trial for males and females. Error bars represent standard errors..... 174

Figure 5.2. Mean choice reaction time (CRT) in seconds for each previous-current preparatory interval (PI) combination..... 175

Figure 6.1. Portable reaction time testing device used in the present study..... 200

Figure 6.2. Altitude profile of the measurement occasions..... 201

Figure 6.3. Group average mean reaction time (RT) and intra-individual standard deviation (RT ISD) across the measurement occasions with superimposed altitude at which the tests were taken..... 202

Figure 6.4. Individual trajectories of mean RT change across measurement occasions with fitted regression line..... 203

Figure 6.5. Predicted mean values of mean reaction time based on Model 1.1.....204

Figure 6.6. Predicted mean reaction time at different altitudes..... 205

Figure 6.7. Individual trajectories of RT ISD change across measurement occasions with fitted regression lines..... 206

Figure 6.8. Predicted mean values for log-transformed reaction time intra-individualstandard deviation (RT ISD) based on Model 2.1..... 207

Chapter 1 -
Introduction

A brief history of reaction time

The interest in individual differences in reaction time (RT) has a long history, dating back to even before the existence of accurate chronometers. In the early 19th century, a Prussian astronomer, F.W. Bessel, noted considerable variability in the way different assistants logged the exact moment that a star crossed the fixed transit line of a telescope. This effect was relatively consistent across individuals, and Bessel's solution was to apply a *personal equation* that calibrated out the individual observers' measurements (Brebner & Welford, 1980; Jensen, 2006).

It did not take long before psychology adopted RT. In the 1860s Donders began to carry out first experimental studies of RT (summarized in Donders, 1969). For example, he compared simple RT tasks (SRT, which required a response to only one stimulus) with choice RT tasks (CRT, which required a different response to each of a number of different stimuli). From differences in these Donders attempted to estimate the speed of mental processes. By the late 19th century Sir Francis Galton had included RT among the various measurements that he collected in his anthropometric lab, which he set up with the aim of determining the extent of individual differences in a range of physical and behavioural variables. Galton only collected a single (and rather crude) RT measurement from each of his participants and, due to insufficient reliability, his results were "disappointing" (Jensen 1982, p. 96). However, with the discovery of *mental chronometry*, the measurement of the speed of mental processes, popularity of RT grew rapidly and it is used by many to this day. This almost two-century long research effort has been summarised in a monograph entitled "Clocking the mind: Mental chronometry and individual differences" (Jensen, 2006).

Most modern psychological studies using RT attempt to estimate the speed of individual's responses (or mean level of RT performance). The central premise of this approach is that mean RT reflects an individual's *true* level of performance, whereas any variation in RTs at different trials are random error or "noise". However, performance of any one person is likely to vary over time, both due to systematic influences such as learning effects, level of arousal, or time of day, and due to unsystematic fluctuations. This variability appears to be more than just error—it is

> a robust phenomenon in which there are reliable individual differences that are manifested consistently across quite different RT tasks (Jensen, 1992, p. 869).

One individual may be inherently more variable across a number of RT measurements than another—even if their average speed is the same (see Figure 1.1. for a schematic illustration of this). Such within-person variability, both in RT and other variables, is beginning to attract increasing recognition from researchers.

Types of variability

Intra-individual (or within-subjects) variability (IIV) refers to the variability observed between different measurements taken from a single individual. Within this very general definition different types of IIV may be discerned.

A useful framework for defining different types of variability more generally is the data box, initially proposed by Cattell (1966). The data box contains three dimensions, which are potential sources of variability: people, measurement occasions, and tests (see Figure 1.2. for an illustration). The definition of variability depends on the dimension of interest (or the source of variability). Three broad types can be identified: *between-people* variability (where a number of people each are tested on one task on a single occasion); *between-tasks* variability (where one person is tested on a number of tasks on a single occasion); and *between-occasions* variability (where one person is tested on one task on a number of occasions).

When considering the central tendency in RT and performance in general, it has been frequent practice to consider the spread of the obtained scores, for example by calculating standard deviation around the mean. This standard deviation denotes the variability between individuals and is sometimes termed *inter-individual variability* or *diversity* (Hale, Myerson, Smith, & Poon, 1988). It is the first of the variability types described above. The remaining two types are *intra-individual*; that is, they concern variability observed within a single person. Between-tasks variability assesses relative performance of an individual on a number of tasks and is often termed *dispersion* (Christensen et al., 1999). Finally, there is variability in performance of a single person on a single task across occasions. This variability is sometimes referred to as *inconsistency* (Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000). Usually, the term *intra-individual variability* is used to describe occasion-to-occasion variability within a single person. It is the latter type of variability that forms the main focus of the present thesis, and both terms (inconsistency and intra-individual variability or, simply, variability) will be used interchangeably.

Sub-types of intra-individual variability

Some types of tasks inherently involve multiple testing "occasions". For example, RT tests usually include a number of trials and the outcome measures (e.g., mean, median, SD) are calculated across these trials. For such tasks (e.g. RT), IIV may be further divided into occasion-to-occasion and trial-to-trial variability. The former refers to performance on the same tasks repeated on different occasions (for example, mean RT on different days, or at different times within a day), whereas the latter refers to variability across multiple trials within a task (for example, variability across all RT trials within a single RT test). These two types of IIV are usually positively correlated (e.g. Rabbitt, Osman, Moore, & Stollery, 2001).

IIV may also be categorised into intra-individual variability and intra-individual change, depending on whether the variability is lasting or transient (Nesselroade, 1991). Nesselroade defined the intra-individual change as

> more or less enduring changes that are construed as developmental by virtue of the nature of their antecedents, their consequences, and their correlates (p.215).

In contrast, intra-individual variability refers to

> relatively short-term changes that are construed as more of less reversible and that occur more rapidly than the intraindividual changes (p.215).

Nesselroade thus distinguishes between variability as a consequence of true intra-individual change over time (such as development or learning) and variability representing transient changes, or fluctuations, in one's level of performance. These two concepts can be considered to broadly reflect trait and state, respectively. However, as we shall see, what Nesselroade refers to as *state* variability can, in fact, be considered a *trait* of an individual. The main interest of this thesis is the transient variability, rather than intra-individual change.

Another issue to be considered when defining IIV is the source of the influences on it. Fiske and Rice (1955) defined *pure* IIV as

> the difference between the two responses of an individual at two points in time under the following conditions: (a) the individual is exposed each time to the same stimulus or to objectively indistinguishable stimulus; (b) the total situation in which the responses are made is the same on both occasions (p. 217).

Fiske and Rice themselves doubt that such pure variability can be achieved; hence, the situation is likely to vary from occasion to occasion. However, they make an important point of distinguishing between the inherent variability of an individual and variability resulting from extraneous factors. For example, a comparison of IIV measured before and after sleep deprivation (e.g. Frey, Badia, & Wright Jr., 2004) represents the effect of external influences. By contrast, comparing IIV of people of different ages represents the inherent variability of an individual. Both these types of variability will be considered in this thesis.

There is stability in variability

Although for a long time IIV was considered to be random error or "noise", it is now thought to be a relatively stable characteristic of an individual that can be reliably measured. For example, Sliwinski et al. (2006) calculated RT IIV indices from odd- and even-numbered trials in three different RT tasks. For both younger and older participants, the odd-even reliabilities of IIV across the three tasks ranged from .65 to .78. Further, Hultsch et al. (2000) found stable individual differences in RT IIV. Specifically, they found that individuals demonstrating more variability on one task were also more variable on other tasks. Moreover, those who showed more variability within a testing occasion also showed more variability across occasions. This extends beyond RT tasks. For example, Nesselroade and Salthouse (2004) estimated within-person variability from two tasks: one requiring participants to track a moving target and one required them to connect subsequent letters, digits or alternating letter-digit sequences. People who were variable within a testing occasion also demonstrated more variability across three separate occasions. It therefore seems that although most people experience fluctuations in performance over short periods of time, the amount of the fluctuation they experience is specific to each person. Nesselroade and Salthouse (2004) state that the finding that

> people who tend to vary more within sessions also tend to vary more across sessions does suggest that there may be a 'characterological' aspect to perceptual-motor intraindividual variability that is consistent with a 'trait-like' interpretation of the phenomenon (p. P53).

Even if RT IIV is a stable characteristic of an individual and can be reliably measured, it may not be of practical importance if its magnitude is small. Investigators who compared intra- with inter-individual variability usually find that the former is about half of the latter (e.g. Li, Aggen, Nesselroade, & Baltes, 2001; Nesselroade & Salthouse, 2004), and Salthouse and Berish (2005) found these two types of variability to be comparable in magnitude. Therefore, it appears that the amount of IIV is substantial. As a stable and reliable measure of considerable magnitude, IIV can be used to describe people's performance and, in itself, can be the subject of investigation.

The importance of RT IIV

Various theorists have proposed that RT IIV is a behavioural marker of the efficiency and integrity of the central nervous system. For example, Hendrickson (1982) suggested that IIV may be due to random errors or "noise" in neural transmission. Li and Lindenberger (1999) further suggested that the efficiency of neural transmission may be affected by catecholamines involved in neuromodulation (specifically, dopamine, norepinephrine, and epinephrine).

The supposition that variability in RT may reflect the underlying state and functioning of the central nervous system is indirectly supported by evidence from a number of different lines of investigation. For example, variability in RT is related to the level of cognitive functioning—individuals with greater variability tend to have lower general intelligence (Deary, Der & Ford, 2001; Jensen, 1992; Rabbitt, Osman et al., 2001) and poorer performance on a number of specific cognitive tasks (Bunce, Tzur, Rarachurn, Gain, & Bond, 2008; Hultsch, MacDonald, & Dixon, 2002). RT IIV also increases with age in normal ageing (Der & Deary, 2006; Hultsch et al., 2002). Moreover, increased RT IIV is apparent in conditions or situations that affect the functioning of the central nervous system, for example in neurodegenerative diseases (Burton, Strauss, Hultsch, Moll, & Hunter, 2006), following brain injury (Stuss, Murphy, Binns, & Alexander, 2003), presence in a hypoxic environment such as high altitude (McFarland, 1937a,b), or alcohol consumption (Maruff, Falleti, Collie, Darby, & McStephen, 2005). Elevated RT IIV is also associated with compromised frontal lobe white matter integrity (e.g. Bunce et al., 2007).

The importance of the RT IIV measure is also reflected in its predictive value. For example, it has been suggested that an increase in inconsistency of performance may be an indicator of an impending change in cognition (e.g., Salthouse, 2007). Indeed, RT IIV has been found to predict cognitive change over 6 years (MacDonald, Hultsch, Dixon, 2003), and even death (Deary and Der, 2005b; Shipley, Der, Taylor, & Deary, 2006).

Therefore, there are a number of theoretical grounds for the interest in RT variability, and the importance of this measure is emphasized by Jensen (2006) who claims that:

> theoretically, individual differences in RTm [mean RT] can be causally derived from whatever causal phenomenon is measured by individual differences in RTSD [RT IIV], but the reverse is not the case. Therefore, intraindividual variability in RT (…) represents a theoretically more fundamental aspect of individual differences in speed of information processing than does RTm (p. 203).

RT IIV as an index of cognitive functioning

Intelligence

It has long been established that central tendency in RT is related to general intelligence, with more intelligent individuals having faster RTs (e.g. Peak & Boring, 1926; Beck, 1933). This observation has provided a foundation for the processing-speed models of intelligence. However, as noted by Vernon (1981),

> it is of theoretical importance that not only the subjects' *speed* of response but also the consistency with which they perform at a given level is related to their intelligence. This finding supports the idea that the overall efficiency—implying both speed *and* a lack of interference or 'noise' which results in greater consistency—with which the person can process information will have a large effect on their performance on intellectual tasks (p. 346).

Indeed, RT IIV has been found to be greater among individuals with lower levels of mental ability measured using a number of different tests; for example, the Armed Services Vocational Aptitude Battery (ASVAB, Larson & Alderton, 1990), Alice Heim 4 Test (AH4; Deary, Der, & Ford, 2001), or Culture Fair Intelligence Test (Rabbitt, Osman et al., 2001). Larson and Alderton (1990) further investigated the magnitude of the associations between general intelligence and different bands of RTs based on speed. For each individual, they ordered RTs from 80 trials from fastest to slowest and grouped them into 16 bands of 5 RTs. They then correlated the mean RT for each band with the estimate of intelligence derived from the ASVAB. A clear pattern of gradually increasing correlations from the fastest (-.20) to slowest

RTs (-.37) was revealed, suggesting that slower RTs are better predictors of intelligence than faster RTs. The pattern of increasingly strong associations with the overall RT variability was also found; the correlations varied from .33 (with the fastest) to .95 (with the slowest band). Crucially, with the slowest band partialled out, the correlation between RT variability and intelligence was no longer significant, suggesting that "intellectually important differences in variability originate in the magnitude of each subjects' slowest RTs" (Larson & Alderton, 1990, p. 319). However, Larson and Alderton's findings were not replicated in a subsequent investigation using a number of different RT tasks (Salthouse, 1996). In that study, stable relationships between mean RT and cognitive performance across the RT distribution were found, suggesting that individual differences are not limited to the slowest portion of the RT distribution.

In a series of investigations using various RT tasks, Jensen (1987a, 1987b, 1991, 1992) consistently found that the (negative) correlation between RT IIV and intelligence is stronger than the correlation between the central tendency in RT (i.e. median or mean) and intelligence. However, this finding, based largely on samples of students, has not been confirmed in a more recent study on a population-representative sample of 900 people aged about 56 (Deary et al., 2001). In that study, significant correlations were found between general mental ability (AH4) scores and RT mean and variability measures from both SRT and CRT tasks, but the correlation coefficients were smaller for measures of variability (-.26 with both SRT IIV and CRT IIV) than for measures of central tendency (-.31 with SRT mean; -.49 with CRT mean). Leaving aside the issue of whether correlation with intelligence in stronger for central tendency or variability, given the available evidence it is well established that both RT measures correlate significantly and negatively with general mental ability scores.

The finding of greater RT variability among individuals with lower intelligence applies not only to the *normal* range, but also to comparisons using the mentally retarded. A number of studies found that mentally retarded individuals (with average IQ ranging from about 50 to 65) have greater RT IIV than non-retarded subjects (e.g. Gardepe & Runcie, 1983; Kellas, 1969; Kirby, Nettlebeck, & Tiggemann, 1977). However, it should be noted that most these studies compared the

retarded groups with groups of students, who are likely to represent above-average intelligence. The correlation between IQ and RT measures (mean, IIV) has also been found within mentally retarded groups (e.g. Vernon, 1981). However, Kirby et al. found significant correlations between IQ and RT variables in only one of four RT tasks they used. It appears that RT variability is related to intelligence across the range, including its normal and abnormal portions.

Specific cognitive abilities

As well as being related to general intelligence, RT IIV is associated with various specific cognitive abilities, with the same general direction: greater variability is linked with poorer performance. For example, Bunce, Tzur et al. (2008) studied correlations between SRT IIV, CRT IIV and three tasks designed to measure components of executive functioning: switching between mental sets, working memory (monitoring and updating), and inhibition. Switching was assessed using a version of verbal fluency tasks requiring an alternate generation of names of animals and countries; updating was measured by a version of n-back task, which involved monitoring of a string of individually presented digits and detecting sequences of three odd numbers; and inhibition was assessed with a Stroop task. Individuals displaying more variability in their SRT and CRT scored lower on switching and inhibition; greater CRT IIV (but not SRT IIV) was also associated with poorer performance on updating, showing that the RT IIV is related to performance on tests of executive functioning. Larson and Alderton (1990) also found that CRT IIV negatively correlated with performance on a task requiring monitoring, retention and updating of the information.

Hultsch et al. (2000) derived RT IIV from four different RT tasks: SRT, CRT, lexical decision RT and semantic decision RT, and found that the variability in each was positively associated with mean RT and negatively associated with accuracy. In other words, more variable individuals had slower RTs and made more errors. This effect was not only observed within tasks but extended, for example, to accuracy in digit symbol substitution task or block design.

In a subsequent study, Hultsch et al. (2002) considered four different clusters of cognitive abilities: perceptual speed, working memory, episodic memory, and crystallised abilities. Perceptual speed was measured using three timed tasks requiring participants to choose two identical pictures out of five line drawings, to decide whether two strings of digits were identical, and to substitute a series of digits with pre-assigned symbols (a version of digit symbol substitution). Working memory tasks involved solving arithmetic problems while remembering the numbers for later recall, and answering simple questions about aurally presented sentences while memorising the last word of each sentence. Episodic memory was assessed using word and story recall tasks. Finally, crystallised abilities were measured with two tests measuring world and vocabulary knowledge, respectively. Hultsch et al. assessed correlations between each of the four composite cognitive abilities scores and four RT measures (SRT, CRT, lexical decision RT and semantic decision RT) in four age groups: young (17-36 years), young-old (54-64 years), mid-old (65-74 years) and old-old (75-94 years). Trial-to-trial variability in the two more complex RT tasks (decision and semantic) were negatively correlated with almost all cognitive scores across the four age groups. The correlations of SRT IIV and CRT IIV with the cognitive composites were most marked in the oldest group. In fact, in the three younger age groups, SRT IIV was only related to perceptual speed and working memory in the young-old and to perceptual speed in the mid-old group; CRT IIV was only significantly correlated with perceptual speed in the young and the mid-old groups.

RT IIV has also been demonstrated to be related to the ability to solve everyday problems from various domains from shopping and telephone use to financial and household management (Burton, Strauss, Hunter, & Hultsch, 2009). Specific problems given to the participants to solve included for example calculating a taxi bill, or working out a number of calories in a serving of cereal. SRT IIV and CRT IIV correlated -.40 and -.34, respectively, with the performance on the everyday problem solving tests. Therefore, RT IIV is associated with a level of performance at a number of cognitive abilities, with less variable individuals demonstrating better performance across tasks.

RT IIV as a marker of neurological functioning

Healthy ageing

Performance on various cognitive tasks declines with age during healthy ageing, and RT IIV is no exception. Older adults have more variable RTs than young adults (Hultsch et al., 2002; Der & Deary, 2006) and the effect is observed even at the older extreme of the age range, with the oldest old having greater RT IIV than the younger old (Dixon et al., 2007). However, Rabbitt, Osman et al. (2001) did not find significant correlations between IIV and age in participants aged 52 to 79 years. Because in their study the three age groups (50 to 61 years, 61 to 70 years, and 71 to 79 years) were matched in terms of their Culture Fair intelligence test scores, Rabbitt and his colleagues suggested that age has no independent contribution to individual differences in RT IIV over and above the Culture Fair scores. This is an interesting idea, suggesting that age effects observed may be due to a general age-related cognitive decline. However, the null finding could also be due to a relatively small sample size (n = 80).

Deary and Der (2005a) investigated both cross-sectional and longitudinal effects of age on SRT IIV and CRT IIV. Cross-sectional differences between young, middle-aged and old cohorts were as expected, with the oldest group demonstrating most variability. Longitudinal increase in SRT IIV was observed over the 8-year follow-up across the three narrow age cohorts (aged about 15, 35 and 55, respectively, at the first wave of testing). However, the pattern was different at different ages for CRT IIV: it decreased in the youngest cohort and increased in the oldest cohort. Sex differences were also noted, with females on average more variable, but especially so in the older groups. Given little evidence of sex differences in the youngest group, aged about 15 years at the first wave of testing, Deary and Der suggested that the post-pubertal accumulation of sex hormones may lead to differentiation of responses between males and females. So far, this hypothesis has not been formally tested. However, Reimers and Maylor (2006) proposed an alternative explanation for the apparent sex difference in RT IIV. They

suggested that it is created largely by different performance of males and females on the initial trials in a RT test, which they demonstrated using a gender classification task. The apparatus used to collect data analysed by Deary and Der did not record individual trial data, and so, Reimers and Maylor's idea could not be tested on the CRT data from that dataset. Therefore, the hypothesis that males and females differ in their CRT across trials remains to be tested.

Although Deary and Der (2005a) found evidence for increase in RT variability over time, other longitudinal investigations have not provided compelling evidence for longitudinal changes in RT IIV with age. For example, Maddox and Douglass (1974) failed to find any evidence for change in CRT IIV over the period of almost 10 years; Fozard, Vercruyssen, Reynolds, Hancock, and Quilter (1994) found longitudinal increases in a SRT task variability, but not in a more complex go/ no-go task. Even a decrease in CRT IIV over a 6-year period, in a group of adults aged 55 to 89 years, has been reported (MacDonald et al., 2003). Therefore, whereas cross-sectional data point to an increase in RT IIV with age, longitudinal evidence is less clear.

As well as different designs of study, using different measures of RT IIV may affect the findings regarding age effects. RT IIV and RT mean correlate positively with each other and some researchers suggest that average speed (i.e. mean RT) should be controlled when analysing RT IIV data (see sections "methodological issues in the study of IIV" and "different measures of intra-individual variability" below for a fuller discussion of this issue). Shammi, Bosman, and Stuss (1998) performed two analyses—with and without controlling for mean RT—and found that age differences were only significant when differences in average speed were not controlled. However, a number of researchers have reported significant age effects on IIV even after controlling for individual mean RT (e.g. Hultsch et al., 2002; Williams, Hultsch, Strauss, Hunter, & Tannock, 2005; Williams, Strauss, Hultsch, & Hunter, 2007). It is not yet clear how different approaches to the study of RT IIV—that is, controlling versus not controlling for mean RT as well as using different methods of controlling for RT mean—affect the subsequent findings in terms of age effect.

The apparent changes in variability with age are not limited to old age. There is evidence that marked change (decrease) in variability occurs in childhood and adolescence and that, across the lifespan, RT IIV is characterised by a U-shape function (Williams et al., 2005, 2007). These changes in RT IIV across the lifespan—a rapid decrease in early years and adolescence, gradually moderating in early adulthood, followed by a decrease in older age—have a parallel with lifespan changes in brain white matter and grey matter volume (MacDonald, Nyberg, & Backman, 2006). Therefore, some have suggested that changes in white and grey matter volume may be the biological bases of changes in RT IIV.

White matter volume increases in childhood, adolescence and young adulthood and decreases in middle and older age (Sowell, et al., 2004), which maps onto the decrease and subsequent increase in RT variability through the lifespan. In addition, a decrease in grey matter volume observed in adolescence (e.g. Gogaty, et al., 2004), coincides with the decrease in RT IIV observed at those ages (e.g. Williams et al. 2005, 2007). Therefore, the decrease in RT IIV observed in children and adolescents and increase in RT IIV observed in old age, might reflect changes in neural transmission efficiency following morphological changes occurring in the brain during maturation and senescence, respectively. In particular, synaptic pruning which involves elimination of some neural connections and myelination could increase the neural transmission efficiency and lead to a reduction in RT variability. Therefore, the findings regarding changes in RT IIV across the lifespan supports the suggestion that variability is a marker of the central nervous system functioning.

Brain damage

Some of the most direct evidence for the link between the central nervous system functioning and RT IIV comes from the study of patients who suffered brain damage. Such patients often demonstrate greater IIV to the extent that Head (1926) stated:

> an inconsistent response is one of the most striking results produced by a lesion of the cerebral cortex (p.145).

Stuss et al. (2003) argued that there may be two broad types of variability: one associated with a particular cognitive process (e.g. on lexical tasks in patients with aphasia, or on bisection tasks in patients with spatial neglect) and a more general one, reflecting attentional deficits that are not specific to any particular cognitive domain. IIV in relatively simple RT tasks would denote the latter type of variability and it is commonly found to be increased among individuals with brain lesions of various origin, severity and location (Bruhn & Parsons, 1971; Stuss et al., 1989).

For example, Bruhn and Parsons (1971) compared two groups of patients: individuals with cortical damage and controls with no brain damage (mainly patients from the orthopaedic ward), who were matched in terms of age and education. Brain-damaged patients had greater RT IIV, measured as both the individual standard deviation and as the difference between the means of the 10 slowest and 10 fastest responses. Since there was no evidence that patients with brain damage demonstrated a larger number of very long responses (blocks or lapses) than controls, the authors suggested that structural damage to the nervous tissue results in general slowing of RT. The implication of this is that increased RT IIV following brain injury may *result from* general slowing rather than being the primary phenomenon. However, other researchers often find that brain-injured patients demonstrate greater variability even when effects of the mean RT are controlled, suggesting that not all of the IIV is a product of general slowing (Hetherington, Stuss, & Finlayson, 1996; Stuss et al., 2003).

Greater variability seems to exist even in heterogeneous groups of patients who have injuries in various locations of the brain. These injuries may have varying extent and result from different causes. However, the location and severity of a brain injury are likely to be related to the magnitude of RT IIV increase. Stuss et al. (2003) investigated the importance of the injury location for RT IIV in groups of patients with different patterns of traumatic brain injury (TBI), in particular frontal vs. non-frontal. Specifically, they compared the performance of patients with non-frontal lesions and non-TBI controls with four groups of patients with focal frontal lesions located primarily in: left dorsolateral frontal, right dorsolateral frontal, inferior medial and superior medial (with the latter group often having inferior medial lesions). They found that patients with lesions in the inferior medial region

performed as well as controls with no TBI, whereas other groups with lesions in the frontal lobes showed increased RT IIV relative to the control group. Those with lesions in non-frontal regions demonstrated greater variability in RT on a complex task, requiring concurrent processing of a number of different features of the stimuli (colour, shape and fill pattern). Therefore, Stuss et al. (2003) suggested that frontal lobes may be of particular importance when it comes to RT IIV.

Another aspect of brain injury that could affect the degree of RT IIV impairment is its severity. Empirical evidence regarding severity of brain injury is not very clear, however. For example, severity as quantified by the length of a post-injury coma, does not appear to be related to the level of RT IIV (Hetherington et al. 1996). Moreover, Stuss et al. (1989) reported on three comparisons performed between different head injury patient groups and healthy controls and only found significant correlations with severity (as assessed by Glasgow Coma Scale and by the duration of post-traumatic amnesia) in one of the three comparisons. They suggested that differences in time since injury in one group and low general severity of injuries in the second group could have reduced the correlations.

A variable related to the severity of injury is the time since injury. It is to be expected that, with time and rehabilitation, the severity of the injury may decrease. Stuss et al. (1989) found a negative correlation between time since injury and RT IIV. Hetherington et al. (1996) explored this finding further and discovered that recovery from TBI in terms of RT IIV may still be taking place several years after the injury. They compared two groups of patients with traumatic brain injury—one 5 years post-injury and another one 10-years post-injury—and found that RT IIV was greater in the 5-year group than in either non-TBI controls or the group 10 years post injury.

Further studies on recovery, include Collins and Long (1996), who investigated mean and IIV on SRT and a go/no-go RT task in healthy adults and two groups of patients with TBI: cognitively impaired or non-impaired. The impairment status was assessed using a revised form of the Impairment Index of the Halstead-Reitan Neuropsychological Test Battery. In line with other empirical findings, Collins and Long found that the TBI patients, whether or not they were considered to be impaired according to the neuropsychological battery, had slower and more

variable RTs on both tasks than did healthy controls. In other words, patients who would be considered to have fully recovered from their injury, still showed abnormal performance on RT tasks. Of particular note is that mean level of performance on RT tasks was better able to discriminate between impaired TBI patients and normal controls, whereas RT IIV better discriminated between non-impaired TBI patients and healthy controls. These results, suggest that RT IIV may be a sensitive measure of neurological dysfunction that may be missed with by a standard neuropsychological assessment.

Some researchers have speculated that IIV differences between patients with brain damage and healthy controls should be greater in more complex tasks (e.g. Stuss et al., 2003). However, empirical evidence regarding this suggestion is rather mixed. For example Bruhn and Parsons (1971) found differences between their brain damage groups and controls on SRT and a go/no-go task to be of similar magnitude. Others found that the differences more pronounced for more complex tasks (e.g. Stuss et al., 1994).

Dementia

Increased RT IIV is often reported in dementia (Walker, Ayre, Cummings et al., 2000; Walker, Ayre, Perry et al., 2000). When different types of dementias are considered, the usual finding is that variability of demented patients is greater than that of normal controls, but there are some differences between different dementia groups as well. For example, Walker and colleagues compared RT IIV in patients with Alzheimer's Disease (AD), patients with Dementia with Lewy Bodies (DLB) and healthy elderly controls. They found that, whereas both dementia groups were more variable than healthy controls, DLB patients demonstrated greater SRT and CRT IIV than AD patients. Moreover, the authors demonstrated that RT IIV could differentiate between the two dementia types, so much so, that they suggested it might be useful to include RT IIV in the standard clinical criteria for DLB. When a further category was added, namely vascular dementia (VaD), RT IIV was consistently greater in DLB that either VaD or AD, however, not all comparisons

were statistically significant (Walker, Ayre, Cummings et al., 2000). For example, the DLB group was more variable than AD group on CRT but not vigilance RT, whereas the inverse pattern was found for the DLB–VaD comparison. When the VaD group was compared with AD, it had greater CRT IIV, whereas there was no significant difference in IIV from a vigilance RT. However, the means pointed towards greater vigilance IIV in the AD group, indicating that IIV in different RT tasks needs to be further investigated. All three investigations by Walker and colleagues show that RT IIV may be positively correlated with clinically assessed fluctuating consciousness or fluctuating cognition (FC). They argue that this is a common feature in many dementias (Walker et al., 1999; Walker, Ayre, Cummings et al., 2000; Walker, Ayre, Perry et al., 2000). FC refers to marked fluctuations in attention and alertness; an example of an extreme case of FC is a patient who is able to have a normal conversation at one point in time and becomes confused, non-conversional and unable to stand unaided at another (Byrne, Lennox, Lowe, & Godwin-Austen, 1989).

Murtha, Cismaru, Waechter, and Chertkow (2002) tested the previously-implied importance of the frontal lobes for RT IIV by comparing IIV of patients with AD and with frontal lobe dementia (FLD). Contrary to expectation, FLD patients did not demonstrate greater RT IIV than AD patients for any of the three RT tests used (SRT, cued CRT and CRT). In fact, the IIV was consistently *lower* for FLD patients than for AD patients, although not significantly so. The only significant difference was between AD patients and healthy controls in CRT IIV, with variability greater among patients than controls. This study was hindered by relatively small numbers of patients, with only 5 FLD and 8 AD cases. Therefore, the lack of significant findings may simply be due to insufficient power. However, the consistency of the pattern of lower variability of the FLD patients relative to the AD patients warrants further investigation of the role of the frontal lobes in the RT IIV. Murtha et al. found some evidence of increased between-subject variability in RT and week-to-week variability in group mean RT in the FLD group. This suggests that some aspects of variability are affected by the frontal lobe dysfunction. Again, future investigations are necessary to shed more light on the issue.

Although it remains to be resolved whether different dementia groups differ in the level of RT IIV, greater RT IIV of dementia patients as a whole in relation to non-demented controls is well-established. However, evidence from comparisons between groups of patients and healthy individuals may not be compelling, because it is possible that the differences observed between the groups are not due the dementia status per se but rather due to presence or absence of a disease more generally. For example, Hultsch et al. (2000) suggested that increased RT IIV may be caused by changes in affective or somatic states, brought about by illnesses which need not be neurological. They addressed this issue by comparing patients with mild dementia (with possible/probable AD or VaD diagnosis) but no other health complaints, patients with osteoarthritis but no neurological conditions, and healthy controls with no history of any major disease. They found, using four different tasks, that RT IIV was larger in the dementia group than either the arthritis or control group and that the latter two groups did not significantly differ from each other. Therefore, it appears that greater RT IIV is specifically associated with neurological disturbances and that altered somatic stated do not lead to a change in RT IIV.

Mild cognitive impairment

As well as in clinically diagnosed dementia, RT IIV seems to be increased in individuals whose cognitive performance is poorer than expected for their age but not sufficient to merit a diagnosis of dementia – termed *mild cognitive impairment* (MCI; e.g. Petersen, 2004). However, partly because of a large heterogeneity in the definition of the MCI among researchers, the findings are somewhat equivocal.

Dixon et al. (2007) investigated RT IIV on a number of tasks, including SRT and CRT, in two samples of older adults aged 64 years and above. In a study with 416 participants from the Victoria Longitudinal Study they found that SRT IIV was higher in MCI than in non-impaired individuals but there was no significant effect for CRT IIV. Similarly, SRT IIV but not CRT IIV predicted cognitive status (MCI vs. no MCI) over and above speed of response (mean RT). A second sample included 304 participants from project MIND, classified as non-impaired, mild MCI or

moderate MCI. SRT IIV and CRT IIV both were higher in moderate MCI than either mild MCI or non-impaired groups. In addition, CRT was significantly higher in the mild MCI group than in the group with no cognitive impairment. Both SRT IIV and CRT IIV predicted cognitive status between no impairment and moderate MCI and between mild and moderate MCI over and above RT speed. Taken together, these two investigations show that both SRT IIV and CRT IIV tend to be greater among individuals with MCI.

Similarly, Christensen et al. (2005) found that in a group of 60 to 64-year-olds SRT IIV and CRT IIV were greater among individuals with some form of cognitive impairment. However, when they investigated differences between subgroups based on the specific impairment (ageing-associated cognitive decline, age-associated memory impairment and MCI) the IIV was no longer significantly greater among individuals with MCI relative to unimpaired individuals. In fact, only those with ageing-associated cognitive decline showed increased SRT IIV and CRT IIV (with small effect sizes of .18 and .25, respectively). Moreover, in contrast to Dixon et al.'s findings, RT IIV did not make a contribution to the diagnostic status (impaired vs. non-impaired) over that made by mean RT.

In another study, which also considered different groups of MCI, but on an older sample (64-92 years), RT IIV was again a more successful measure than mean RT (Strauss, Bielak, Bunce, Hunter, & Hultsch, 2007). In this study, four MCI groups were created based on the presence or absence of memory impairment and the extent of impairment (single or multiple cognitive domains affected): memory single, memory multiple, non-memory single and non-memory multiple. A wide variety of RT tasks were administered, from a tapping task in which "RT" was purely motor in nature (with no cognitive component), through standard CRT-type tasks to more complex tasks such as 1-back (involving working memory) or a task involving inhibition of irrelevant information. Strauss et al. used principal components analysis to derive four factors from these tasks – motor, basic RT, complex RT and interference RT. Latencies from the tapping task loaded on the motor factor, CRT tasks loaded on basic RT factor, 1-back and switching task loaded on complex RT factor, and a task involving interference control loaded on the interference RT factor. The only group significantly more variable on the motor component than non-

impaired individuals was memory multiple group. All four MCI groups had greater variability from CRT tasks (basic RT component) than those who were cognitively intact. For the more complex tasks (complex and interference components) both groups with impairments in multiple cognitive domains performed poorer than non-impaired controls. When the groups were collapsed into two broader categories based on severity of impairment (intact/single domain impairment vs. multiple domains), variability scores predicted group membership over and above mean RT.

Parkinson's Disease

Elevated RT IIV can also be observed among patients with Parkinson's Disease (PD). For example, Crawford, Goodwich, Henderson, and Kennard (1989) demonstrated that patients with PD had more variable SRTs than controls. Reeds and Franks (1998) compared two groups of PD patients varying in the severity to which they were affected. Those more affected showed increased variability relative to the less severely affected. A subsequent study by de Frias, Dixon, Fisher and Camicioli (2007) found that whereas PD patients had slower RTs on three different CRT tasks, they were significantly more variable than healthy controls only on the most complex task, including 8 possible choices. No group differences were found on SRT, suggesting that task complexity may be an important factor. In comparing groups of varying severity, de Frias et al. found that more severely affected patients performed worse on two out of three CRT tasks (with 2 and 8 possible choices respectively).

Burton et al. (2006) compared PD patients with AD patients and healthy controls and found that AD patients had more variable SRT and CRT than PD patients and that PD patients had more variable RTs than healthy controls. This finding could be taken to mean that RT IIV reflects specific neurological dysfunction rather than a more general compromising of the central nervous system. However, it is of note that in Burton et al.'s sample, AD patients had lowest current levels of cognitive performance than either PD patients or healthy controls as well as lower estimated pre-morbid IQ. Therefore, some differences in RT IIV performance between the groups might be expected regardless of their disease status. Moreover,

PD patients were medicated, with all receiving dopamine replacement therapy. Considering that dopaminergic systems have been implicated in RT IIV (Li & Lindenberger, 1999), this might have a direct positive effect on their RT IIV.

Bringing it together

Taken together, the available evidence suggests that RT IIV is elevated in a number of conditions characterised by neurological dysfunction. There is some evidence to suggest that the amount of variability is related to the severity of the impairment. Whereas some suggestions have been made that group differences in RT IIV are greater in more complex tasks, the empirical evidence is equivocal. The importance of frontal lobes for intact RT IIV has been studied in brain damage and dementias, and there is some evidence to support this proposition, but again the findings are not completely uniform.

Predictive value of RT IIV

As outlined earlier, RT IIV is related to the level of cognitive performance on a number of different tasks. It can also predict neurological status. For example, Bruhn and Parsons (1977) found that RT IIV differentiated between healthy and impaired groups, although not between different types of impairments (brain damage vs. epilepsy). Hultsch et al. (2000) found that SRT IIV and CRT IIV made a unique contribution, over and above that of mean RT, to predicting neurological status (dementia vs. non-dementia group membership). However, RT IIV is not limited to contemporaneous associations – it can also predict future cognition, cognitive decline and even mortality.

For example, RT IIV has been found to predict cognitive decline over several years (Lövdén, Li, Shing, & Lindenberger, 2007; MacDonald et al., 2003). In a recent study, Cherbuin, Sachdev, and Anstey (2011) showed that IIV could predict progression from healthy ageing to MCI. RT IIV is not only predictive of cognition.

For example, among patients undergoing knee or hip replacement operations, pre-operative levels of CRT IIV predicted the likelihood of developing delirium after the surgery (Lowery, Wesnes, & Ballard, 2007). Specifically, patients with CRT IIV 1 standard deviation above the sample mean were five times more likely to develop delirium. However, this effect was not significant for SRT IIV and only a trend was found for RT IIV from a vigilance task (p = .051).

Ultimately, RT IIV predicts mortality. This has been first demonstrated by Deary and Der (2005b) in the oldest cohort of the West of Scotland Twenty-07 study. SRT and CRT data were collected from over 1,000 people aged about 56 years in 1988. They were followed up for mortality until the end of the year 2002. The study found that individuals who had more variable SRT and CRT at the first wave of testing were less likely to still be alive about 15 years later. One standard deviation increase in SRT ISD was associated with a 29% increase in the likelihood of being dead at follow-up, compared with a 17% increase in the risk of dying per standard deviation of CRT ISD. SRT IIV and CRT IIV predicted mortality even after adjusting for sex, smoking, social class and education. RT measures (specifically, CRT mean and SRT IIV) predicted mortality better than did general intelligence assessed by AH4 Test.

Subsequent investigations by Shipley, Deary and Der confirmed the association between RT IIV and death. For example, Shipley et al. (2006) investigated data from the UK Health and Lifestyle Survey (HALS) and found that variability on SRT and CRT tasks were significant predictors of mortality; people with less variable RTs were more likely to be alive 19 years later. The effects were slightly attenuated by adjusting for socio-economic status (social class and education) and health behaviours (smoking, alcohol consumption and physical activity); controlling for health status (forced expiratory volume, body mass index and blood pressure) attenuated the effects of RT IIV so that they were no longer significant. When the sample was split into three age groups: young (aged 20 to 39 years), middle-aged (40 to 59 years) and old (aged 60 and above), both SRT and CRT variability predicted mortality in the oldest group (in which there were most deaths). CRT variability was also a significant predictor of mortality in the youngest group, suggesting that RT's relationship with death is not limited to the older

individuals. Non-significant associations in the middle-aged group could be attributed to the smaller number of deaths in that group. In a further study on the same sample, Shipley et al. (2007) investigated the association between change in RT IIV over 7 years and mortality. A greater increase in SRT and CRT variability was associated with all-cause mortality as well as with mortality from specific causes (all cardio-vascular, coronary heart disease, stroke, respiratory disease). These associations remained significant even when controlling for SES, health behaviours, health status, as well as age, sex and the baseline cognitive ability. The results suggest that the association of RT IIV with death is robust and independent.

The association between RT IIV on semantic and lexical decision tasks and death has also been investigated and confirmed by MacDonald, Hultsch, and Dixon (2008). They found, using two samples from the Victoria Longitudinal Study, that RT IIV was greater among those who died before May 2002 than among survivors. The effect size was larger in the older (80-95) than younger (59-79) individuals. They also found that RT IIV was a better predictor of mortality than mean RT.

Methodological issues in the study of IIV

RT IIV seems to be a characteristic of an individual (e.g. Hultsch et al., 2002; Rabbitt, Osman et al., 2001; Nesselroade & Salthouse, 2004). In addition, its importance has many manifestations in the literature, as reviewed above. However, there are several methodological issues that make the study of RT IIV problematic.

Correlation with the RT mean

Among the major problems facing researchers into the IIV is its correlation with RT mean. Specifically, individuals who exhibit large IIV usually have high mean RT as well. The correlation between RT mean and IIV is usually medium to large in magnitude. For example, Deary and Der (2005a) found the correlation between mean and variability to be .58 for SRT and .55 for CRT; Hultsch et al. (2000) report values of .54 and .82 for SRT and CRT. It has also been suggested that

the relationship between RT mean and RT IIV may not be linear (Schmiedek, Lovden & Lindenberger, 2009).

Despite the correlation between mean RT and IIV being well established, the causal direction of this relationship is not known. Some authors indicate that they consider slowing to be the primary phenomenon. For example, Flehmig, Steinborn, Lagner, Scholz, and Westhoff (2007) argue that

> differences in SD_{RT} [RT IIV] between younger and older adults might simply reflect the fact that older adults are on average slower than younger adults (p.136).

However, Jensen (1982) points out that

> assuming an inherent periodicity in the nervous system, the average speed of RT can be seen as a *consequence* [emphasis added] of variability of RT more easily than the reverse relationship (p. 103).

Regardless of whether changes in RT mean lead to an increase in variability or whether greater variability results in greater mean RT or, indeed, whether both RT IIV and mean RT are correlated though independent parameters (Jensen, 1992), RT IIV can only be considered of practical importance if its effects are independent of average speed. Therefore a number of researchers use a measure of IIV which controls for mean RT. These are described and discussed in the section "different measures of intra-individual variability" below.

For example, Salthouse and Berish (2005) showed (for occasion-to-occasion variability) that controlling for mean RT removed or reversed the correlations of RT IIV with age and scores on a number of cognitive tests. However, when ISD was controlled, correlations between mean RT and age and cognitive performance were not substantially altered. Therefore, it appeared that occasion-to-occasion IIV is secondary to mean RT. However, trial-to-trial RT IIV has been shown to be significantly associated with the level of performance on various cognitive tests even when mean RT was controlled. Therefore, it may be that trial-to-trial is a primary phenomenon whereas occasion-to-occasion variability is not. This possibility merits further investigation and controlling for mean RT when studying RT IIV is advisable.

Skewed distribution

Another problem is that the distribution of RTs is usually positively skewed. This largely reflects the fact that RTs are left-censored, in that they cannot be negative. In addition, a theoretical minimum is often applied to the RT data, to eliminate responses which are likely to reflect accidental or anticipatory responses (i.e., responses which were initiated before the appearance of a stimulus, but completed after the stimulus presentation). The same is not true of the slow (long) RTs, for which there is no theoretical maximum. The skew affects commonly used measures of central tendency and dispersion (mean and SD). Although the use of mean and SD to summarise RT data, albeit not fully adequate, is now generally accepted, the summaries themselves also have a skewed distribution. This poses problems for analysis as most statistical tests assume a normal distribution of the outcome variable. However, a number of transformations are available to normalise the data prior to analyses, including using the reciprocal or a logarithmic transformation. A more sophisticated method is also available, which chooses the optimal degree of transformation based on the model to be fitted: Box-Cox transformation (Box & Cox, 1964). This particular method is described in more detail in Chapter 3 of this thesis.

Task and procedural factors

Finally, there are problems associated with the exact procedure employed in a RT task. There are a number of factors that affect RT, including modality of the stimulus, sequence of stimuli, the length of the preparatory interval (PI) and whether or not it is variable (Brebner & Welford, 1980). If these factors affect RTs on individual trials then they may also affect RT IIV calculated across trials. For example, RTs are shorter to stimuli which have been repeated (i.e. the same as the stimulus on the previous trial). If two otherwise identical tasks differ in the number of stimuli repetitions, then RT IIV obtained from them may also be different. This can become an issue when the effects of the procedure differ between groups to be studied. For example, Lahtela, Niemi, and Kuusela (1985) found that, when PIs are

variable, RTs of older adults are more affected by short PIs than RTs of younger adults. In other words, older adults, respond disproportionately slower on trials with short PIs. Thus, a difference in RT IIV between older and younger adults might be an artefact of variable and short PIs.

These problems are by no means specific to the study of RT variability. There is also no clear solution to them, other than to consider them when comparing results of different studies. This is done in the systematic review of age differences in RT IIV reported in Chapter 2.

Different measures of intra-individual variability

Raw measures of IIV

Intra-individual standard deviation and intra-individual variance

The most commonly used measure of RT IIV by far is intra-individual standard deviation (ISD). It is the standard deviation around a person's own mean RT, calculated from RTs across all trials of the test (although sometimes trials with errors, or unusually long or short RTs, are excluded prior to the calculation). The standard deviation is commonly used to describe the amount of variability between subjects, so applying this statistic to within-person variability is a natural choice. The resulting ISD measure is both easy to calculate and intuitively understandable. A variation of this measure uses intra-individual variance (IVar) of the RTs across trials instead of their standard deviation. These two measures are obviously closely related.

Although commonly used, these measures are not without their limitations. Firstly, RT distributions tend to be positively skewed, and so ISD is not a perfect measure of IIV in RT. Moreover, both SD and variance calculated across trials are potentially affected by outliers. Many investigators attempt to trim the data prior to the calculation of IIV. For example, it is common to exclude trials with abnormally short or long RTs, which might result from accidental key presses or distractions.

However, it is difficult to separate aberrant RTs from RTs that are very short or very fast, but validly so. Often arbitrary a priori cut-off values are set to facilitate the decision process. However, using the same values across participants with different speeds of response (e.g., different age groups), means that the strictness of the criteria is not consistent for different individuals.

Inter-quartile range and percentile differences

A group of measures that address some of the limitations of ISD and IVar calculate variability excluding a portion of slowest and fastest ends of each person's RT distribution. One example is the inter-quartile range (IQR), which excludes the top and bottom quartile of RTs. IQR is a special case of a family of measures based on percentile differences and occasionally, percentiles other than 25th and 75th are used. For example, Adam, Retey, Khatami, and Landolt (2006) used a difference between the 90th and 10th as an index of IIV.

Percentile difference measures of IIV have the advantage of trimming an equal portion of RTs from all individuals. Moreover, instead of using arbitrary values, cut-off points are determined for each individual based on their own RTs. However, a major shortcoming of all range measures is that they disregard much potentially informative data. For example, in the case of IQR, 50% of the RT data are effectively discarded. In addition, these measures are crude and do not account for what happens in between the two percentiles

Methods of controlling for mean RT

Coefficient of variation

Among the simplest ways of controlling for mean RT is using the coefficient of variation (CV). It is obtained for each individual by dividing their RT ISD by intra-individual mean. The central premise behind the use of CV as a measure of

variability is that if IIV increases in proportion to RT mean, then no difference in CV will be found between individuals with different average speed. Therefore, if differences in CVs are found, then it can be assumed that individual differences in RT IIV are not attributable to individual differences in mean RT.

The simplicity of CV and its ease of computation place it among the most commonly used mean-adjusted measures of intra-individual variability. However, CV is not without its limitations. For example, Hultsch et al. (2008) point out that it confounds unsystematic variability with systematic effects (such as practice, fatigue etc.). For example if trial-to-trial RT performance improves across a block of trials, then this systematic decrease in RT across trials would add to the variability estimated using the CV. Further, they argue that the usefulness of CV is limited because it is a cross-product of the main effect of ISD, the main effect mean RT, and their interaction. Therefore, any group differences that are found using the CV may be due to differences in mean RT, differences in RT ISD, or a mixture of both. Considering CV alone, does not allow these effects to be separated. However, this issue is less applicable if both ISD and CV are used in a single study and the CV is used to verify effects established previously with ISD.

Including mean RT as a covariate

Mean RT may also be controlled by including it as a covariate in the models of RT IIV. Although, as with CV, this method does not control for systematic sources of variation, it allows a more precise modelling of the relationship between mean and ISD. For example, non-linear associations may be modelled and tested by including polynomial terms of mean RT in models of RT IIV.

Regression method and purified residuals

Finally, some authors (e.g. Hultsch et al., 2000; Hultsch, et al., 2002, Williams et al., 2005) opt for a regression method which allows them to remove various systematic effects on IIV, for example, group differences in mean RT, the

effects of practice or fatigue. In this approach, RTs are regressed on the potential confounders, such as age group, gender, or trial number. The resulting residuals are then used to calculate the ISD. To highlight that these residuals are independent of the effects of the confounders, they are sometimes referred to as *purified* residuals.

Although regression method controls for some systematic effects on variability, it does not do so with much precision. Predicted values from regression model controlling age and trial number will be identical for individuals of the same age on the same trial. Their residual scores will be calculated as deviations from that one predicted value, common for the whole group. This method becomes even more problematic if the model is mis-specified, for example, by assuming a linear relationship between RT IIV and age.

Whole distribution analysis

Some researchers attempt to characterise the entire RT distribution. This can be achieved by fitting a convolution of a normal (Gaussian) distribution and an exponential distribution, known as ex-Gaussian function, to the RT data (e.g. McAuley et al., 2006; West et al., 2002). Three parameters are obtained from this procedure: μ and σ, representing, respectively, the mean and standard deviation of the normal component, and τ representing the exponential component. Therefore μ reflects the average performance (speed), σ reflects the variability, and τ captures the exponentially distributed tail of the distribution.

Outstanding questions; unresolved issues

Age differences in RT IIV

Clearly, the vast majority of research evidence points to greater variability of older relative to younger adults. However, as mentioned before, the finding tends to be stronger in cross-sectional rather than longitudinal investigations and when not

controlling rather than when controlling for mean RT. In addition, differences in procedures used in different studies may confound the age differences found. To address these issues a systematic review of the literature on age differences in RT IIV was performed and it is reported in Chapter 2. The review attempted to establish and estimate the magnitude of age differences in RT IIV. Given some suggestions that group differences are more marked in tasks of greater difficulty, two RT tasks were considered: SRT and CRT. The systematic review also considered a number of procedural factors and tested their importance for the study of age differences in RT IIV.

As well as the systematic review, two empirical studies were concerned with age effects on SRT IIV and CRT IIV (Chapters 3 and 4). A body of cross-sectional studies report greater IIV among older than younger adults, but few studies investigated age effect across the lifespan, trying to model actual age effects rather than simply comparing different age groups. The study described in Chapter 3 addressed this gap and investigated age effects across the life span, in a large, cross-sectional sample with a wide age range including children, adolescents and adults up to the age of 75. Given that the evidence regarding age effects on RT variability from longitudinal studies is less clear, in Chapter 4 a longitudinal study of age effects is reported.

Sex differences in RT IIV

The available evidence regarding sex differences in RT IIV suggests that adult females are more variable; there is no clear evidence of sex differences in the younger samples. Deary and Der (2005a) suggested that sex differences in RT variability may not be present in children and occur only in adulthood. To date, there has been no investigation of sex differences in RT IIV across the lifespan. In Chapter 3, sex differences are investigated across the lifespan to test whether the hypothesis that they occur post-puberty is substantiated.

The alternative hypothesis—that males and females respond differently across trials (Reimers & Maylor, 2006)—was also tested and is reported in Chapter

5. In this study trial-by-trial variability was investigated in relation to sex, age and task-characteristic factors.

Disentangling variability from mean

Across the studies reported in this thesis, various measures of RT IIV were used. For each study a measure of raw and mean-adjusted variability were used and the results obtained from them compared.

The final empirical study, summarised in Chapter 6, investigated the effect of altitude-related hypoxia on RT IIV. This study had two main aims. The first was to investigate whether altitude affected RT IIV. The second was to investigate to what degree RT IIV and mean RT are independent. Observational studies are problematic because it is often not possible to tease out effects of various confounders. Although one may not age somebody experimentally, it is possible to experimentally manipulate RT IIV. It has been suggested before that hypoxia leads to an increase both mean and RT variability. Of particular interest to the study in Chapter 6 was the investigation of whether an increase in RT IIV independent of increase in mean RT could be induced by reducing the availability of oxygen at high altitude.

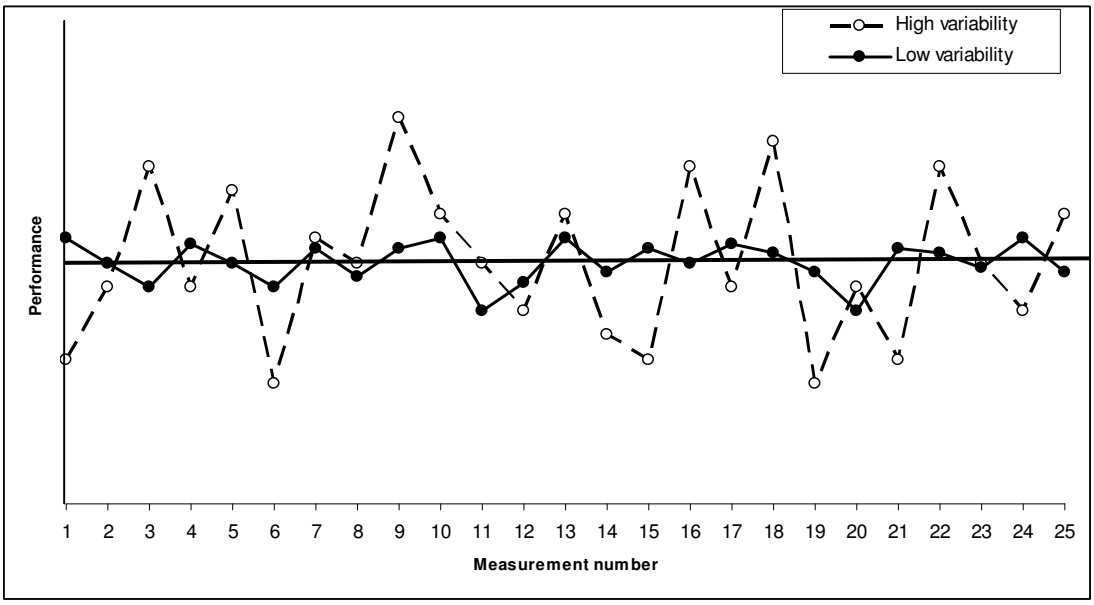

Figure 1.1. A hypothetical trajectory of performance across measurement occasions of individuals with equal mean performance but different variability.

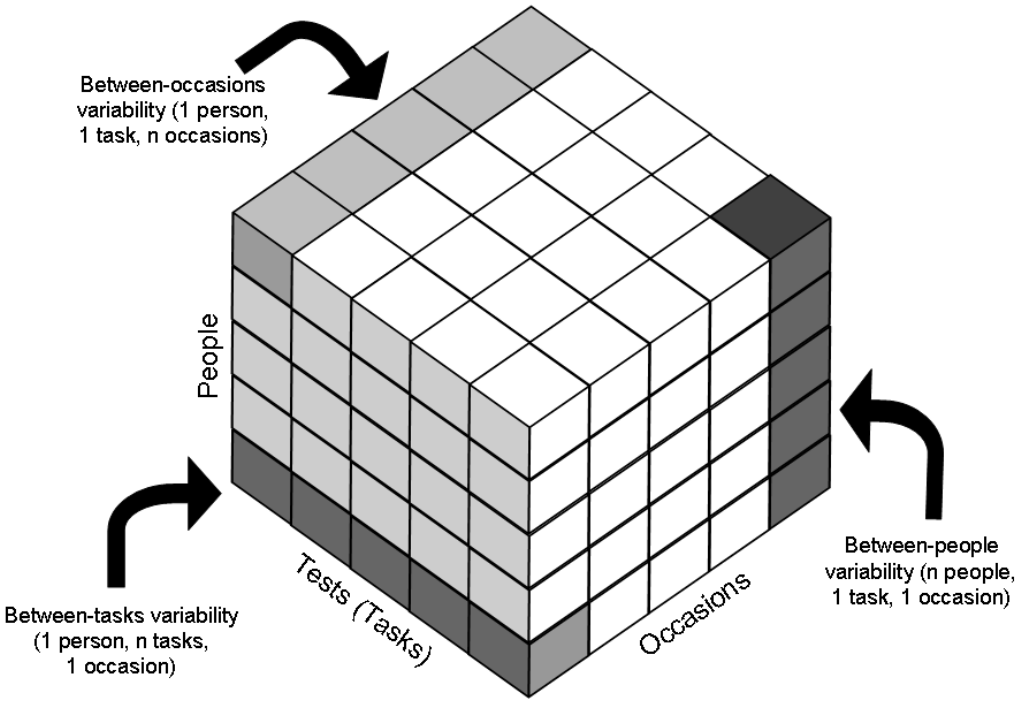

Figure 1.2. Three-dimensional data box (based on Cattell, 1966; Hultsch & MacDonald, 2004), illustrating different types of variability.

Chapter 2 -
Age effects on intra-individual variability in simple and choice reaction time: A systematic review

Intra-individual variability in reaction time (RT IIV)—generally understood as the variability of the responses of one individual on a single test within a single testing occasion—is often reported to increase with age in adulthood (e.g. Anstey, 1999; Der & Deary, 2006; Hultsch, MacDonald, & Dixon, 2002; see also Hultsch, Strauss, Hunter, & MacDonald, 2008 for review). As already discussed in Chapter 1, RT IIV is thought to be an indicator of the functioning of the central nervous system (e.g. Hendrickson, 1982). The available empirical evidence supports this notion, in that greater RT IIV is observed in a number of conditions affecting the central nervous system; for example, in neurodegenerative diseases (Alzheimer's or Parkinson's Disease; Burton, Strauss, Hultsch, Moll, & Hunter, 2006) or following a traumatic brain injury (Stuss, et al., 1989; Stuss, Pogue, Buckle, & Bondar, 1994). Greater RT IIV has also been reported in states that temporarily affect the central nervous system functioning, such as alcohol consumption (Maruff, Falleti, Collie, Darby, & McStephen, 2005), or presence at high altitude (McFarland, 1937 a, b). Moreover, RT IIV is associated with white matter integrity in otherwise healthy adults (Bunce et al., 2010). Taken together, this evidence suggests that increased RT IIV marks a deterioration of the central nervous system functioning.

As well as marking concurrent neurological dysfunction, RT IIV has predictive validity. For example, it predicts cognitive decline over 6 years (MacDonald, Hultsch, & Dixon, 2003), progression from healthy ageing to mild cognitive impairment (Cherbuin, Sachdev, & Anstey, 2010), and even mortality (Shipley, Der, Taylor, & Deary, 2006).

As this review will show, there is also evidence that RT IIV increases with age in normal healthy ageing—that is, even in adults who do not demonstrate clinical impairment. Given its predictive value, RT IIV may be a useful screening tool for

early signs of age-related neuropathology. However, there are a number of issues relating to RT IIV and ageing that remain unresolved to date and that limit the conclusions that can be drawn from the available research. Although many investigations find significant age effect on RT IIV, the magnitude of the effect varies between studies. There are also reports of there being no significant age effects on RT IIV (e.g. Waugh, Fozard, Talland, & Erwin, 1973; Shammi, Bosman, & Stuss, 1998). Therefore, it is desirable to investigate the issue systematically and to attempt to quantify the differences between older and younger individuals as well as to investigate potential sources of heterogeneity in the findings of various investigators.

Magnitude of age differences in RT IIV and the effects of procedural factors

As yet, there have been no documented attempts to investigate or to quantify the differences in RT IIV between older and younger individuals systematically. Many studies that have investigated age effects on RT IIV have found greater levels of variability in older than younger individuals (e.g. Der & Deary, 2006; Hultsch et al., 2002). However, some have not found a significant age effect on RT IIV (e.g. Waugh et al, 1973). Because RT can be obtained from a number of different tasks, it is expected that some of the differences between study findings may be due to differences in the tasks and procedures adopted by different studies. A classical RT task involves a presentation of a stimulus following which a simple response is required. However, there is great scope for variability in how the response is obtained. For example, there might be different stimuli, different intervals before their presentation, and different responses required of participants. All these might alter any age effects on RT IIV.

Considering stimuli, the main source of potential influence on the result is their mode of presentation. It is well-established that RTs differ between modalities (Brebner & Welford, 1980)—they tend to be faster in response to auditory than visual stimuli, and intermediate for tactile stimuli. Therefore, it is possible that RT IIV of responses to stimuli in different modalities is also different, and age effects on

them may also be different. Another aspect of stimulus presentation which might have an effect is the length of the preparatory interval (PI); that is, the time between the beginning of an RT trial and the stimulus presentation. For example, this could be the interval between a *get ready* signal and stimulus presentation. In simple reaction time (SRT), where there is only one stimulus and one response, it is common practice to use variable PIs to minimise anticipatory responses. However, in choice reaction time (CRT) tasks, where there are a number of different stimuli presented each requiring a different response, PIs can be either variable or have a fixed length. RTs of older adults to stimuli presented following a short PI (where PIs are variable) tend to be disproportionately longer than RTs of younger adults (Lahtela, Niemi, & Kuusela, 1985). Thus, differences in RT IIV between older and younger adults might be amplified in tasks which use variable (and short) PIs.

It has also been suggested that age effects on RT IIV tend to be larger in tasks that are more complex (Dixon, Garrett, Lentz, MacDonald, Strauss, & Hultsch, 2007; West, Murphy, Armilio, Craik, & Stuss, 2002). However, the tasks that are considered *more complex* often involve executive processing to a greater extent. For example, both Dixon et al. and West et al. used a 1-back task, which involves retaining a stimulus presented at a previous trial in the working memory and comparing it with the current stimulus. On the other hand, Hultsch et al.'s (2002) data show the opposite pattern: the RT IIV differences between younger and older groups were less marked for more complex tasks (requiring a semantic or lexical decisions) than for SRT or CRT tasks. Given these findings, it becomes apparent that some more complex tasks may require different cognitive mechanisms than *simple* tasks and, as such, these classes of tasks may not be directly comparable. Any comparisons of age-group differences in RT IIV from such different tasks are, at best, difficult to interpret due to multiple potential sources of any differences found.

This review will address the issue of age effects on RT IIV from tasks of different difficulty. However, the tasks considered herein will not involve complex processing or memory demands. Comparisons will be made for SRT, which requires the detection of a stimulus and response execution; and for CRT, which requires additional steps of stimulus identification and the selection of an appropriate

response. In addition, the number of possible choices in the CRT task will be considered as a measure of task difficulty.

Comparing the old with the young and the middle-aged

Studies which set out specifically to test age differences in RT IIV often employ extreme age groups; that is, young and old groups with narrow age ranges and separated by a large gap (e.g. Smulders, Kenemans, Jonkman, & Kok, 1997: 18-24 versus 62-73; Yan, Thomas, & Stelmach, 1998: 20-30 versus 65-80). Few studies include middle-aged adults in their samples. Because it has been suggested that adult-age developmental trajectories of different RT tasks differ (e.g. SRT and CRT; Der & Deary, 2006), considering only youngest and oldest individuals may lead to some important information about mid-life RT IIV being overlooked. It is useful to determine the magnitude of differences between old and young adults, as well as between old and middle-aged adults. Therefore, reviewed here are studies which collected data on SRT IIV and CRT IIV from older adults and allowed comparisons of that group with either young adults or with middle-aged adults. It was not within the scope of this review to compare young and middle-aged groups, although it would be a valuable future addition.

Different measures of RT IIV

A further issue which needed addressing was whether age differences found in RT IIV depend on the variability measure used. There is no consensus on how to best conceptualise IIV and, consequently, several different measures are currently used (see Chapter 1 for descriptions of various RT IIV measures). The main distinction that can be drawn between the different methods of obtaining IIV is whether or not they adjust for RT mean. As discussed in Chapter 1, a major methodological issue in the study of IIV is its correlation with RT mean. The direction of causality of the association is not well understood. An increase in mean RT may drive the increase in IIV but, the opposite may be true if some variability-producing forces are at play and it is the IIV that drives increases in mean RT. The

third possibility is that mean RT and RT IIV are not causally linked at all, but instead both reflect influences from another, shared source or sources. To overcome this problem, many researchers adopted various methods in an attempt to account for the association between mean RT and RT IIV. These include partialling out the effects of age on individual trial RTs prior to calculating ISDs (thus obtaining what is sometimes termed *purified* residuals, e.g. Hultsch et al., 2002), including mean RT as a covariate in models of RT IIV, or simply dividing ISD by the individual mean RT to obtain a CV. In this review, the measures which take mean RT into account will be referred to as *mean-adjusted* measures, and those that do not will be referred to as *raw* measures of IIV.

There are some suggestions that RT IIV increases with age only when a raw measure of IIV is used. For example, Shammi et al. (1998) found a significant age difference in CRT IIV, but only when mean RT was not controlled. This finding suggests that links between older age and IIV unadjusted for RT mean may be spurious. If any increase in IIV results from general slowing of RTs, then individual differences in IIV might not be of clinical or practical importance. However, a number of researchers report significant age effects on IIV even after controlling for individual mean RT (e.g. Hultsch et al., 2002; Williams, Hultsch, Strauss, Hunter, & Tannock, 2005, Williams, Strauss, Hultsch, & Hunter, 2007). Others, who use more than one measure of IIV, typically find an attenuation of effect sizes for differences in IIV adjusted for the RT mean (e.g. Der & Deary, 2006; Dykiert, Der, Starr, & Deary, in revision; and Chapter 3 of this thesis). Given the lack of consistency in findings using different measures on RT IIV, two questions remain: (1) are there age differences in IIV that are over and above age differences in mean RT?; and, if so, (2) are they similar across different measures of mean-adjusted IIV?

Given the issues outlined above, the aims of this review were as follows:

1. To establish and quantify the differences in RT IIV between older adults and two groups of younger adults (young and middle-aged).
2. To investigate whether any differences found are more prominent in more difficult tasks (e.g., in CRT rather than SRT)
3. To investigate differences between age groups in the two broad types of variability measures: raw IIV and IIV adjusted for RT mean.

4. To investigate whether and how various procedural factors (such as stimulus modality, length and variability of PI, or response type) affect the age differences found

Methods

Search strategy

Searching for relevant studies was performed in two stages. First, a search of electronic databases was performed mainly in the first half of the year 2008. From this point onwards, this search will be referred to as *the general search*. Because of the large number of studies from which the full text needed reviewing, assessing eligibility of studies identified was a lengthy process. Therefore, another search was performed, aiming at including any relevant studies published between the time when the general search was performed and the end of evaluation of the studies identified through it. The second search will be referred to as *the update search*.

The general search. Four electronic databases, PsycINFO (accessed via EBSCOhost), Medline (via OvidSP), Web of Science (via ISI Web of Knowledge), and EMBASE (via OvidSP) were searched for relevant studies from their respective inception to the date when the search was performed. PsycINFO was searched from 1806 to 4/01/08, MEDLINE from 1966 to 13/03/08, Web of Science from 1900 to 23/06/08, and EMBASE from 1980 to 31/07/08. The broad aim of the search strategies was to identify studies which used a RT test and considered variability in the responses. Therefore, the general format of the strategies was as follows: (1) to identify a set of studies which used a RT test; (2) to identify a set of studies that considered IIV; and (3) to identify studies which were flagged as belonging to both these sets (i.e., those that considered IIV in an RT task). Where possible, the fields searched included title, abstract and keyword or subject heading, but these varied between databases. Full lists of search terms for each database are shown in Appendix A. Hand searching involved scanning the reference lists of all studies selected for inclusion.

The update search. The same four databases were searched again for studies published in years 2008 and 2009. Consequently, all databases were searched from their inception to 31/12/2009. The search strategy was the same as before, with the exception of PsycINFO which, during the update search, was accessed via Ovid SP (access to EBSCOhost was no longer available). Again, reference lists of the relevant articles were scanned for additional studies for potential inclusion.

Study selection

The present systematic review, aiming to compare RT IIV differences between three age groups, was part of a larger review investigating RT IIV in different contexts. Therefore, the study selection was performed in two stages. In the first instance, studies from the general search were evaluated using the general inclusion criteria outlined below. All studies that were retained after the first selection were later screened for relevance to the present systematic review using an additional set of criteria (see ageing-specific inclusion criteria below). Throughout the screening process, any uncertainties were discussed with the thesis supervisors and consensus was sought. On no occasion was it necessary to contact authors of original papers to resolve the issues.

General inclusion criteria

Studies to be included in the review were evaluated using the following criteria: language and publication status, study sample, availability of empirical data, RT task, IIV measure, and relevance of the subject matter. These are described in detail below.

Language and publication status. Included studies had to be journal articles, books or book chapters published in the English language. Unpublished studies, datasets not yet analysed or reports published in sources that are not peer-reviewed (e.g., technical reports), were not included. The lack of full listing of such sources precludes a systematic approach to their review.

Study sample. Studies were included if they employed a sample of human adults, aged 18 years or above. Younger participants, aged 16 or 17, were deemed acceptable if they were included in a broader age group, for example 16 to 25. If the age of participants was not specified but a general description was provided, indicating that the sample included adults (e.g. *university students*, *sophomores*, *young adults*, or *elderly*), the criterion was considered to be met. Any studies that used simulated data rather than data collected from human participants were excluded.

Availability of empirical data. To be considered eligible for the review, studies had to analyse quantitative empirical data. Thus, opinions, commentaries, theoretical or review papers were not included.

RT Task. The studies were included in the review if they used a qualifying SRT or CRT task. A task was considered eligible if it met all the following criteria: (1) it was a SRT or CRT task – with one or multiple possible responses, respectively; (2) it was of a stimulus-response nature, where each stimulus had a pre-determined response assigned to it (for example stimuli 1 and 2 were mapped to response buttons 1 and 2, respectively); (3) it required conscious and voluntary responses (e.g., not reflexes or responses resulting from transcranial stimulation), the latency of which could be objectively determined: for example, a button press or release was acceptable but an EMG record was not; (4) it was administered as a single task (with no concurrent task being carried out as in a dual-task paradigm, for example); (5) any experimental manipulations were not related to the modality of the stimulus presentation or response (e.g. no visual degradation of stimuli in a visual RT or introducing of extraneous noise in an auditory RT task). A task was not considered eligible if it involved at least one of the following: (1) higher cognitive decision or judgement (e.g., true or false; same or different); (2) categorisation (e.g., animal/ not animal); (3) executive function (e.g., inhibiting a response in a go/no-go or stop-signal task); (4) memory (e.g., in an n-back task).

IIV measure. To be eligible, a study needed to consider RT IIV across trials within a single testing occasion. There was no restriction in terms of the measure of IIV used; any measure claimed by the study's authors to reflect within-subject

variability was accepted. However, measures of IIV calculated across occasions or across different tasks were not eligible for this review.

Relevance of the subject matter. Following piloting of the selection criteria, five broad categories of interest were identified. These were: the effect of age in healthy ageing; the effect of diseases or ill health; the effects of drugs or states (e.g. psychoactive drugs, sleep deprivation, altitude); links with bodily functioning; biological or neurological measures (e.g., forced expiratory volume [FEV]; grip strength); and mortality risk. Studies were deemed to be relevant to the review if they investigated (or provided sufficient information for the reviewer to investigate) issues falling into any of these five categories. Consequently, even if a study met all the earlier criteria but was merely concerned with, for example, task manipulation such as PI variability, number of potential responses, or stimulus intensity, it was not considered relevant.

Ageing-specific inclusion criteria

For the systematic review reported here, additional inclusion criteria were adopted in terms of the sample. Firstly, the sample had to include at least 10 healthy adults aged 60 years or older. Secondly, it had to include one or both of the following: 10 or more healthy young subjects (aged about 18 to 39 years); 10 or more healthy middle-aged subjects (aged about 40 to 59 years).

Multiple publications

In the case of multiple publications from a single study (on the same or related datasets) only one was selected for inclusion in the review. The primary rule used was to select the publication from which most relevant data were available.

Quality assessment

Quality assessment of the included studies comprised evaluation of the sample representativeness and data treatment. The degree of sample representativeness was assessed depending on the sampling procedure and coded as

representative (random sampling or whole population sampled), likely to be representative (systematic or stratified sampling), and not likely to be representative (non-probability sampling, e.g. purposive, convenience, or snowball). The former two categories were considered representative and the latter one non-representative for the purpose of later subgroup analysis. If insufficient information was provided in the paper, a judgement was made based on the sample size and the likelihood that the sample was purposive or convenience.

Another aspect of study quality assessment involved the assessment of data treatment. Specifically, it was considered whether or not trial-level trimming was performed. Trial-level trimming usually includes eliminating extremely fast and slow responses, which are likely to result from accidental key presses or a distraction or lapse in concentration, respectively. Excluding such aberrant responses prior to estimating RT IIV improves the reliability of the measure.

Data extraction

Data from the included studies were extracted using a specially-designed data extraction form (see Appendix B). For each study, information about the participants, tasks and RT IIV data for each age group of interest were extracted. Detailed information on tasks were collected, including the number of trials, stimulus type and modality, PI, etc., for later use as possible moderating variables.

The aim of this review was to compare RT IIV between age groups (particularly, older versus younger adults and older versus middle-aged adults) and the target age ranges for these groups were 18 to 30, 40 to 50 and 60 to 80 years. Where no age groups were created in the publication, the authors were contacted and asked to provide data for these groups. If the sample in a study was already subdivided into age ranges, then the relevant groups were selected. The criteria for selecting an age group were as follows. Groups which were contained in the ranges 20 to 30, 40 to 50 and 60 to 70 years were selected over those that crossed the boundaries. If more than one group was contained within a single age range, then the one closest to or containing the mid-point of the respective range (25, 45, or 65) was selected. For the youngest group, ranges excluding ages of below 18 years were

given priority over those including teenagers. It was not always possible to obtain age groups falling within the age ranges specified, so groups that included participants with ages outside them were accepted. To avoid overlap between groups, restrictions were imposed such that the boundaries (with 1 year tolerance) were, 16 to 39 for young, 40 to 59 for middle-aged, and 60 years and above for old (no upper limit was defined for the old group).

The measure of interest in the present review was RT, defined as time elapsed between a stimulus to execute a pre-determined response and the execution of this response. However, occasionally, rather that reporting the overall RT, authors record RTs fractionated into decision (DT) and movement (MT) components; for example, when a response involves releasing a home key and pressing a response key (e.g., Spirduso & Clifford, 1978). For these studies, DT rather than MT was selected, as it is intended to capture the time taken to complete the cognitive, rather than motor, component of the task.

All authors who provided only one type of IIV measure (raw or mean-adjusted) were contacted and asked to provide the other. Authors of studies from which no IIV data were reported, but where it was implied in the report that they have been collected and/or considered (for example, when RT distributions were considered; Smulders et al., 1997), were also contacted. To assist comparability across studies, whenever contacting authors, the most commonly used raw and mean-adjusted measures were requested: RT ISD and CV, respectively.

If an author could not be contacted or the original data were not available, the paper was screened for other usable data sources. For studies which reported significance tests for comparisons between two groups, effect sizes were estimated from these statistics using conversion formulae provided by Wolf (1986). This was done for three studies (Hogan, 2003; Shammi et al, 1998; Smulders et al., 1997) and details are given in Appendix C. If the relevant data were only presented graphically, the graphs were digitised using Engauge Digitizer software (version 2.15). Each graph was digitised twice and the values obtained from each were averaged to form more reliable estimates. Digitisation of graphs was performed for four studies (Hogan, 2003; Pierson & Montoye, 1958; Sparrow, Begg, & Parker, 2006; Wilkinson & Allison, 1989) and the details are also presented in Appendix C.

Results

Figure 2.1 illustrates the flow of studies through the review process. Titles and abstracts identified by the general search (n = 11,544) were screened for relevance using the general inclusion criteria specified above. If the title and abstract did not provide sufficient information to justify exclusion of a study, a full text was obtained for further screening. Overall, 1,036 full texts were retrieved for detailed consideration, out of which 848 were deemed not relevant and excluded (see Appendix D for a list of all excluded studies, with reasons), and 188 were retained for further review.

For the purpose of this review of age differences in RT IIV, the 188 retained studies were screened further for the appropriate population, ages and group sizes (ageing-specific inclusion criteria). Following this additional screening, 156 studies were excluded (123 because of an inadequate age range, 22 because of insufficient group size in one or more groups, and 11 which did not include healthy participants; see Appendix E). Four additional relevant studies were identified through hand searching of reference lists of the retained papers. Therefore, the total number of eligible studies identified via the first, general search was 36. Five studies reported on data that overlapped with other studies included in the review, and were therefore excluded. A further three studies were excluded because the data necessary for them to be included in the review were unobtainable. These exclusions left 28 studies to be reviewed.

The update search resulted in 2,417 hits and titles and abstracts of these were screened for relevance to the healthy aging systematic review using the specific criteria including age and group size restrictions. In total 90 full texts were obtained for detailed review, out of which six studies met all inclusion criteria (see Appendix F for a list of considered studies from the update search, with reasons for exclusion where relevant). One study reported data which overlapped with another paper already included in the review. Therefore, five eligible studies identified in the update search were included in the review. Reference lists of these papers revealed no additional studies eligible for inclusion.

In total, 33 studies were included in the review; 18 studies had data on SRT IIV and 24 had data on CRT IIV. The results pertaining to SRT and CRT IIV will be presented in turn. Details of the samples and age groups of all included studies are presented in Table 2.1.

SRT IIV

Eighteen studies contributed SRT IIV data to this review and the details of their tasks and data are summarised in Table 2.2 (see also Table 2.1 for information on age groups). There was an identifiable methodological heterogeneity between the studies. In terms of stimulus modality, most studies used a visual mode of stimulus presentation. The visual stimuli used varied between studies, but were usually static; i.e., they appeared at the beginning of a trial and remained unchanged until a response was made. Commonly-used stimuli were a light (n=5), letter (n=3), shape (n=2), or digit (n=2). Two studies used a dynamic stimulus, in the form of a timer which started at the beginning of a trial and continued to increment until a response was made. Four studies used auditory stimuli (i.e., tones or a buzzer sound). Gorus, De Raedt, & Mets (2006) tested both visual and auditory SRTs and data on the former were selected in keeping with the majority of the remaining studies. Most of the responses to the stimuli in the reviewed studies involved either pressing a response key or releasing a home key. In one study that differed from the others (Yan et al., 1998), participants were asked to move a hand-held stylus repeatedly between two target circles. The RT was measured between the "go" signal and the initiation of the back-to-front movement series. Given that SRT does not require a response selection, if stimuli occur at regular intervals it may elicit anticipatory responses. Not surprisingly, most SRT studies used variable PIs. However, this could not be determined for three studies. The length of PIs varied greatly across studies; the shortest PI was 150 ms and the longest 25 s (overall median = 3 s). The number of trials per study varied from 8 to 100 (median = 30.5).

The obtained raw SRT IIV measures were either dispersion around the individual mean, (i.e., ISD or variance; n=12) or percentile difference, such as inter-quartile range (IQR; n=3). Adjusted SRT IIV measures were either coefficient of

variation (CV; six based on ISD, e.g., ISD/mean RT or ISD/median RT; and two based on percentile differences, e.g., IQR/mean RT) or variability index (VI = ISD/(1000/mean RT); n=1). Three studies used a regression method to control for RT mean differences and calculated ISD on the purified residuals. Hogan (2003) also used regression method to control for central tendency in RT (median in this case), but did not calculate ISD from residuals. Instead, R^2 from regression models including age and median RT were used to estimate RT IIV differences between age groups. Finally, Pierson and Montoye (1958) provided a crude measure of the frequency of mode (explained in detail at the end of this section), which was then averaged across participants in each age group.

Out of the 18 studies included in the review, 15 had sufficient data to allow their inclusion in the meta-analysis: 13 provided data on raw SRT IIV, and 13 provided data on SRT IIV adjusted for mean SRT. Sufficient data were not available from three studies and they could not be included in the meta-analysis (Pierson & Montoye, 1958; Spirduso & Clifford, 1978; Surwillo, 1963). These studies and their findings are briefly summarised at the end of this section. The remaining SRT IIV studies contributed data to four meta-analyses: raw SRT IIV in old versus young participants, adjusted SRT IIV in old versus young participants, raw SRT IIV in old versus middle-aged participants, and adjusted SRT IIV in old versus middle-aged participants.

Heterogeneity between studies was assessed in each meta-analysis using Cochran's Q test. There was significant heterogeneity among studies in all four comparisons (all ps < .001; I^2 range 74.89 to 87.70); therefore, random effects method was applied to pool effect sizes.

Raw SRT IIV: old versus young

Effect sizes for the old-young group difference in raw SRT IIV could be obtained from 13 studies. The forest plot in Figure 2.2 presents effect sizes (Cohen's *d*) for all old-young comparisons. To evaluate the magnitude of the effect sizes, values of 0.20 for a small effect, 0.50 for a medium effect, and 0.80 for a large effect,

were used as suggested by Cohen (1992). The pooled effect size was medium in magnitude, $d = 0.582$ ($Z = 10.220$, $p < .001$), and indicated that raw IIV was larger in older than in younger groups.

To investigate possible sources of between study heterogeneity, subgroup and meta-regression analyses were performed. The groups were defined based on sample size (small: <100, medium: 100-1000, large: >1000), stimulus modality (auditory vs. visual), and two measures of study quality: sample representativeness (likely to be representative vs. not likely), and trial-level data trimming (performed vs. not performed). The effect sizes for each subgroup are presented in Table 2.3. Given that the number of studies in subgroups was generally small, τ^2 (the true variance between studies) was estimated for each subgroup separately and later pooled. This procedure does not reflect the assumption that variance between studies is the same for all subgroups, but is used to overcome the imprecision in the estimate of τ^2 that is likely in subgroups comprising a small (<5) number of studies (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Sample size or sample representativeness did not explain much between study heterogeneity ($p = .584$ and .149, respectively), although effects were slightly larger in representative samples. Effect sizes were also a little larger in the subgroup of studies which adopted trial-level trimming, but not significantly so ($p = .253$). The effects tended to be slightly larger in studies in which the stimuli were presented visually rather than aurally ($p = .062$).

Meta-regression was performed on three continuous variables: age of the older group, the number of trials within the test, and the length of PI (see Figure 2.3). Age of the older group was mean, median or midpoint of ages for the group, depending on which measure was available. For studies in which PI was fixed, the exact value of PI was used; where PIs were variable, the median or the middle value was used.

For studies which considered raw SRT IIV, effect sizes were larger for older old groups ($B = 0.031$, $se = 0.015$, $p = .042$), and smaller with longer PIs ($B = -0.040$, $se = 0.019$, $p = .034$). There was no significant relationship between effect size and the number of SRT trials.

Raw SRT IIV: old versus middle-aged

Six studies provided data which contributed to the old versus middle-aged comparisons of raw SRT IIV (see Figure 2.4). When older and middle-aged groups were compared, the former group had greater raw SRT IIV, and the overall effect size was small, $d = 0.327$ ($Z = 5.002$, $p < .001$). The modest number of studies for this comparison did not permit an investigation of sources of heterogeneity.

Mean-adjusted SRT IIV: old versus young

Thirteen studies considered old-young differences in SRT IIV adjusted for mean SRT and provided sufficient data to allow meta-analysis (see Figure 2.5). The overall effect size for the comparison was small, $d = 0.370$ ($Z = 4.960$, $p < .001$). It was in the expected direction, with older adults demonstrating greater variability than younger adults, even when the differences in mean SRT were controlled.

Subgroup analyses were performed, as was done with raw SRT IIV, using sample size, modality, representativeness and trial-level data trimming. In addition, the measure of mean-adjusted IIV was considered, with four groups: CV based on variance or SD (e.g. ISD/mean); CV based on percentile difference (e.g. IQR/mean); ISD calculated from residuals purified of influences of mean SRT; and other, for measures which did not fall in either of the three categories. τ^2 were pooled across subgroups to reduce the imprecision of within subgroup heterogeneity estimates where the number of studies is small. The effect sizes for the subgroups considered are given in Table 2.3.

Studies with medium sample sizes tended to produce largest effect sizes ($p = .073$ for overall between group heterogeneity). Both small and large studies had lower estimates, and the effect size estimate for the <100 subgroup was not statistically significant. The difference between older and younger individuals appeared to be larger in studies with less representative samples, and the effect was not significant in the subgroup with more representative samples. The effect sizes from representative and non-representative samples were significantly different from each other, $p = .002$. In terms of trial-level data trimming, the effect sizes were a

little larger for studies that adopted some form of trimming than those that did not. However, this difference did not reach statistical significance (p = .323). There was no significant difference in effect sizes between studies visually presenting stimuli and a study (n=1) with an auditory presentation (p = .710), although the effect size was larger for visual presentation. Finally, when subgroups based on the SRT IIV measure were considered, it appeared that effect size was smaller for CV (based on ISD or variance) than for the ISD calculated from residuals purified of mean RT. Effect size for CV which used a percentile difference divided by individual mean was close to 0 and not statistically significant. However, the overall differences between subgroups based on SRT IIV measures were not statistically significant (p = .199).

Meta-regression was performed with mean age of the old group, number of trials, and PI (see Figure 2.3). Effect sizes were larger for older groups with higher mean age (B = 0.042, se = 0.019, p = .025); the number of SRT trials and PI were not significantly related to effect sizes.

Mean-adjusted SRT IIV: old versus middle-aged

A comparison of SRT IIV adjusted for SRT mean was possible for seven studies (see Figure 2.6). The pooled effect size was small (d = 0.167, Z = 2.745, p < .001) and again the direction was as expected; that is, older groups showed more IIV than younger groups.

Subgroup analyses or meta-regression were not performed for this comparison, due to a small number of studies included.

Attenuation of age group difference effect sizes by adjusting SRT IIV for SRT mean

Having carried out the four meta-analyses that examined age differences in SRT IIV, a pattern emerged in which effect sizes appeared to be larger for older than younger groups, and larger if SRT IIV was not adjusted for SRT mean. In order to

estimate the degree to which effect sizes are attenuated by adjusting SRT IIV measures for mean SRT, pooled estimates were obtained from studies that contributed data to both analyses within each age group comparison. In other words, meta-analyses were re-run for both old versus young and old versus middle-aged groups, but only on a subset of studies that had usable data on both raw and mean-adjusted SRT IIV. There were 11 such studies which considered old versus young differences and six studies that considered old versus middle-aged differences. When comparing older with younger groups, effect sizes are attenuated by 44.8% by using a SRT IIV measure that is adjusted for mean SRT (raw SRT IIV d = 0.592; mean-adjusted SRT IIV d = 0.327). For old versus middle-aged differences the attenuation of effect sizes was 55.1% (raw SRT IIV d = 0.343; mean-adjusted SRT IIV d = 0.161).

Evidence from studies not included in meta-analyses of SRT IIV

A relatively early study by Pierson and Montoye (1958) used an unusual measure of SRT IIV, namely, frequency of mode. This simple measure reflects the consistency of responding, with higher values reflecting less variability. Data were only presented graphically and the estimates were obtained from a graph. Authors themselves report that IIV decreases until the age of about 30, followed by an increase. The mean frequencies of mode obtained from the digitised graph revealed that they were largest in the youngest group (4.518), smallest in the oldest group (1.972) and intermediate in the middle-aged group (3.120). Thus consistency was lower in the old group than either the young or middle-aged group, and the difference was most marked for the young-old comparison.

Spirduso and Clifford (1978) reported mean SRT ISD of younger and older participants who were either physically active (racketball players or runners, with a history of training four times a week) or not active (never engaged in any sports on a regular basis). In the non-active participants SRT ISDs were slightly lower for the young (32 ms) than the old group (38 ms). However, the difference between active

and non-active groups was much greater than between younger and older groups, with the latter not statistically significant.

The final study which considered age and SRT IIV but was not included in the meta-analysis, is Surwillo (1963). In that study, SRT ISD correlated positively with age, with a small effect size (r = 0.26). The correlation coefficient reported was not used in the meta-analysis because it was calculated for the age range crossing the age group boundaries defined for this review. Consequently, it would not be possible to obtain an estimate of an effect size for old versus middle-aged difference without it being influenced by younger subjects in the sample. Instead, a regression equation provided by the author was used to calculate predicted ISD values at ages equivalent midpoints of middle-aged and old age ranges (45 and 65). No estimate was obtained for midpoint of the young group because it was outside the age range of the sample used in this study. The obtained values were 27.39 and 30.99, indicating that IIV of adults aged 65 was 3.6 ms greater than IIV of adults 20 years younger.

CRT IIV

The review process identified 24 studies with CRT IIV. Their task characteristics and data obtained from them are summarised in Table 2.4 (information about participants and age groups are presented in Table 2.1). All studies used visual stimuli, but one used both visual and auditory presentation (Bherer et al., 2006). Stimuli differed between studies. The most frequently used stimuli were lights (n=7), digits (n=5), letters (n=5), and circles (n=4). Three studies used stimuli indicating direction: arrow (Duchek et al., 2009; Hogan, 2003) and fish (like an arrow, but adapted to make the task more appealing to children; Williams, et al., 2007). In most studies, responses involved either pressing a response key or releasing a home key (n=23). One study used an unusual response type, with two keys pressed in order determined by the stimulus (e.g. LR required pressing the left key first, followed by the right key; Shammi et al., 1998). Concerning the PI, most CRT tasks had a fixed PI (n=14), with only six studies using a variable PI; three studies did not report the information and one used both fixed and variable PIs. PIs ranged from 0 (immediate stimulus onset following a response) to 20s, with an

overall median of 800 ms. The number of trials per study ranged from 15 to 513 trials, with a median of 40.

Raw CRT IIV measures obtained from the included studies were either ISD (n=18) or IQR (n=2). For most studies adjusted IIV measures was the CV (n=11; 1 based on IQR). Eight studies used various versions of regression method to partial out effects of RT mean and calculated ISD from the "purified" residuals. For two studies, other measures were available, including age group R^2 from regression with median RT entered as a covariate (Hogan, 2003), and an unusual conceptualisation of IIV as RT mean^2/variance (Fozard, Thomas, & Waugh, 1976), which is equivalent to 1/CV^2.

Of the 24 studies which considered IIV in CRT in the relevant age groups, 22 contributed to meta-analyses, including 18 studies with data on raw CRT IIV, and 19 studies with data on CRT IIV adjusted for mean CRT. Sufficient data to allow inclusion in the meta-analysis could not be obtained for either CRT IIV measure from two studies (Fozard et al., 1976; Spirduso & Clifford, 1978). In addition, Shammi et al. (1998), which reported sufficient information for raw CRT IIV and was included in a meta-analysis, only provided a verbal account of age effect on mean-adjusted CRT IIV. These studies and their findings are summarised briefly at the end of this section. Main analyses of the remaining studies with sufficient CRT IIV data were comparable with those performed for SRT IIV. That is, four meta-analyses were performed: raw IIV in old versus young participants, adjusted IIV in old versus young participants, raw IIV in old versus middle-aged participants, and adjusted IIV in old versus middle-aged participants.

Cochran's Q test was used to assess the heterogeneity between studies in each meta-analysis. There was significant heterogeneity among studies in all four comparisons (all ps < .001; I^2 range 80.56 to 91.45); therefore, random effects method was applied to pool effect sizes.

Raw CRT IIV: old versus young

Of the identified studies, 18 contributed data to the old versus young comparison of raw CRT IIV. A forest plot summarising individual and pooled effect

sizes for this analysis is presented in Figure 2.7. Cohen's *d* pooled from all studies was 0.960 ($Z = 10.380$, $p < .001$), indicating a large difference between older (more variable) and younger (less variable) groups.

Subgroup analyses were performed to identify some source of variance in the effect sizes of different studies. Subgroups were created based on the following characteristics: sample size (small: <100, medium 10-1000, large >1000), sample representativeness (likely representative vs. likely not representative), and trial-level data trimming (preformed vs. not performed). Subgroups were also created based on CRT task characteristics: whether the responses were spatially determined (spatial vs. non-spatial), on the number of choices (two vs. four), and on whether PI was fixed or variable across task trials. Modality of stimulus presentation was not used here, because all studies employed visual stimuli. Because most raw CRT IIV subgroups included at least five studies, τ^2 was estimated separately for each subgroup and not pooled. The only exception was the analysis based on sample size, which included subgroups with n=3 and n=4.

The results of the subgroup analyses are summarised in Table 2.5. Studies with smaller samples appeared to produce larger differences between older and younger individuals than studies with medium or large samples. Effect size for studies with large samples was larger than effect sizes for studies with medium samples. However, the differences were not statistically significant ($p = .169$). Effect sizes obtained from studies with more representative samples were slightly lower than from studies with less representative samples, but again, there was no significant between subgroup heterogeneity ($p = .561$). Effect sizes were significantly larger in studies which did not perform trial-level data trimming ($p = .029$). The difference between older and younger participants was more marked in tasks with no spatial component; for example, where a response involved discrimination between different letters or digits, rather than between their spatial position ($p = .001$). The number of choices and PI variability did not explain much between-study variance, with subgroups in each very similar in effect sizes ($p = .811$ and .556, respectively).

Bivariate meta-regression was performed with old group age, the number of CRT trials and the PI length as covariates. None of the proposed variables explained

a significant amount of variance in raw CRT IIV difference between old and young groups, although some trends could be observed in scatter plots for old group age (positive); see Figure 2.8.

Raw CRT IIV: old versus middle-aged

Eight studies provided data on differences in raw CRT IIV between older and middle aged groups (see Figure 2.9 for a forest plot). The pooled effect size for this comparison, $d = 0.524$ ($Z = 6.461$, $p < .001$), was medium in magnitude, and lower than the effect size for difference between old and young participants. However, the direction remained unchanged, with older people demonstrating greater IIV. Due to a modest number of studies, sources of heterogeneity were not explored for this comparison

Mean-adjusted CRT IIV: old versus young

Differences between old and young groups in the mean-adjusted CRT IIV were pooled from 19 studies (see Figure 2.10). Older participants showed greater variability, even when it was adjusted for the CRT mean, with a medium effect size, $d = 0.563$ ($Z = 6.344$, $p < .001$). One study, McAuley Yap, Christ, and White (2006), provided an estimate which was a clear outlier. When this study was removed, d increased to 0.632.

Subgroup analyses were performed, including subgroups considered in raw CRT IIV comparison (sample size, sample representativeness, trial-level data trimming, spatial nature of responses, the number of choices, and ISI variability) with the addition of mean-adjusted CRT IIV measure, including CV, ISD from purified residuals, and other. For analyses of subgroups based on sample size and IIV measure, pooled τ^2 was used; all other subgroups comprised more than five studies and so, separate τ^2 were estimated for each.

The outcome of subgroup analyses are presented in Table 2.5. There was little difference in effect sizes obtained from small and medium studies. The

difference between older and younger participants in the largest sample subgroup was smaller than either in small or medium samples, but there was no significant variance between the groups (p = .184). The effect size was larger for the subgroup with less representative samples, but again no significant difference was found between the two groups (p = .135). Whether individual trial data were trimmed or not appeared to explain a degree of between study heterogeneity; the effect size was larger with trimming than without (although, this was only a trend, p = .079). The difference between older and younger participants was slightly smaller in tasks concerned with spatial location of stimuli rather than their formal discrimination. However, there was no significant heterogeneity between these two groups of studies (p = .464). The number of choices also did not appear to explain much heterogeneity (p = .965). There was a marked difference in effect sizes between subgroups based on PI variability. When PIs varied, the effect size for old-young difference in mean-adjusted CRT IIV was negligible (0.027). However, with fixed ISI, the effect size was large and significant (0.923). This large difference was statistically significant (p < .001). Further investigation of this finding revealed that median PI length was greater for variable (2,000 ms) than for fixed PIs (500 ms). Finally, when different measures of IIV were considered, the effect sizes obtained from studies using purified residuals to calculate ISD were notably larger than effect sizes from studies using CV (either based on ISD or percentile differences) or other measures (p < .001).

Meta-regression was performed to assess whether age of the old group, the number of trials or the length of PI could explain some of the between-study heterogeneity (see Figure 2.8). Effect sizes were smaller at larger PIs (B = -0.170, se = 0.081, p = 0.035), but there was no significant effect of either the age of the older group or the number of CRT trials.

Mean-adjusted CRT IIV: old versus middle-aged

Ten studies contributed to the comparison of mean-adjusted CRT IIV between old and middle-aged groups (forest plot can be seen in Figure 2.11). The

pooled effect size was small in magnitude, d = 0.344 (Z = 4.979, p < .001) and revealed that older participants demonstrated greater mean-adjusted IIV in CRT than did middle-aged individuals.

Attenuation of age group difference effect sizes by adjusting CRT IIV for CRT mean

Effect sizes obtained from the four meta-analyses performed on the CRT IIV data, showed a similar pattern to that observed in SRT IIV data. That is, the differences were larger for greater age difference and larger if CRT IIV was not adjusted for CRT mean. The degree of effect size attenuation by adjusting CRT IIV measures for CRT mean was investigated as before: pooled effect sizes were obtained from studies which provided data on both raw and mean-adjusted CRT IIV within both age group comparisons (old versus young and old versus middle-aged). Overall, 16 studies were used for the calculation of effect size for old versus young differences and eight studies contributed data to the old versus middle-aged comparisons. If a mean-adjusted CRT IIV measure is used, the effect sizes are attenuated by 52.4 % in old versus young comparison (raw CRT IIV d = 0.967; mean-adjusted CRT IIV d = 0.460), and by 46.2 % in old versus middle-aged comparison (raw CRT IIV d = 0.524; mean-adjusted CRT IIV d = 0.282).

The number of possible choices in CRT

Some studies provided data on more than one version of CRT task in terms of the number of possible choices. However, to ensure the independence of effect sizes included in meta-analyses, only one level of choice was selected for the main analysis. Three studies reported data from 2-, 4-, and 8-choice RT tasks, with a further two studies reporting data on both 2-, and 4-choice RT tasks. From these, only the 4-choice RT data were included in the analyses performed previously. To allow more of the available data from these studies to be used, additional analyses were performed whereby a separate estimate of effect size was obtained for both

mean-adjusted and raw measures of CRT IIV for 2 and 4-choice RT (there were only three studies which reported data on 8-choice task, rendered insufficient for meta-analysis).

There were altogether 11 studies which provided data on raw 2-choice RT IIV, and 12 on raw 4-choice RT IIV. The effect sizes estimated from these were 0.925 ($Z = 8.574$, $p < .001$) and 0.893 ($Z = 8.292$, $p < .001$), respectively. There were 13 studies that provided mean-adjusted measures of 2-choice RT IIV and 11 studies with mean-adjusted 4-choice RT IIV data. Effect sizes for old-young difference in these were 0.713 ($Z = 4.459$, $p < .001$) and 0.577 ($Z = 4.642$, $p < .001$), indicating a somewhat smaller difference in 4-choice than in 2-choice RT task.

Evidence from CRT IIV studies not included in meta-analyses

Fozard et al. (1976) reported on CRT IIV of participants in the Normative Aging Study. They noted that CRT ISDs were larger for older than middle-aged or younger groups. An ANOVA performed on all age groups included in the sample revealed a significant effect of age. Fozard et al. also considered a measure of CRT IIV adjusted for mean, $\hat{n}$, which represents mean^2/variance ratio (with greater values indicating less variability). There was a significant effect of age on this measure, but post-hoc analysis revealed that only the differences between younger and middle-aged group were significant. However, a significant negative correlation between $\hat{n}$ and age, indicated that intra-individual variability in RT increases with age (and does so to a greater extent that mean RT).

Spirduso and Clifford (1978) measured raw CRT ISD of younger and older participants engaging in different levels of physical activity. For non-active groups, mean CRT ISDs were greater for older (44 ms) than for younger individuals (39 ms). However, the difference was much larger between active and non-active groups, than between older and younger groups. Indeed, activity level, but not age, had a significant effect on CRT ISD.

Finally, Shammi et al. (1998) considered both raw and mean-adjusted measures of CRT IIV. Data provided for the former were sufficient for inclusion in the meta-analysis; however, only a verbal account of the results are provided for the CV. Shammi et al. noted that the difference between older and younger participants was only significant when IIV was conceptualised as a raw ISD. When CV was used, the effect of age was not significant.

Assessing publication bias

The file-drawer problem (Rosenthal, 1979), that is an under-representation of non-significant studies which are not published, poses a genuine threat to the estimated effect size in meta-analytical procedures. Therefore, evidence of publication bias in the extracted data was assessed. This was done for old-young comparisons only, because there were greater numbers of studies included in those than in the old versus middle-aged comparisons. Three common procedures were adopted to assess publication bias. Firstly, the fail-safe N was calculated. This method estimates the number of studies with non-significant results which, if they were included in meta-analysis, would nullify the effect. Secondly, funnel plots were visually inspected and checked for symmetry. Finally, trim and fill technique (Duval & Tweedie, 2000 a, b) was adopted to re-calculate "corrected" effect sizes taking into account potential missing studies with negative effects.

The fail-safe N for raw and mean-adjusted SRT IIV were 1,503 (equivalent to approximately 116 missing studies for each observed study) and 410 (approximately 32 missing studies for each observed study), respectively. For raw CRT IIV, the fail-safe N was 4,109 (228 missing studies per each one observed), and for mean-adjusted CRT IIV the number of non-significant studies which would nullify the effect was 1,236 (65 per each one observed). Overall, the fail-safe Ns were large for all comparisons considered.

Funnel plots of effects based on raw and mean-adjusted SRT IIV and CRT IIV are presented in Figure 2.12. For both SRT IIV measures, studies are centred around the overall effect size. The plots do not reflect the usual funnel shape, but departures from symmetry are not marked. There are relatively few studies with large

standard errors; however, they are under-represented among both positive and negative poles. Trim and fill procedure did not identify any "missing" points and the overall effect size remained unchanged for both raw SRT IIV and SRT IIV adjusted for mean SRT. Funnel plots for CRT IIV were clearly non-symmetrical. For both raw and mean-adjusted CRT IIV, trim and fill procedure (under random effects model) identified a number of "missing" studies. Consequently, the pooled effect sizes were reduced from 0.960 to 0.842 for raw CRT IIV and from 0.563 to 0.451 for mean-adjusted CRT IIV.

Discussion

This meta-analysis consistently found greater IIV at older than younger ages. The finding was consistent in comparisons performed for both SRT and CRT, and in comparisons between different age groups (old versus young as well as old versus middle-aged). Effect sizes were larger for CRT than SRT, supporting the notion that age effects are more pronounced in more demanding tasks (Hultsch et al., 2008). SRT and CRT differ in that CRT requires an additional processing step: i.e., response selection. An obvious explanation for greater difference between older and younger participants in IIV in CRT than SRT is that an increased number of mental operations provides more scope for variability. Perhaps the effects of age on variability occur at different steps involved in a task and are additive. However, neither of the tasks considered in this meta-analysis were thought to involve executive processes, such as planning, problem solving, or inhibition. Therefore a greater age effect in the more difficult task (CRT) could not be attributed to age-related increases in executive control fluctuations as was the case in West et al. (2002).

The investigation of CRT difficulty by comparing effect sizes obtained from CRT tasks involving 2 or 4 stimuli did not reveal a similar pattern, however. When the number of possible choices was considered in subgroup analysis, no significant differences were found between the two groups for either raw or mean-adjusted CRT IIV. When separate meta-analyses were performed on all studies that provided data on 2- and 4-choice RT IIV, slightly smaller effects for the more difficult task with 4

possible choices were obtained. Admittedly, the differences were rather small and further empirical investigations are required to test whether the number of possible choices in a CRT task has an effect on age differences in CRT IIV.

The finding of greater effect sizes in the old versus young comparisons than in the old versus middle-aged comparisons was expected given that age-related increases in RT IIV occur throughout adulthood and that IIV reaches a lifetime minimum in the early 20s (Der & Deary, 2006). However, the fact that the pattern was apparent in both SRT and CRT was somewhat surprising. Earlier accounts suggested that CRT IIV increases throughout the adult age range, whereas SRT IIV remains relatively stable (Dykiert et al., in revision) or even decrease slightly in the early 20s (Der & Deary, 2006). One explanation may be that many of the middle-aged groups included individuals up to the age of 60. Even if RT IIV increases only in old age (defined here as 60 years and above), it is reasonable to assume some increases can be observed among individuals just below this arbitrary cut-off. Therefore, the middle-aged groups might have contained both younger middle-aged individuals (whose SRT IIV are still stable) and older middle-aged participants (whose SRT variability might have already started increasing). The presence of such older and more variable individuals would inflate the mean SRT IIV of the middle-aged group. A relatively small number of studies in this review precluded further analyses to test this hypothesis, such as a meta-regression with mean age as a covariate. In any case, the pattern of smaller differences between the middle and old group than between the young and old group indirectly suggests that whereas older adults have markedly more variable RTs than young adults, some increase in RT IIV occurs already before the age of 60.

An alternative explanation for the finding may be in terms of cohort effects. Given that the age gap between the groups of interest in this review could be as large as 40 years and all studies were cross-sectional, it is possible that cohort effects created (or at least added to) the observed younger-older differences in RT IIV. However, this explanation is not likely to be true, as studies included in this review spanned a few decades and all found similar pattern of age effects. If cohort effects were at play, such that older adults in current times had more variable RTs than young adults because of some influences that were specific to them, then an opposite

effect would be found in studies carried out when the older adults would have been young themselves. By implication, they should have greater variability than their older counterparts who did not shared the consistency-impairing influences. In fact, even in studies published 20 (Wilkinson & Allison, 1989) or even 60 years ago (Obrist, 1953), the same pattern of age effects is apparent: older adults are significantly more variable in RTs than younger adults, rendering the cohort effect explanation unsupported.

Another finding of this review was that greater RT IIV among older individuals was found regardless of whether an IIV measure was or was not adjusted for RT mean. Effect sizes for mean-adjusted IIV were smaller than for raw IIV measures, implying that some (but not all) of the age-related increase in RT IIV is mediated by the slowing of RTs with age. One explanation for this finding is that there might be different "components" of RT IIV. For example, let us assume that RT mean is the driving force increasing IIV (see Chapter 1 for a discussion of other possibilities, such as that greater IIV increases RT mean or that both are influenced by a common other factor). The total age-related increase in IIV may be a combination of greater variability which shares common variance with the slowing of responding (and is therefore removed by adjusting IIV for mean RT) and greater variability due to other causes. RT variability is somewhat "constrained" by the mean in the bottom part of the distribution (i.e. shortest RTs). Considering that there is minimum RT (it cannot be less than 0, and researchers often adopt a theoretical minimum RT of 100 ms or 150 ms thought to be the minimum amount of time required to execute a motor response) but not a maximum limit, an increase in RT mean could be expected to have an effect on the shape of the distribution whereby it becomes less positively skewed. In other words, if all IIV was due to mean RT increase, then one could expect more increase in the number of very short responses (which are more constrained by short overall RTs) than very long responses. There is no reason to expect a disproportionate increase in very long RTs with increasing RT mean, as these are theoretically possible at any level of overall speed of responding. However, some authors note that such an increase in very long RTs (sometimes termed attentional blocks) is indeed observed among older adults (e.g. West et al., 2002; Williams et al., 2005). It was not the focus of this meta-analysis to separately

investigate slow and fast portions of the RT distribution; therefore, the proposition of different effects of mean RT on these remains untested. However, the finding of the present review is consistent with the view that both general and specific variability-producing influences may be at play at older ages (Williams et al., 2005). Therefore, it is possible that age-related increase in IIV comprises greater variability brought about by larger mean RT as well as a larger number of very long RTs with other underlying mechanisms (and so, not removed by the adjustment for RT mean). This possibility warrants further empirical investigation.

Explaining heterogeneity

Subgroup analysis and meta-regression identified a few potential sources of between-study heterogeneity. However, these were rarely replicated across different comparisons; i.e., few were common for all SRT IIV, mean-adjusted SRT IIV, raw CRT IIV and mean-adjusted CRT IIV studies.

Sample size appeared to explain some heterogeneity, although it only approached significance in mean-adjusted SRT IIV. Moreover, only in studies of adjusted CRT was the pattern as one could expect; that is, effect sizes were larger for small and medium studies than for large studies. However, this lack of a pattern of larger effect for smaller studies may simply reflect a relatively low "publication" bias in the present review, as it included effect sizes which were not even reported in their original publications. On the contrary, the degree of sample representativeness produced similar results in all comparisons, with effect sizes larger for less representative samples. Although not reaching statistical significance other than in mean-adjusted SRT IIV, sample representativeness led to a pattern consistent across all four measures of RT IIV. This is not surprising, given that many of non-representative samples in studies included in this review were selected specifically to test old-young differences and some young groups consisted primarily of students. Purposive samples of older and younger participants are likely to be more different than similar age groups taken from the population. If different age groups are recruited from different populations, it is likely that the real age effects are exaggerated.

Trial-level trimming did not explain much between study heterogeneity. There was no significant difference between subgroups based on trimming in either raw or mean–adjusted SRT IIV. For raw CRT IIV studies, trial-level trimming was associated with larger effect sizes; however, a trend for the opposite pattern was observed in mean-adjusted CRT IIV. Therefore, no clear conclusion can be drawn from this subgroup analysis. Since trimming often involves excluding aberrant RTs thought to result from accidental responses (very short) or distractions/loss of concentration (very long), this procedure improves the precision of estimates of RT IIV. It follows that with a better measure, more precise effects are obtained, which should produce a clearer picture of the actual age differences in RT IIV. That is, the effect should be larger in studies which performed the trimming. On the other hand, trimming decreases the estimates of variability. In addition, given that RT of older adults tend to include more very long responses (West et al., 2002; Williams et al., 2005), it is likely to reduce the IIVs of older people more than those of younger individuals. Consequently, the effect size would be *smaller* than when no trimming is performed. These opposing forces could act together to render effects of trial-level trimming non-significant.

For studies of adjusted SRT and CRT, different measures of IIV were considered and significant heterogeneity between subgroups based on the IIV measures was found for CRT IIV. Age effects were larger for IIV obtained from purified residuals, than for either CV or other methods. Although CV is often criticised as a method of adjusting RT ISD for mean RT, it does provide adjustment of each individual's ISD for his or her own mean RT. On the contrary, purified residuals are typically obtained from a regression line which is fitted to all participants, and are rarely adjusted for individual speed of responding. Therefore, CV may actually provide a more precise measure of mean-adjusted IIV than ISD of purified residuals.

There was a trend toward larger effect sizes in studies of SRT IIV which used visual rather than auditory stimuli. Given that RTs are usually shorter in response to auditory than visual stimuli (Brebner & Welford, 1980), this effect may reflect greater IIV among older adults brought about by their longer RTs. The lack of a similar trend with mean-adjusted SRT IIV measures provides some support for this

explanation. However, since there was only one study which used an auditory stimulus and provided SRT IIV, this explanation remains tentative.

The comparisons of procedural factors for studies of CRT, included the response type (spatial vs. non-spatial), number of possible choices (2 vs. 4), and PI variability (variable vs. fixed). The finding of larger effect sizes when a response was non-spatial (statistically significant for raw CRT IIV) is a novel one. A possible explanation for the effect may be that tasks with spatial components are largely perceptual rather than cognitive, hence requiring less processing and leaving less scope for age-related variability. The remaining procedural factors did not explain heterogeneity in raw CRT, although for mean-adjusted CRT effects were larger with fixed rather than variable PIs. Given that this effect did not replicate across the two CRT IIV measures, the apparent differences should be treated with caution. However, it is worth noting the magnitude of the difference between the two subgroups: 0.027 for variable (0.172 with McAuley et al., 2006 removed) and 0.923 for fixed PI. This large difference is in the opposite direction to that expected from the existing literature. Given that older adults seem particularly affected by encountering very short PIs among longer ones (Lahtela et al., 1985), the old-young IIV difference should be larger, not smaller, in studies adopting variable PIs. A further investigation of the finding revealed that the PI variability was confounded with PI length, in that fixed PIs were notably shorter than variable PIs. Therefore, it may not be PI variability per se, but rather the length of the interval used that is related to age differences in CRT IIV. The finding would be worth exploring further in future empirical studies.

Finally, results from meta-regression analyses provide some candidate covariates. However, none replicate across the four RT measures considered in this review. Given the relatively small number of studies included in the meta-regression analyses, some null findings could be due to insufficient power. Old group age was related to the old-young effect sizes in raw and adjusted SRT, with the differences larger for older old groups (i.e. old groups with higher mean age). The mean age of old group, however, did not predict effect sizes in CRT. This pattern of findings is in contrast to what could be expected from earlier investigations, suggesting that CRT IIV start to increase with age earlier than SRT IIV (e.g. Der & Deary, 2006). The

value of PI was associated with effect sizes in raw SRT and adjusted CRT, with smaller effects at longer PIs. The number of trials did not predict the magnitude of old-young difference for any IIV measure, which was again contrary to the expectation.

Publication bias

None of the three methods of publication bias assessment revealed bias among studies of either raw or adjusted SRT IIV. There were relatively few studies with large standard errors, but they were "missing" equally from higher as well as lower ends of the effect size distributions. Some publication bias was detected among studies of CRT IIV, however. Funnel plots presenting both raw and mean-adjusted CRT IIV studies were asymmetrical and more studies than expected had large and positive effect sizes (indicating greater variability among older groups). Trim and fill procedure led to a reduction of the pooled effect sizes for both sets of studies.

It is possible that, since effect sizes tend to be larger for raw than adjusted IIV measure, authors choose to publish only the results obtained from raw measures when adjusted IIV shows no difference between young and old. However, given that in this review attempts were made to include data on both measures from each relevant study, this explanation is not likely to be accurate. It should also be noted that the "missing" studies identified by the trim and fill procedure largely fall within the realm of negative effect sizes. In other words, the procedure implies that most of the missed studies rather than finding no difference between the age groups, would find younger groups to have greater variability. This scenario is highly unlikely, and so the extent of the publication bias suggested may be overestimated.

Strengths and limitations

There were a number of strengths of this review. Firstly, it was a large, thorough and comprehensive review of published studies on SRT IIV and CRT IIV. Secondly, strict inclusion and exclusion criteria ensured that the reviewed studies were acceptably similar in terms of RT tasks and participant groups. Thirdly, this

review included all studies from which the relevant data could be obtained, even if age differences in RT IIV were not the main focus of a study. Therefore, it can be expected that effects of publication bias would be reduced. Finally, attempts were made to obtain both raw and mean-adjusted measures from all included studies and this has been successfully achieved for a large proportion of the identified studies.

Among the limitations is the cross-sectional design of the included studies. The number of longitudinal investigations into RT IIV was scarce, and the length of follow-up did not allow comparison of IIV over 60 years with either 20 or 40. An obvious problem associated with cross-sectional investigation is the potential confounding of age differences with cohort effects. However, as already mentioned, given that age group differences are observed in studies carried out a few decades ago as well as those more recent, there are sound grounds for concluding that greater IIV in older than younger ages is a genuine effect.

Another limitation of the review is that despite the attempts to keep studies similar by controlling the nature of RT tasks administered, there were notable procedural differences between those that were included. Studies differed in the type of stimuli used, the PI, the number of trials, and also in the treatment of data (including data preparation prior to analysis). These resulted in significant between-study heterogeneity, yet no clear sources of it could be identified. This issue have been addressed by using random effects models in order to provide more accurate estimates of effect sizes in light of the between-study heterogeneity. One striking finding was that despite the heterogeneity, there was commonality – the vast majority of the studies reported an age effect in RT IIV in the expected direction (i.e., with older groups showing greater RT IIV). Therefore, although the magnitude of the effect varied across the studies, the direction of it did not, suggesting that older-younger difference in RT IIV is relatively robust to procedural differences.

Summary

To summarise, this review established that RT IIV is larger in older than younger individuals. The difference between old and young groups was larger than that between old and middle-aged groups, suggesting that increase in RT IIV is not

limited to old age, but occurs already in mid-adulthood. Age effect on RT IIV was also larger in CRT than SRT, but not for different number of choices in CRT tasks. The differences in RT IIV between older and younger adults were larger for raw than RT mean-adjusted measures of variability, indicating that some of the increase in IIV with age shares common variance with increases in mean RT. However, this finding does not support claims that all of the observed increase in RT IIV can be explained by slowing of responses with age. Procedural factors did not account for much of the between-study heterogeneity, potentially due to a modest number of studies in most identified subgroups. However, the direction of effect sizes across studies and subgroups pointed towards a larger RT IIV in older individuals.

Table 2.1. Sample characteristics of studies included in the review of simple and choice reaction time intra-individual variability studies

Study	Sex	Representa-tive?	Young group		Middle-aged group		Old group		Review in which included
			N	Mean age (range)	N	Mean age (range)	N	Mean age (range)	
Adam et al. (2006)	M	Not likely	12	25.2 (21-31)	-	-	11	66.4 (61-70)	SRT
Anstey et al. (2005)	M,F	Likely	2,404	NR (20-24)	2,530	NR (40-44)	2,551	NR (60-64)	SRT, CRT
Bherer et al. (2006)	M,F	Not likely	12	20 (NR)	-	-	12	70 (NR)	CRT
Bunce et al. (2004)	NR	Not likely	24	25.5 (20-30)	-	-	24	69.3 (60-85)	CRT
Bunce, Handley & Gaines (2008)	M,F	Not likely	77	23.9 (18-30)	38	45.3 (41-50)	96	71.1 (61-80)	SRT, CRT
Bunce, Tzur et al. (2008)	M,F	Not likely	54	22.7 (18-30)	28	45.8 (41-50)	34	64.9 (61-70)	CRT
Deary & Der (2005)	M,F	Likely	658	Approx 24 (23-26)	741	Approx 44 (39-50)	696	Approx 63 (62-66)	SRT, CRT
Der & Deary (2006)	M,F	Likely	1,706	24.1 (18-30)	1,341	44.7 (40-50)	1,563	67.9 (60-80)	SRT, CRT
Duchek et al. (2009)	NR	Not likely	35	20.29 (NR)	-	-	220	71.75 (NR)	CRT
Finkel & McGue (2007)	M,F	Likely	35	32.8 (27-35)	40	44.4 (40-50)	175	66.0 (60-80)	SRT, CRT
Fontani et al. (2004)	M,F	Not likely	17	25.0 (18-29)	17	51.0 (46-57)	17	66.0 (61-77)	SRT
Fozard et al. (1976)	M	Not likely	24	MD=34	24	MD=50	24	MD=69	CRT

Study	Sex	Representa-tive?	Young group		Middle-aged group		Old group		Review in which included
			N	Mean age (range)	N	Mean age (range)	N	Mean age (range)	
				(25-?)					
Fozard et al. (1994)	M,F	Not likely	226	NR (25-34)	187	NR (35-44)	232	NR (65-74)	SRT
Gooch et al. (2009)	M,F	Not likely	16	22.6 (19-29)	-	-	16	72.8 (62-81)	CRT
Gorus et al. (2006)	M,F	Not likely	27	28.5 (19-37)	-	-	27	74.7 (64-84)	SRT, CRT
Hogan (2003)	M,F	Not likely	78	18.8 (NR)	-	-	94	70.1 (60+)	SRT, CRT
Hultsch et al. (2002)	M,F	Not likely	99	23.2 (17-36)	-	-	361	69.6 (65-74)	SRT, CRT
Li et al. (2009)	M,F	Likely	25	25.3 (18-30)	26	45.6 (40-50)	68	70.5 (60-80)	CRT
Obrist (1953)	M	Not likely	25	27.5 (18-39)	-	-	57	71.5 (65-75)	SRT
Martin et al. (2009)	M	Not likely	-	-	29	44.4 (39-50)	39	63.7 (61-70)	CRT
McAuley et al. (2006)	M,F	Not likely	43	19.6 (17-22)	-	-	33	72.9 (61-82)	CRT
Pierson & Montoye (1958)	M	Not likely	60	19-30	40	41-55	40	66-85	SRT
Rakitin et al. (2006)	M,F	Not likely	31	24.4 (18-35)	-	-	32	71.2 (60-86)	CRT

Study	Sex	Representa-tive?	Young group		Middle-aged group		Old group		Review in which included
			N	Mean age (range)	N	Mean age (range)	N	Mean age (range)	
Shammi et al. (1998)	F	Not likely	18	27.8 (20-35)	-	-	18	68.2 (60-75)	CRT
Smulders (1997)	M	Not likely	12	20.7 (18-24)	-	-	12	66.8 (62-73)	CRT
Sparrow et al. (2006)	M	Not likely	10	26.3 (20-32)	-	-	10	71.1 (64-78)	SRT
Spirduso & Clifford (1978)	M	Not likely	15	22.2 (20-30)	-	-	15	64.2 (60-70)	SRT, CRT
Surwillo (1963)	M	Not likely			N/A	45	N/A	65	SRT
West et al. (2002)	NR	Not likely	20	23.9 (19-29)	-	-	20	73.8 (65-83)	CRT
Wilkinson & Allison (1989)	M,F	Not likely	1,189	NR (20-29)	208	NR (40-49)	50	NR (60-69)	SRT
Williams et al. (2005)	M,F	Not likely	47	NR (18-29)	28	NR (45-59)	25	NR (60-81)	CRT
Williams et al. (2007)	M,F	Not likely	80	24.8 (20-29)	93	45.2 (40-49)	27	66.6 (60-76)	CRT
Yan et al. (1998)	M,F	Not likely	20	24.4 (20-30)	-	-	20	70.4 (65-80)	SRT

Note. Numbers of participants reflect actual numbers used to obtain estimates of intra-individual variability and may differ from those reported in papers. M = males, F = females, NR = not reported, SRT = Simple Reaction Time, CRT = Choice Reaction Time

Table 2.2. Summary of task characteristics and data available from simple reaction time intra-individual variability studies

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Response	Data source	Trial-level trimming	Raw IIV measure	Adjusted IIV measure
Adam et al. (2006)	Visual	Timer	Variable, 2,000-10,000	Approx. 100 (NR)	Pressing a response key	Provided by author(s)	Yes - Percentile difference IIV measure eliminates extreme RTs	90th-10th percentile	90th-10th percentile/ mean RT
Anstey et al. (2005)	Visual	Light	NR	80 (NR)	Pressing a response key	Provided by author(s)	Performed	ISD	Mean-independent variability
Bunce, Handley & Gaines (2008)	Visual	Letter	Variable, 300-1,000	48 (8)	Pressing a response key	Provided by author(s)	Performed	ISD	Purified ISD
Deary & Der (2005)	Visual	Digits	Variable, 1,000-3,000	20 (8)	Pressing a response key	Provided by author(s)	Not performed (trial data not available)	ISD	CV
Der & Deary (2006)	Visual	Digit	Variable, 1,000-3,000	20 (8)	Pressing a response key	Provided by author(s)	Not performed (trial data not available)	ISD	CV
Finkel & McGue (2007)	Visual	Light	Variable, approx. 5,000	15 (3)	Releasing a home key and pressing a response key (RT=time to release, DT)	Provided by author(s)	Performed	ISD	CV
Fontani et al. (2004)	Visual	Letter	Variable, 2,000-4,000	80 (5 minutes)	Pressing a response key	Reported in paper	Performed	-	VI = ISD/ (1000/mean RT)

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Response	Data source	Trial-level trimming	Raw IIV measure	Adjusted IIV measure
Fozard et al. (1994)	Auditory	Tones	Variable, 6,000-13,000	20 (46)	Pressing a response key	From a Master's thesis of a co-author	Performed	Variance	Variance/ mean
Gorus et al. (2006)	Visual (Auditory also available but not used here)	Light	Variable, 3,000-6,000	28 (NR)	Releasing a home button and pressing a response button (RT=DT+MT used)	Provided by author(s)	Performed	IQR	IQR/MD*100
Hogan (2003)	Visual	Square	Variable, 1,000-3,000	Approx. 33 (NR)	Pressing a response key	Raw: reported in paper; adjusted: digitised graph	Performed	MD ISD	MD ISD adjusted for MD RT
Hultsch et al. (2002)	Visual	Plus sign	Variable, 500-1,000	50 (NR)	Pressing a response key	Provided by author(s)	Performed	ISD	Purified ISD
Obrist (1953)	Auditory	Blended click and buzzer sound	Variable, 1,000-2,000	50 (25+)	Pressing a response key	Reported in paper	Yes - Semi IQR eliminates extreme RTs	Semi IQR	-
Pierson & Montoye (1958)	Visual	Light	?, Approx. 2,000	15 (15)	Releasing a key (DT)	Digitised graph	Frequency of mode disregards extreme RTs	Mean frequency of mode	-

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Response	Data source	Trial-level trimming	Raw IIV measure	Adjusted IIV measure
Sparrow et al. (2006)	Visual	Letter	Variable, 6,000-18,000	90 (NR)	Pressing a response key	Digitised graph	Performed	ISD	CV
Spirduso & Clifford (1978)	Visual	Light	Fixed, 20,000	50 (NR)	Releasing a home key and press a response key (DT used here)	Reported in paper	NR	Mean of ISD across 5 blocks of 10 trials	-
Surwillo (1963)	Auditory	Tone	Variable, 10,000-25,000	Approx. 20 (NR)	Pressing a response key	Estimated from a regression equation	NR	ISD	-
Wilkinson & Allison (1989)	Visual	Timer	Variable, 1,000-10,000	8 (2)	Pressing a response key	Digitised graph	Performed	-	CV
Yan et al. (1998)	Auditory	Tone	Variable, 150 – 1,000	10 (5)	Beginning of a back to front movement series; removing stylus from a home position	Reported in paper	NR	ISD	-

Note. CV = coefficient of variation, IIV = intra-individual variability, IQR = inter-quartile range, ISD = intra-individual standard deviation, MD = median, NR = not reported, PI = preparatory interval, VI = Variability index

Table 2.3. Summary of subgroup analysis results for simple reaction time intra-individual variability studies

Groups compared	Raw SRT IIV				Mean-adjusted SRT IIV			
	N	ES	Z	p	N	ES	Z	p
All studies	13	0.582	10.220	<.001	13	0.370	4.960	<.001
Sample size				.584 [a]				.073[a]
<100	5	0.531	3.328	.001	4	0.251	1.250	.211
100-1000	5	0.657	7.095	<.001	5	0.543	5.284	<.001
>1000	3	0.531	5.859	<.001	4	0.231	2.371	.018
Sample representativeness				.149 [a]				.002[a]
Not likely	9	0.669	7.990	<.001	9	0.541	6.334	<.001
Likely	4	0.499	6.049	<.001	4	0.156	1.714	.086
Data trimming				.253 [a]				.323[a]
Not performed	3	0.479	4.290	<.001	2	0.203	1.029	.303
Performed	10	0.632	8.449	<.001	11	0.423	4.109	<.001
Stimulus modality				.062 [a]				.710[a]
Auditory	3	0.362	2.815	.005	1	0.287	1.197	.231
Visual	10	0.627	10.488	<.001	12	0.381	4.689	<.001
Mean-adjusted IIV measure								.199[a]
CV (based on variance or SD)	-	-	-	-	6	0.268	2.014	.044
CV (based on percentile difference)	-	-	-	-	2	-0.003	-0.009	.993
ISD on residuals purified of mean	-	-	-	-	3	0.653	3.665	<.001
Other	-	-	-	-	2	0.464	1.751	.080

Note. SRT IIV = simple reaction time intra-individual variability, CV = coefficient of variation, ES = effect size (Cohen's d), ISD = intra-individual standard deviation, PI = preparatory interval

[a] p value for overall between subgroup heterogeneity

Table 2.4. Summary of task characteristics and data available from choice reaction time intra-individual variability studies

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Number of choices, spatial/ non-spatial, task objective	Response	Data source	Trial-level trimming	raw IIV measure	adjusted IIV measure
Anstey et al. (2005)	Visual	Light	NR, NR	40 (NR)	2, spatial, identify location	Pressing a response key	Provided by author	Performed	ISD	Mean-independent variability
Bherer et al. (2006)	Visual and auditory	Letter/tone	Fixed, 500	20 (probably 0, as data from a baseline session)	2, non-spatial, discrimination (B or C; high or low tone)	Pressing a response key	Provided by author	Performed	ISD	CV
Bunce et al. (2004)	Visual	Light	Fixed, 0 (pressing of home key initiated next trial)	100 (20)	2 4 8, spatial, Location of stimulus light	Pressing a response key	Reported in paper	Performed	ISD	Purified ISD
Bunce, Tzur et al. (2008)	Visual	Circle	NR, NR	100 (20)	2,4, spatial, identify location	Pressing a response key	Provided by author	Performed	ISD	Purified ISD
Bunce, Handley & Gaines (2008)	Visual	Circle	Fixed, 500	48 (12)	2,4, spatial, identify location	Pressing a response key	Provided by author	Performed	ISD	Purified ISD
Deary & Der (2005)	Visual	Digits	Variable, 1,000-3,000	40 (8)	4, non-spatial, discriminate between digits (1,2,3,4)	Pressing a response key	Provided by author	Not Performed (trial data not available)	ISD	CV

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Number of choices, spatial/ non-spatial, task objective	Response	Data source	Trial-level trimming	raw IIV measure	adjusted IIV measure
Der & Deary (2006)	Visual	Digit	Variable, 1,000-3,000	40 (8)	4, non-spatial, discriminate between digits (1,2,3,4)	Pressing a response key	Provided by author	Not Performed (trial data not available)	ISD	CV
Duchek et al. (2009)	Visual	Arrow	Fixed, 500	40 (4)	2, non-spatial, identify the direction of arrow	Pressing a response key	Provided by author	NR	-	CV
Finkel & McGue (2007)	Visual	Light	Variable, approx 5,000	15 (3)	4, spatial, discriminate between colours/locations	Releasing a home key and pressing a response key (RT=time to release, DT)	Provided by author	Performed	ISD	CV
Fozard et al. (1976)	Visual	Light	Fixed, 3,000	121 (0)	2, spatial, discriminate between colours /locations	Releasing a key	Reported in paper	NR	ISD	$Mean^2$/variance
Gooch et al. (2009)	Visual	Letter	Fixed, 250	40 (8)	4, spatial, location of X	Pressing a response key	Provided by author	Performed	ISD	CV

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Number of choices, spatial/ non-spatial, task objective	Response	Data source	Trial-level trimming	raw IIV measure	adjusted IIV measure
Gorus et al. (2006)	Visual	Light	Variable, 3,000-6,000	28 (NR)	4, spatial, location of light,	Releasing home key and pressing response key (RT=DT+MT used here)	Provided by author	Performed	IQR	IQR/MD*100
Hogan (2003)	Visual	Arrow	Variable, 1,000-3,000	33 (10)	2, non-spatial, identify the direction of arrow	Pressing a response key	Raw: reported in paper; adjusted: digitised graph	Performed	MD ISD	Median ISD adjusted for median RT
Hultsch et al. (2002)	Visual	Letter	Fixed, 1,000	20 for each level of choice (NR)	2 4 8, spatial, identify location of O	Releasing a home key and pressing a response key (RT used)	Provided by author	Performed	ISD	Purified ISD
Li et al. (2009)	Visual	Circle	Fixed, 1,000	Approx 42 (6 relevant)	2, non-spatial, identify changing colour	Pressing a response key	Provided by author	Performed	ISD	Purified ISD
Martin et al. (2009)	Visual	Light	NR, NR	60 (20)	2 4 8, spatial, identify location	Releasing a home key and pressing a response key (DT used)	Provided by author	Performed	ISD	CV

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Number of choices, spatial/ non-spatial, task objective	Response	Data source	Trial-level trimming	raw IIV measure	adjusted IIV measure
McAuley et al. (2006)	Visual	Circle	Variable, 600-1,000	40 (10 relevant)	2, spatial, identify location of stimulus changing colour	Pressing a response key	Provided by author	Performed	ISD	CV
Rakitin et al. (2006)	Visual	Digit	Fixed, 500	30 (NR)	4, non-spatial, discriminate between digits (1,2,3,4)	Pressing a response key	Provided by author	Performed	ISD	CV
Shammi et al. (1998)	Visual	Letter	Fixed, 750	180 (10)	2, non-spatial, distinguish between LR and RL (order)	Pressing two keys (L and R) in order determined by stimulus	Estimated from F	NR	ISD	CV
Smulders (1997)	Visual	Digit	Both: Variable 1.520-5.680, Fixed 3.020	Fixed PI: 277; variable PI: 236 (100 per block)	2, non-spatial, distinguish between different digits	Pressing a response key	Estimated from F	Performed	IQR	-
Spirduso & Clifford (1978)	Visual	Light	Fixed, 20,000	50 (NR)	3, spatial, identify the location of stimulus light	Releasing a home key and pressing a response key (DT used)	Reported in paper	NR	Mean of ISDs from 5 blocks of 10 trials	-

Study	Modality	Stimulus	PI: variable/ fixed, values (s)	Trials: test (practice)	Number of choices, spatial/ non-spatial, task objective	Response	Data source	Trial-level trimming	raw IIV measure	adjusted IIV measure
West et al. (2002)	Visual	Digit	Fixed, 200	50 (8)	4, non-spatial, discriminate between digits (1,2,3,4)	Pressing a response key	Provided by author	NR	ISD	CV
Williams et al. (2005)	Visual	Letter	Fixed, 500	32 (0 – data from practice block)	2, non-spatial, distinguish between X and O	Pressing a response key	Reported in paper	Performed	-	Purified ISD
Williams et al. (2007)	Visual	Fish	Fixed, 250	20 (10+ mixed conditions)	2, non-spatial, identify direction	Pressing a response key	Reported in paper	Performed	-	Purified ISD

Note. CV = coefficient of variation, IIV = intra-individual variability, IQR = inter-quartile range, ISD = intra-individual standard deviation, MD = median, NR = not reported, PI = preparatory interval.

Table 2.5. Summary of subgroup analysis results for choice reaction time intra-individual variability studies

Groups compared	Raw CRT IIV				Mean-adjusted CRT IIV			
	N	ES	Z	p	N	ES	Z	p
All studies	18	0.960	10.380	<.001	19	0.563	6.344	<.001
Sample size				.169[a]				.184[a]
<100	11	1.127	8.317	<.001	10	0.661	4.804	<.001
100-1000	4	0.710	4.010	<.001	6	0.643	4.268	<.001
>1000	3	0.924	5.004	<.001	3	0.259	1.379	.168
Sample representativeness				.561[a]				.135[a]
Not likely	13	1.019	6.859	<.001	14	0.673	3.526	<.001
Likely	5	0.897	6.091	<.001	5	0.348	3.370	.001
Data trimming				.029[a]				.079[a]
Not performed	5	1.132	14.153	<.001	5	0.318	2.219	.026
Performed	13	0.851	8.461	<.001	14	0.678	4.623	<.001
Response type				.001[a]				.464[a]
Non-spatial	9	1.122	20.908	<.001	10	0.654	4.520	<.001
Spatial	9	0.738	7.124	<.001	9	0.494	3.019	.003
Number of possible choices				.811[a]				.965[a]
Two	7	0.988	5.986	<.001	8	0.589	2.642	.008
Four	11	0.942	8.693	<.001	11	0.577	4.642	<.001
PI variability				.556[a]				<.001[a]
Variable	6	1.072	19.733	<.001	6	0.027	0.191	.848
Fixed	9	0.955	4.980	<.001	11	0.923	4.960	<.001

Groups compared	Raw CRT IIV				Mean-adjusted CRT IIV			
	N	ES	Z	p	N	ES	Z	p
Mean-adjusted IIV measure								.004[a]
CV (based on variance or SD)	-	-	-	-	9	0.271	1.757	.079
CV (based on percentile difference)	-	-	-	-	1	0.262	0.543	.587
ISD on residuals purified of mean	-	-	-	-	8	1.012	6.355	<.001
Other	-	-	-	-	1	0.050	0.117	.907

Note. CRT IIV = choice reaction time intra-individual variability, CV = coefficient of variation, ES = effect size (Cohen's d), ISD = intra-individual standard deviation, PI = preparatory interval.

[a] p value for overall between subgroup heterogeneity

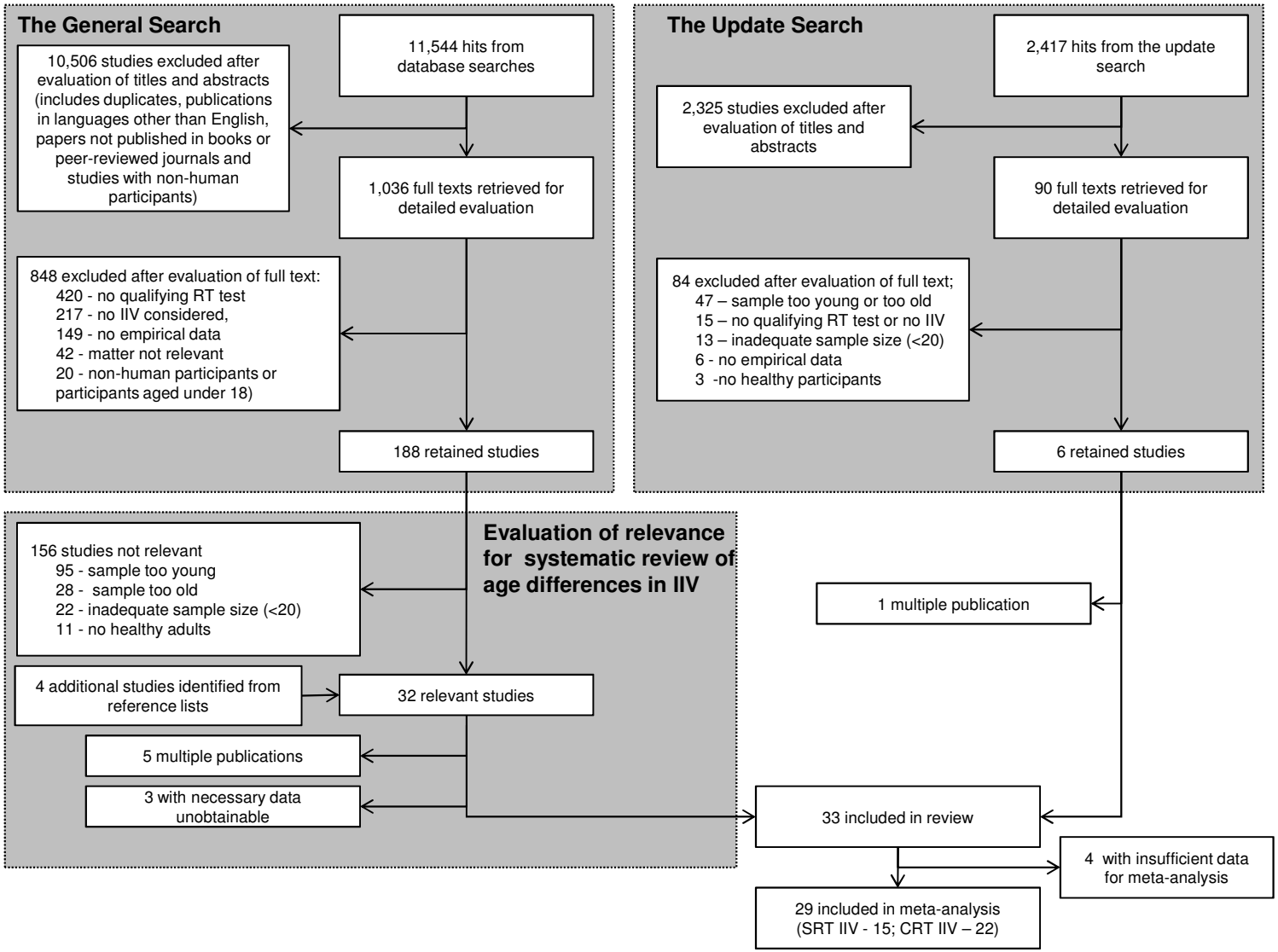

Figure 2.1. Flow diagram of the review process

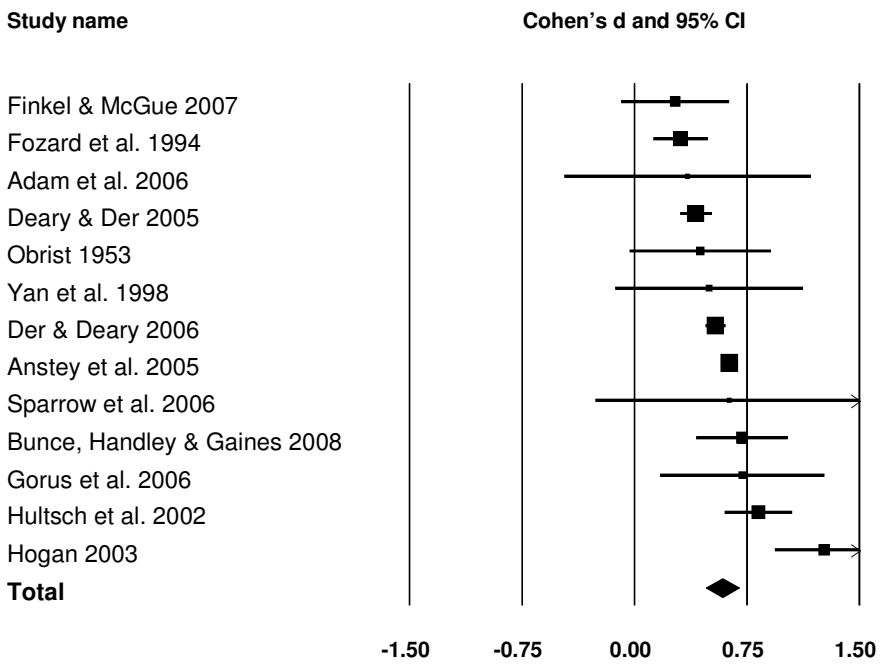

Figure 2.2. Effect sizes for old versus young comparisons in raw simple reaction time intra-individual variability studies

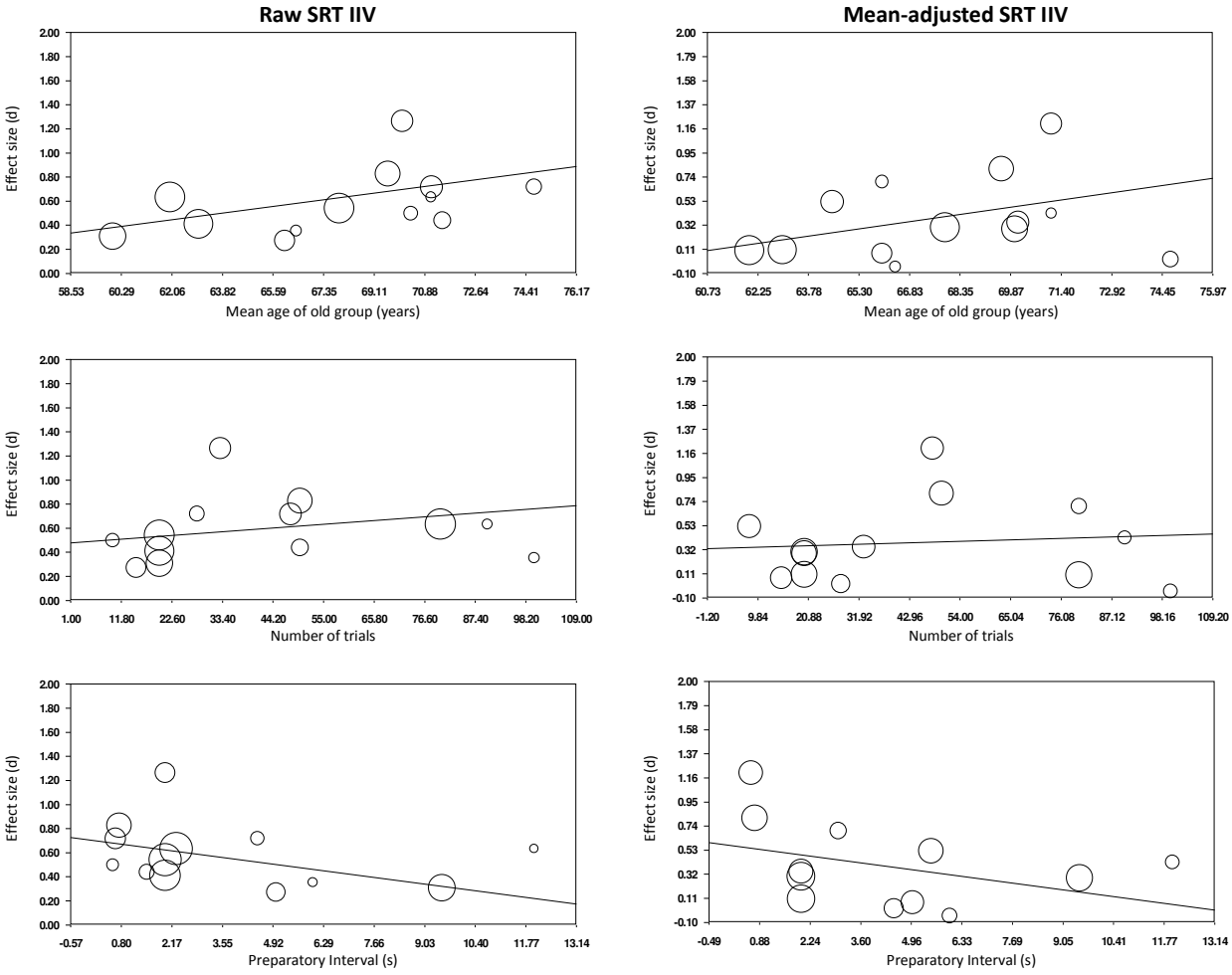

Figure 2.3. Scatterplots of effect sizes for old-young differences in raw (left panel) and mean-adjusted (right panel) simple reaction time intra-individual variability (SRT IIV) and the three covariates: old group age mean (top panel), number of trials (middle panel) and the length of preparatory interval (bottom panel). Circles are proportional to study weights.

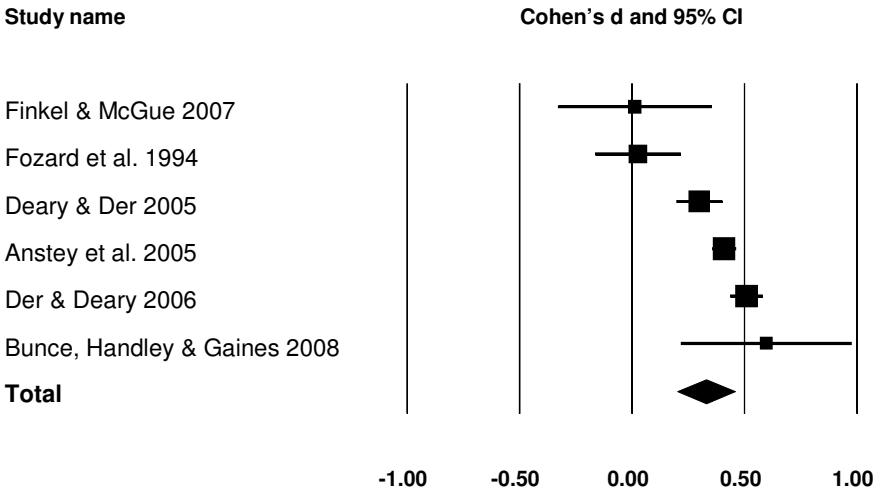

Figure 2.4. Effect sizes for old versus middle-aged comparisons in raw simple reaction time intra-individual variability studies.

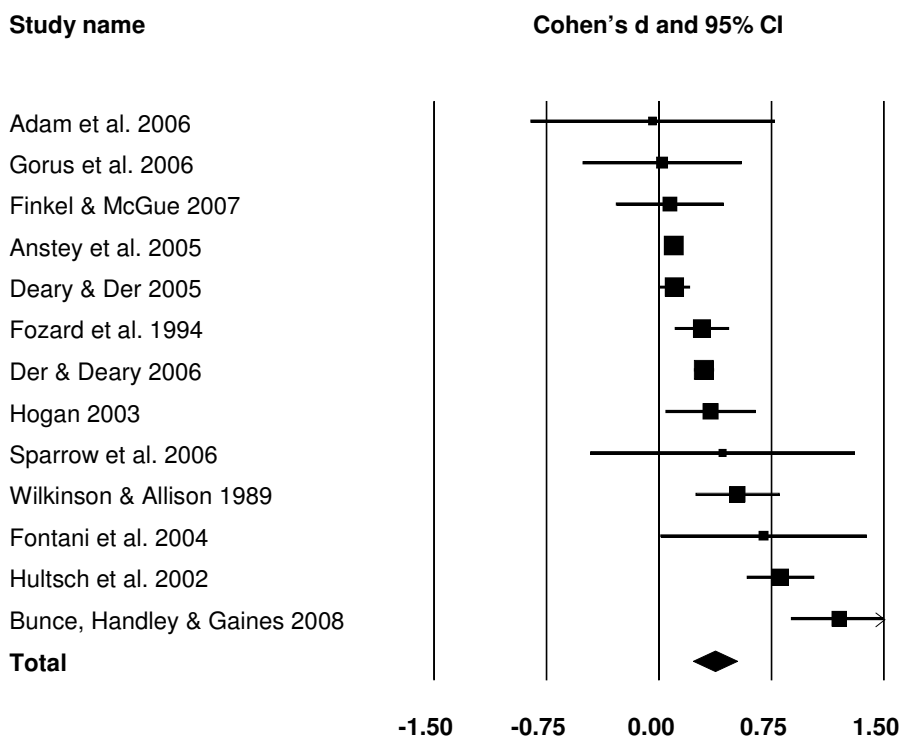

Figure 2.5. Effect sizes for old versus young comparisons in mean-adjusted simple reaction time intra-individual variability studies

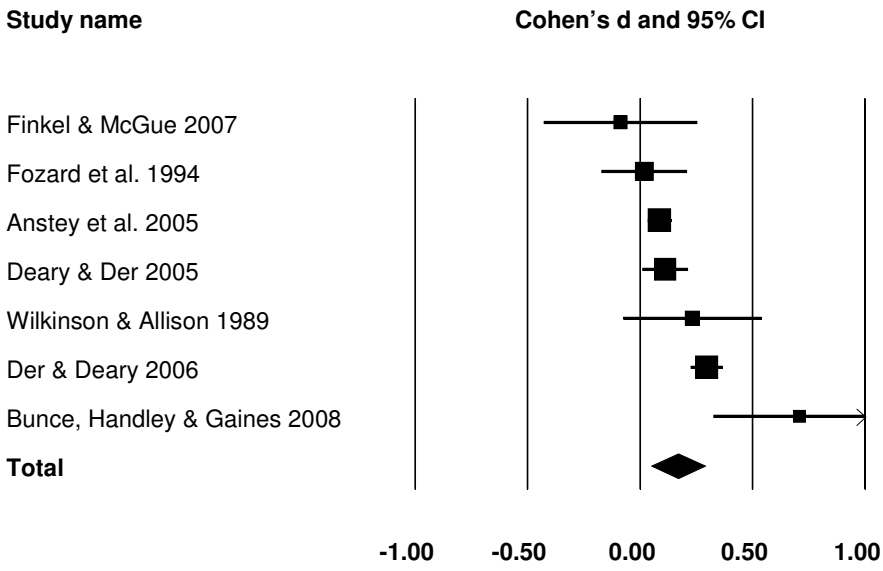

Figure 2.6. Effect sizes for old versus middle-aged comparisons in mean-adjusted simple reaction time intra-individual variability studies

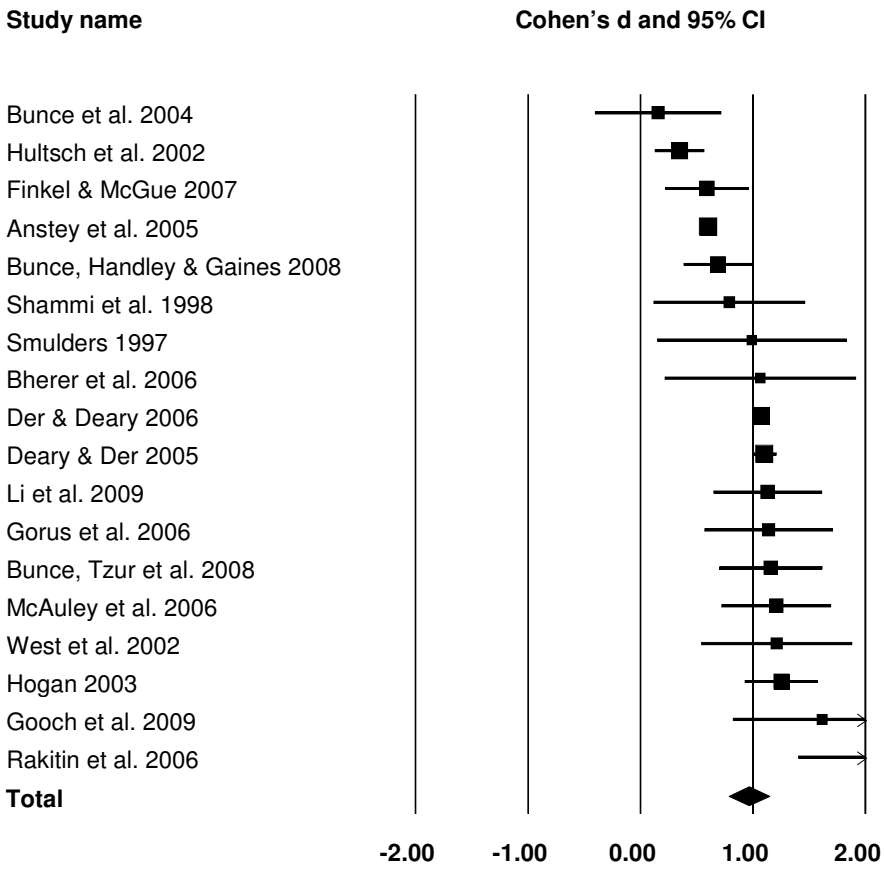

Figure 2.7. Effect sizes for old versus young comparisons in raw choice reaction time intra-individual variability studies

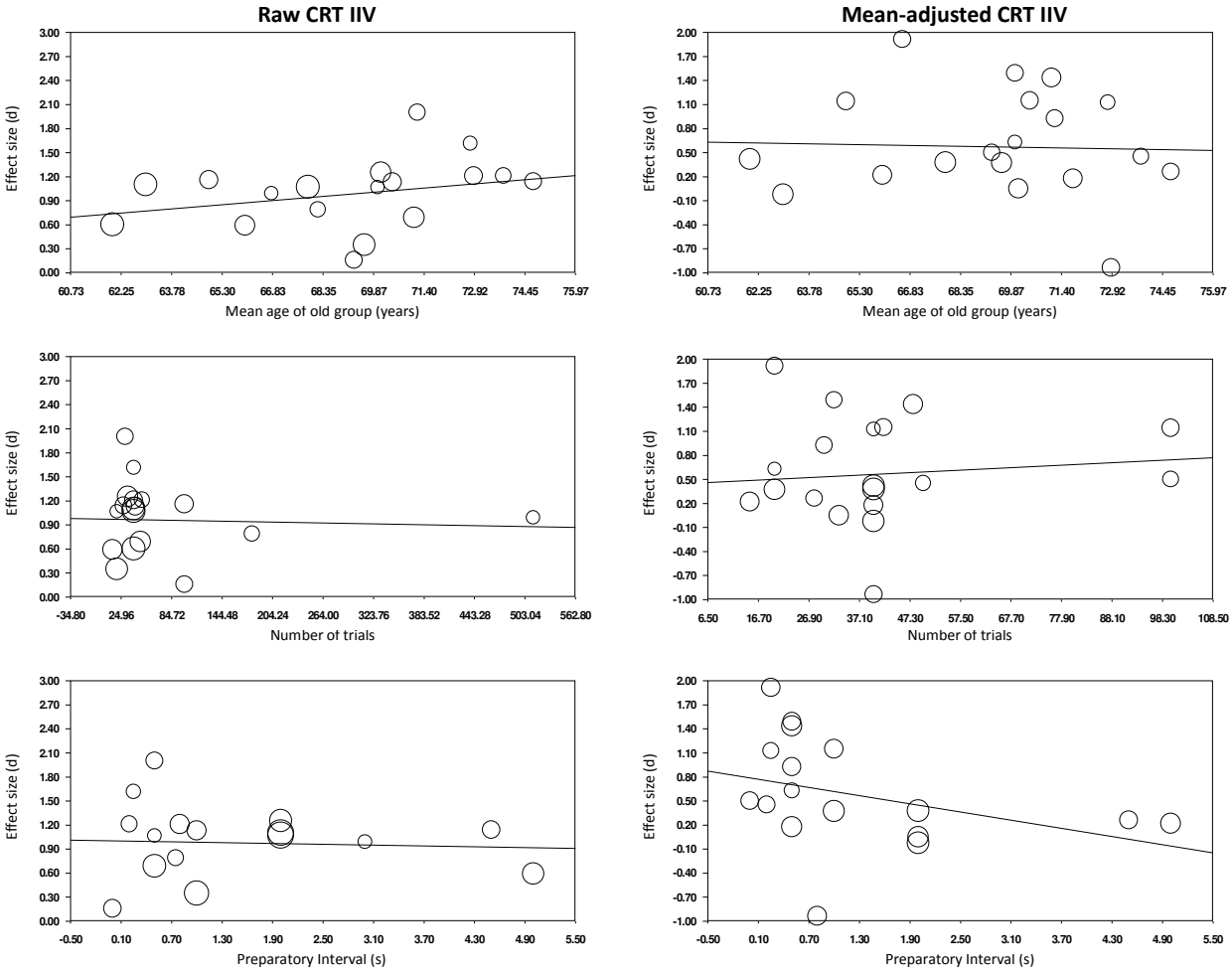

Figure 2.8. Scatterplots of effect sizes for old-young differences in raw (left panel) and mean-adjusted (right panel) choice reaction time intra-individual variability (CRT IIV) and the three covariates: old group age mean (top panel), number of trials (middle panel) and the length of preparatory interval (bottom panel). Circles are proportional to study weights.

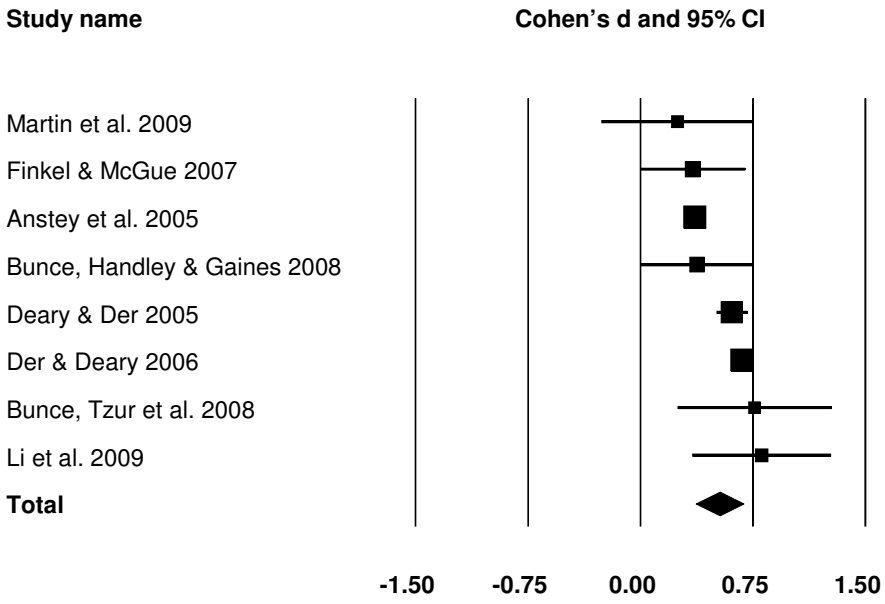

Figure 2.9. Effect sizes for old versus middle-aged comparisons in raw choice reaction time intra-individual variability studies

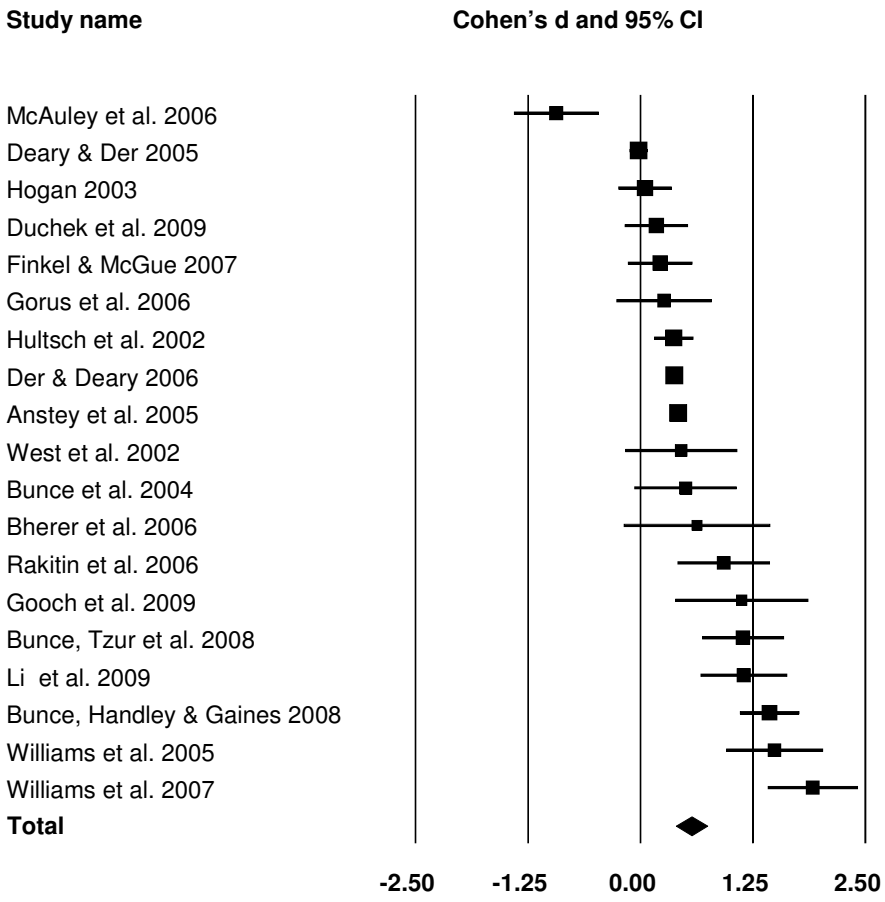

Figure 2.10. Effect sizes for old versus young comparisons in mean-adjusted choice reaction time intra-individual variability studies

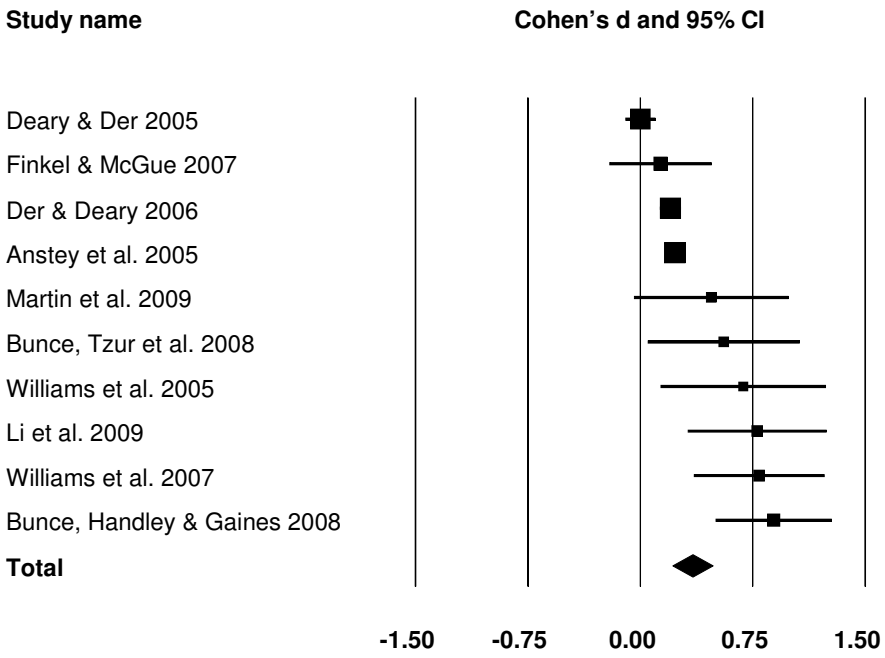

Figure 2.11. Effect sizes for old versus middle-aged comparisons in mean-adjusted choice reaction time intra-individual variability studies

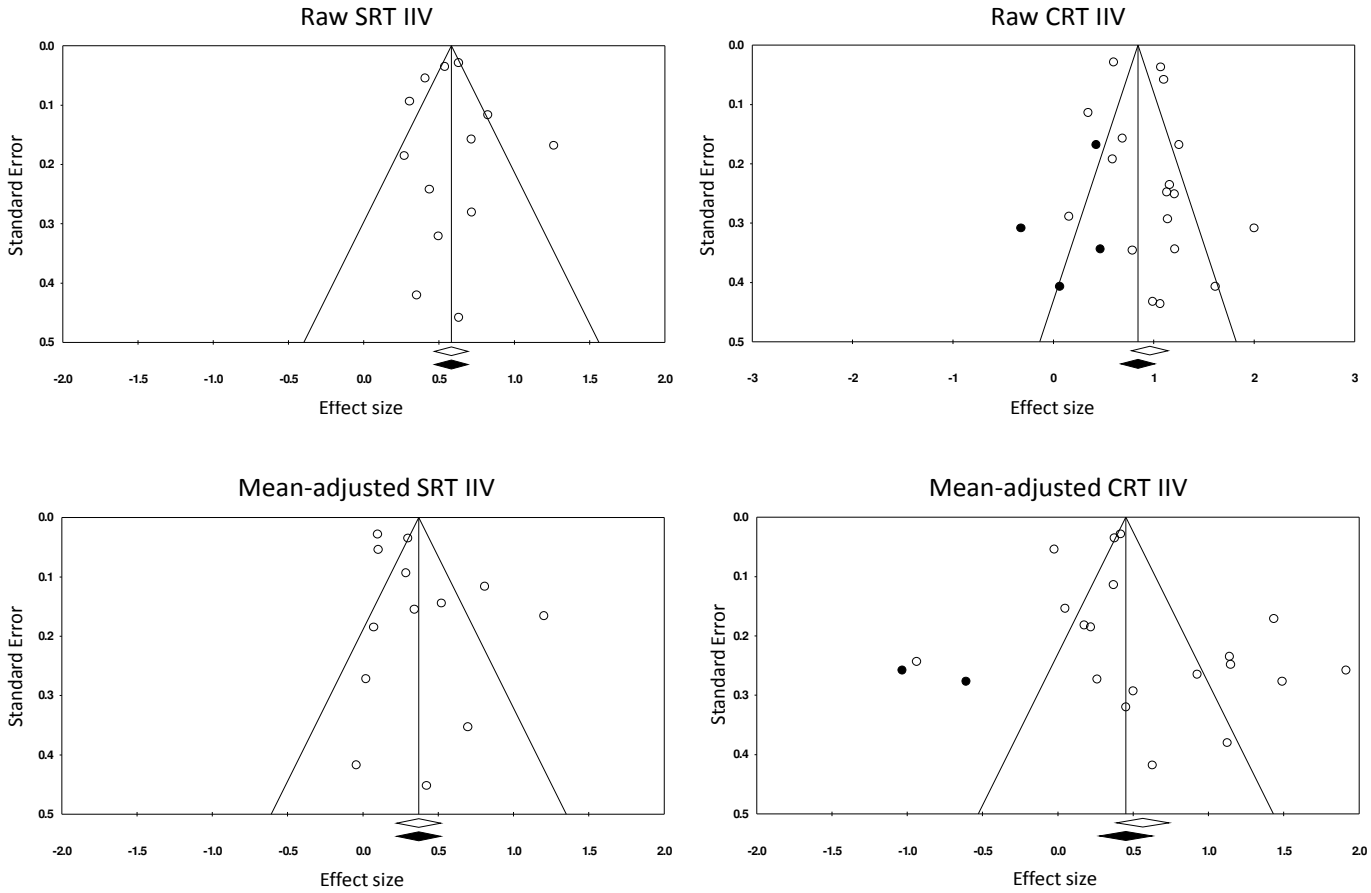

Figure 2.12. Funnel plots showing little evidence of publication bias of raw or mean-adjusted simple reaction time intra-individual variability studies (left panel), with some bias apparent among studies of raw and mean-adjusted choice reaction time intra-individual variability (right panel). Open symbols = actual studies, filled symbols = "missing" data points identified by the trim and fill procedure (Duval & Tweedie, 2000a,b).

Chapter 3 -
Sex differences in reaction time mean and intra-individual variability across the lifespan

The relationship between age and reaction time (RT) has been widely researched and is now well-established. RT follows an approximation of a U-shape through the lifespan: RTs are slowest in childhood, become faster until early adulthood and then decline into older ages. A similar pattern has been found for intra-individual variability (IIV) in RT (Williams, Hultsch, Strauss, Hunter, & Tannock, 2005; Williams, Strauss, Hultsch, & Hunter, 2007). The effect of sex on RT has also been the subject of much research, with females often reported to have greater RT IIV than males (e.g., Der & Deary, 2006). However, these issues are far from resolved: the shape of the relationship with age in different RT tasks is not well understood, the evidence regarding sex differences is inconsistent and—crucially—there is a shortage of large studies across a wide age range, especially including children (Der & Deary, 2006). A pattern of sex differences was found by Deary and Der (2005a), whereby females had more variable RTs than males at older ages, but not in teens or early 20s. They suggested that sex differences in RT IIV may not be apparent in children and only occur post-puberty. The present study set out to test this hypothesis by investigating sex and age effects on simple and choice reaction time (SRT and CRT, respectively) mean and variability. This was done on a large sample including children, adolescents and adults up to the age of 75. I shall start this chapter with presenting a review of the existing literature describing the effects of age and sex on RT mean and IIV.

Reaction time and age

In general, RTs become faster with age in childhood (Gilbert, 1894; Goodenough, 1935) and slower with increasing age in adulthood (Deary & Der, 2005a; Der & Deary, 2006; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994). Koga and Morant (1923) were among the first researchers to investigate changes in mean RT through the lifespan. Using a subset of Galton's SRT data of 3,379 males, they found that auditory and visual RTs were longest in childhood, decreased to minimum in the early 20s and slowly increased with age throughout the adult age ranges. Although some concerns have been expressed about the sample's representativeness (the vast majority of participants in Galton's samples belonged to higher social classes or were students, and they paid to have their measurements taken) and about the RT data reliability (it appears only a single measurement was taken for visual and auditory RTs; Johnson et al., 1985), the findings were later replicated, with many studies finding a U-shaped relationship between age and RT. For example, Wilkinson and Allison (1989) analysed data from 5,325 individuals, aged from under 10 to over 70, who visited an exhibition at the Science Museum in London. They found that, when participants were split into age groups, fastest RTs were achieved by the 20 to 29-year-olds (although not significantly different from 30 to 39-year-olds); RTs were slowest in under 10-year-olds, followed by over-70-year-olds.

The general shape of the relationship has more recently been confirmed on a large and representative sample (though not including children) by Der and Deary (2006). Their investigations from that, and an earlier study (Deary & Der, 2005a), further suggested that the point of inflection of the curve—when it begins noticeably to slow in adulthood—may be different depending on the RT task used. Deary and Der (2005a) investigated cross-sectional and longitudinal effects of age on RT in data from the Scottish Twenty-07 study. There were three cohorts aged about 16, 36 and 56 at the first wave of data collection, with longitudinal data available from an 8-year follow-up. Cross-sectional data showed that both SRT and 4-CRT increased with age, i.e. the youngest cohort was faster than the middle cohort, which in turn was faster than the oldest cohort. However, longitudinal analysis

showed that, over the 8-year period, SRT increased (most notably in the oldest cohort), while CRT actually decreased in the youngest cohort (that is, from 16 to 24), remained relatively stable in the middle one, and increased in the oldest cohort. This difference has been demonstrated more clearly on the UK Health and Lifestyle Survey (HALS) sample, with SRT and CRT data available from 7,130 adults representing all ages between 18 and 94 (Der & Deary, 2006). The shape of the age-RT relationship was different for the two tasks; SRT remained stable until the middle ages, whereas CRT showed slowing throughout the age range. Therefore, it seems that, although the general patterns of initial decrease, stability, and then increase in RT with age holds for both SRT and CRT, the critical ages at which the trends change differ between the two tasks.

Reaction time intra-individual variability and age

The effects of age on intra-individual variability in RT have also attracted research interest. Changes in variability in older ages have received a particularly large amount of research attention and a chapter of this thesis is devoted to a systematic review of literature on RT variability and age in adults (see Chapter 2).

Wilkinson and Allison (1989) carried out one of the first studies to investigate RT intra-individual variability in a large sample and across a wide age range. They found that greatest variability was observed among young children and the oldest age group, whereas least variability was observed in the second and third decades. The finding of Wilkinson and Allison has been replicated by Williams and his colleagues (Williams et al., 2005, 2007) who statistically modelled age-RT intra-individual standard deviation (RT ISD) relationship on a 2-CRT task. They found that RT ISD is a quadratic function of age, confirming the observed pattern of RT ISD decreasing in childhood and adolescence and increasing through adulthood.

There is evidence that RT intra-individual variability continues to increase into very old ages. For example, Anstey (1999) found that there was a positive correlation between age (ranging from 60 to 90) and ISD on three out of four RT tasks (visual SRT and visual and auditory CRT). However, Maddox and Douglass (1974) failed to find a compelling evidence of a longitudinal increase in RT

variability in a smaller sample with a comparable age range drawn as part of Duke Longitudinal Studies.

It appears that, like mean RT, intra-individual variability seems to follow a U-shape. However, again Der and Deary suggested that the age at which the decline begins is different for SRT and CRT tasks. In the Twenty-07 data, intra-individual variability in SRT tended to rise over the 8-year follow-up period, whereas CRT ISD decreased in the young adulthood and increased in the older age (Deary & Der, 2005a). The HALS data also showed a differential trajectory of RTs across ages (Der & Deary, 2006): CRT ISD started to increase at younger ages than did SRT ISD. There was actually an indication that SRT ISD *decreased* until late 30s before it started increasing.

Sex differences in reaction time

It is often reported that men have faster RTs than women. This effect has been found in a number of samples from university students aged between 18 to 25 years (Reed, Vernon, & Johnson, 2004), to representative samples of middle-aged and older adults (Christensen, Mackinnon, Korten, & Jorn, 2001; Lahtela, Niemi, & Kuusela, 1985). Gilbert (1894) found the same pattern in SRT in school children, although the effect was not clear when a 2-way CRT was considered. Goodenough (1935) found that, already at age 3, boys responded faster than girls. However, null findings (Bunce, Tzur, Ramchurn, Gain, & Bond, 2008; Kalb, Jansen, Reulbach, & Kalb, 2004) and even the opposite pattern (e.g. Fairweather & Hutt, 1972) have also been reported. Landauer, Armstrong and Digwood (1980) found that, when RT is separated into decision and movement components, females outperform males on the former, while males outperform females on the latter. These two effects were thought to counteract each other so that no sex difference in the overall RT was found. Although samples used in these studies are not all representative of the wider population, there is no reason to suspect systematic bias between the sexes in sample selection.

Recently, interest in RT has centred not only on mean response time, but also on the inconsistency of an individual's responses across trials within a test (i.e., the

IIV). Only a few studies to date have investigated sex differences in RT IIV and they suggest that adult females are less consistent than males (Deary & Der, 2005a; Der & Deary, 2006; Fozard et al., 1994; Goodenough, 1935). Of note is that this trend did not occur in Eckert and Eichorn's (1977) childhood data.

Given inconsistencies in the literature, it remains unclear whether males and females differ in their RT performance in terms of mean or variability. One factor that may explain some of the discrepancies in the research findings is the different ages of groups that have been studied. Although it is now well-established that, in general, RTs become shorter with age in childhood and longer with increasing age in adulthood (with a similar pattern for RT IIV), few investigators examined the developmental patterns of RT separately for males and females. Those who did found some indication that these may, in fact, differ between the sexes (e.g. Fairweather & Hutt 1972; Noble, Baker, & Jones, 1964). Therefore, regardless of whether there is an overall difference in RT between the sexes, there might or might not also be variations in the pattern of any differences across the lifespan.

Sex differences in reaction time through the lifespan

Among the earliest studies that had investigated RT in males and females across a wide age range (4-60 years) was the one carried out by Bellis (1933). Although no statistical tests were performed, in all age groups tested, auditory and visual SRTs of males were faster than those of females[1]. Bellis found that the male advantage was particularly marked in the two youngest groups (4-10 and 11-20 years old). This finding may, however, result from inadequate age matching of males and females in these groups (Goodenough, 1935). From the investigation of mean ages it is clear that females were on average 1.9 years younger than males in the youngest group and 1.7 years younger in the second youngest group. Given the rapid improvement in RT in childhood, it seems most likely that Bellis' finding is spurious.

Eckert and Eichorn (1977) report on the effects of age and sex on auditory SRT in children. They presented results from two longitudinal studies of

[1] When t-tests were performed on the summaries reported in Bellis' article, sex differences in the two youngest groups were significant at $p < .001$. Overall, 8 out of 12 comparisons were significant.

development: the Berkeley Growth Study (56 subjects followed-up from 4.5 to 11.5 years) and the Oakland Growth Study (164 subjects followed up from 10 to 16 years). An improvement in RT with age was seen in both groups, with a trend toward a decrease in the magnitude of improvement; that is, the improvement from year to year was greater at younger than at older ages. There was some evidence of boys responding faster than girls when all the ages were pooled together, but no significant sex difference when only a single age level was considered. For girls, RTs became shorter until 15, whereas at 16 they were similar to those obtained at age 12; i.e., they became longer. Boys, on the other hand, seemed to improve from 10 through to 16. This trend suggests a differential developmental trajectory of RT between the sexes. However, due to a cut-off at age 16, it cannot be verified whether girls' poorer performance was a sign of the start of a decline, or whether it was simply an irregularity in the data.

Fairweather and Hutt (1972) showed that variations in the magnitude of sex differences can be found even in a relatively narrow age range and already in childhood (6-11 years). However, they found that girls were *faster* than boys on 2-, 4- and 8-CRT tests; older children were faster than younger ones. However, the authors noted that at the age of 7.5 years, boys and girls were performing at a similar level, following which the RTs of girls decreased, while those of boys remained relatively unchanged. The difference between the sexes was greatest at 8.5 years old. Following this, RTs of boys began to decrease and, at age 10, almost reached the girls' level of performance. A similar pattern was observed in SRT. Fairweather and Hutt provided an investigation of sex and age effects in different tasks and demonstrated clearly that both age and sex effects are larger in more demanding tasks. However, their study was hindered by a small sample size (n=36). The proposed differential developmental curves of boys' and girls' CRTs were based on the interpretation of the means and, by the authors' own admission, the sample size did not allow robust statistical analysis. Therefore, even if the described trends are genuine, the precision of Fairweather and Hutt's age estimates for the main changes would need to be verified.

Only a few studies have investigated sex and age effects on RT in large samples (n > 500), and even fewer in representative ones. Noble et al. (1964) studied

mean CRTs of 600 people aged 8 to 87 years. They found that males had faster RTs than females and that, for both sexes, RT followed the expected pattern of improvement from childhood to early adulthood, a period of stability, and a subsequent steady decline. However, they also found a significant interaction between sex and age and, from visual inspection of means, concluded the following: males and females have similar CRTs until the age of about 16; from that age females start to become slower, whereas men's CRTs continue to improve until about age 20. The decline parts of the slopes are different for men and women, with the result that the gap between the sexes narrows, with men's performance falling below that of women in the oldest age group. From Noble et al.'s data, it appears that any male advantage might arise from their prolonged period of improvement in RT relative to girls, which enables them to achieve a higher level of performance before they start declining with age. The paper by Noble et al. raises some interesting possibilities. However, many of the observations have not been subject to statistical testing and are based merely on visual inspection of graphs.

Several more recent studies provide evidence based on a more rigorous treatment of data (Fozard et al., 1994; Deary & Der, 2005a, Der & Deary, 2006), but none have included children in their samples. Fozard et al. analysed data on 1,265 adult participants in the Baltimore Longitudinal Study of Aging. RTs were available from two tasks: auditory SRT and a go/no-go task, requiring a response to a high tone and inhibiting a response to a low tone. Across ages 20 to 90, both RTs became slower, but the change was larger for RTs from the go/no-go task than SRTs. There was a significant sex difference in SRT and go/no-go tasks, with males responding faster, but there was no interaction between sex and age indicating that the amount of change in RT with age in both sexes was comparable. Fozard et al. also considered intra-individual variability in the two RT tasks and found that it too increased with age, but a sex difference was only found in the more complex go/no-go task. Again, there was no sex by age interaction.

In the previously-mentioned study on the HALS sample, Der and Deary (2006) found significant sex differences in RT, with women's RTs being slower and more variable than those of men. The sex difference was most marked in CRT intra-individual standard deviation (ISD). Unlike in Fozard et al.'s (1994) data, significant

age by sex interactions were found in SRT mean and SRT and CRT ISD, suggesting that the magnitude of sex differences varies with age.

Deary and Der (2005a) investigated cross-sectional and longitudinal effects of age on RT in data from the West of Scotland Twenty-07 study. They found that women were generally more variable, but, for SRT, this was only significant in the mid-40s, whereas CRT ISD was greater in women in the two older cohorts, but not significantly so in the youngest. The pattern of results suggested that sex differences in RT variability are observed among older rather than younger adults. Deary and Der proposed an explanation for both the existence of the sex difference and its manifestation in the middle-aged to older groups (and absence in the younger group), in terms of hormonal effects. As they pointed out, estrogens receptors are present in many brain areas known to play a role in information processing, motor performance, attention, as well as in systems thought to affect variability in information processing. Given that estrogens have different effects in female and male brains (McEwen, 2001), they suggested that the post-pubertal accumulation of sex hormones may lead, with time, to differentiation of responding between the genders. If this hypothesis is correct, then sex differences in RT intra-individual variability should be found in adults, but not in children or young teenagers. This hypothesis was tested in the present study. Another explanation of the effect, in terms of differences in trial-to-trial performance between males and females, was proposed by Reimers and Maylor (2005). This theory will be described in detail and tested in Chapter 5.

The relationship between mean reaction time and intra-individual variability

The correlation between mean RT and RT IIV, its implications, and different methods of adjusting for RT mean when studying RT IIV have been discussed in Chapter 1. The effects using these different approaches to studying RT IIV have on the findings in terms of age and sex are presented next.

Whether mean RT is controlled when studying intra-individual variability, and how it is achieved, can have important implications for the study results. For

example, Fozard et al. (1994) found that, when intra-individual variability/mean ratio was used as the dependent variable, age effects found in an unadjusted variability measure remained significant in both SRT and go/no-go tasks, but the sex difference was no longer significant. Der and Deary (2006), by contrast, found that the sex difference in CRT ISD remained significant even when CRT ISD was modelled while controlling for mean RT, although the age effect did not. When CRT CV was used as an outcome, both sex and age effects remained significant. For SRT ISD, the effect of sex was removed by controlling for the SRT mean in either way, whereas age remained significant for both. Deary and Der (2005a) found that, when the regression method is used to control for the mean, the effects of age (but not of sex) on RT intra-individual variability are largely removed.

Although it cannot yet be determined which effects are resilient to controlling for RT mean and which are not, the findings of Fozard et al. (1994) and of Der and Deary (Deary & Der, 2005a; Der & Deary, 2006) suggest that the effects of age and sex on RT variability may have different origins and mechanisms. In the present study, two methods of controlling for RT mean are used and their results are compared in terms of age and sex effect.

The present study

The present study uses SRT and CRT data from a sample that included both children and adults and can, therefore, answer some of the outstanding questions following Deary and Der's studies (Deary & Der, 2005a, Der & Deary, 2006). In this study the shape of the age–mean RT and age–RT variability relationships were investigated across an age range including children and adults that include older individuals. It was expected that age would have an effect on the RT measures (means and ISDs) and that the relationship between age and RT measures would be negative in children and adolescents and positive in adults. In terms of sex differences, it was hypothesised that greater female RT intra-individual variability would be found among adults but not children. Finally, to shed light on the issue of controlling for the mean RT when investigating RT ISD, in this study, both RT ISD

and RT CV are used as outcomes. In addition, RT ISD was modelled with and without controlling for mean RT.

Method

Participants

Participants were visitors to Science Festivals held in Cambridge, Edinburgh and Cheltenham in 2008, and Edinburgh in 2009. Because the festivals were aimed at children, participants under 18 years old comprised the bulk of our sample. However, the accompanying parents and guardians were also invited to participate, so that the overall age range was 3 to 87 years. Over 2,400 participants attempted the RT task, and 2,392 records from the RT apparatus were successfully matched with participant information. Ninety five records were incomplete (33 individuals did not complete both RT tasks, and the age of 62 people was not available). Records from participants who made 8 or more errors (n = 131), were also excluded from the analyses, because a 20% error rate indicates that an individual was unable or unwilling to follow the instructions. In addition, participants younger than 4 (n = 5) and older than 75 years (n = 4) were excluded, as there were insufficient numbers in these age groups to estimate reliably their respective age groups' levels of performance. A further 34 participants were excluded following individual trial RT data trimming, which is described in detail in the "data preparation" section to follow. Finally, as most of the analyses here are concerned with gender differences, records in which gender was not reported were also excluded to ascertain equivalent samples for models including and excluding sex. These exclusions left a working sample of 1,994 (848 males, 1,146 females), ranging in age from 4 to 75 years, with a mean = 15.35, median = 10.34, and SD = 13.44 (see Table 3.1 for numbers of males and females broken down by age groups).

Because the sample consisted of visitors to Science Festivals, who had to pay an entrance fee, it was expected that they would mainly represent professional social classes and relatively affluent backgrounds. Index of Multiple Deprivation (IMD) scores were used as a proxy for socio-economic status of our sample. Consenting

participants from within the UK provided full postcodes which were later linked to their respective data zone if in Scotland or lower super output area (LSOA) if in England (Office for National Statistics, 2006). Here, only Scotland and England are considered, as there were few participants from outside these countries for whom valid postcodes were available (9 from Wales, 5 from Northern Ireland, and 1 from the Channel Islands). Data zones and LSOAs are clusters of neighbouring areas, containing respectively 500-1,000 and 1,000-1,800 residents (Office for National Statistics, n.d.). IMD scores were obtained for Scottish data zones (Scottish Government, 2006) and English LSOAs (Department of Communities and Local Government, 2007) in which the participants resided. Because the deprivation scores are derived differently for Scotland and England, each nation is considered separately.

The IMD scores obtained were grouped into country-specific quintiles (with lower quintiles representing more deprivation). Figure 3.1 presents the relative numbers of participants from each quintile by gender and age (child/adult). As expected, the most deprived areas were under-represented and the least deprived areas were over-represented in the sample. About half of males, females, children and adults from Scotland and England lived in the area considered to be among the 20% least deprived areas in their respective country. About 25% of the Scottish and close to 30% of the English sample were made up by individuals living in the three most deprived quintiles. It appears that the males and females as well as children and adults come from similar backgrounds. Importantly, there was no significant sex (male/female) or age group (child/adult) difference in the overall IMD scores for participants from England (sex: $t(465) = 1.589$, $p = .113$; age: $t(478) = -1.003$, $p = .316$) or Scotland (sex: $t(1152) = -.100$, $p = .920$; age group: $t(1152) = -.148$, $p = .883$).

Reaction time tests and procedures

RT testing was one of the activities offered by the UK's Medical Research Council at Science Festivals held in Cambridge, Cheltenham and Edinburgh. The

study was approved by the Glasgow University Faculty of Law, Business, and Social Sciences Ethics Committee.

Simple and 4-CRTs were measured using a RT testing device, which was an upgraded version of apparatus designed for and used in the UK HALS study (Cox, Huppert, & Whichelow, 1993). The device had an LCD screen for the presentation of stimuli and five response buttons labelled 0 to 4 (see Figure 3.2 for illustration if the device). Buttons 1 and 2 were placed on the left side of the box and were to be operated with the middle and index finger of a left hand; buttons 3 and 4 were placed on the right and were to be operated with the index and middle fingers of a right hand; and a central 0 button (located between buttons 2 and 3) was used in SRT with the preferred hand. In the SRT task a "0" (zero) would appear on the display screen and participants were required to press the 0 button as soon as the stimulus appeared; there were 8 practice and 20 test trials. In the CRT task one of the numbers from 1 to 4 would appear and the participants were required to press the corresponding response button. Eight practice trials were given prior to 40 test trials. The tasks were always administered in the same order; that is, SRT followed by CRT, with each task preceded by its respective practice session.

The same sequence of stimuli for CRT was used for all participants. The interval between a response and the onset of the next stimulus varied between trials and ranged from 1 to 3s. The pattern of the intervals was kept constant across participants in both SRT and CRT tasks. For both tasks, response latency (time elapsed between the stimulus onset and pressing of a button) was recorded for each trial with millisecond accuracy. Means and standard deviations were calculated for each individual in both tasks and, in the case of CRT, separately for correct and incorrect responses.

Potential participants were briefed on the task prior to participating and, if aged under 16 years, the consent of their parents or guardians was obtained. Between 1 and 6 individuals could participate at any one time, with most tests being carried out in groups of 3 or 4.

Data preparation

All pre-presses, that is, responses made before the appearance of the stimulus, were identified and removed. There were 391 trials with pre-presses out of a total of 43,140 in SRT and 49 out of 86,280 in CRT. All SRTs faster than 100 ms (n = 202) and CRTs faster than 150 ms (n = 42) were also removed, because they indicate accidental responses. To determine the upper cut-off value for RTs, a first minimal trimming of the most extreme responses was performed, in which all SRTs above 3,000 ms and CRTs above 5,000 ms were trimmed, resulting in the exclusion of 37 SRT and 29 CRT trials.

Following these preliminary trimmings, age-specific means and standard deviations were calculated, and any values falling above 5 SDs of their respective mean were also removed. Due to a greater number of children than adults, age groups used for the calculations were 1-year bands for ages up to 14 years, and 5-year bands thereafter (see Table 3.1). This procedure resulted in the removal of 235 individual trials for SRT and 198 for CRT. The cut-off values used for trimming were less stringent than in many other studies using RT data, in which the lower cut-off of 150ms is used regardless of the task and the upper cut-off is determined by 3 SDs above the mean (e.g. Hultsch, MacDonald, & Dixon, 2002, Williams et al., 2005). It was my intention to perform only minimal trimming of most aberrant responses, because it has been suggested that an increased number of very slow responses, especially among older adults, may be a genuine phenomenon (e.g., West, Murphy, Armilio, Craik, & Stuss, 2002).

Next, cases in which more than 25% of an individual's trials were removed following the above exclusions (in either RT task) were identified. These cases were removed completely because a high frequency of pre-presses and/or very long and short responses may indicate a participant's lack of ability or effort to carry out the task properly. There were 25 SRT tests with more than 5 invalid trials, whereas all CRT tests had fewer than 10 invalid RTs and no records were excluded on that basis. In total, 2.1% of individual SRT and 0.4% of CRT trials were removed. Following these exclusions, means and ISDs were re-calculated for each test and each participant from the trials deemed valid. The final step was to exclude nine

participants whose mean RT or RT ISDs fell beyond 3 inter-quartile ranges above their age group-specific 75th percentile.

To obtain a simple measure of RT intra-individual variability which is independent of the mean RT, the coefficient of variability (RT CV) was calculated from the trimmed RT data (CV = RT ISD/intra-individual mean RT*100).

Statistical analyses

Analyses were performed using SAS software (version 9.2, SAS Institute). Polynomial regression models were fitted to all RT measures using PROC GLM. As mentioned before, the effect of mean RT on RT ISD was controlled in two ways: by using RT CV as an outcome and by including mean RT as a covariate in the RT ISD models. Therefore, for both SRT and CRT tests, four models were fitted: one for mean RT, one for RT CV, and two models for RT ISD – one with and one without adjusting for RT mean. This analysis was initially performed on the whole sample and later repeated on only participants below the age of 18 and separately on the adult portion of the sample (aged 18 and above). For these analyses, age was centred on the relevant mean (15.35 years for the whole sample, 9.73 for under 18s, and 42.03 for individuals aged 18 and over).

Because residuals from the preliminary models fitted to the data were not normally distributed and were heteroscedastic across the age groups, the data were transformed prior to the final analyses. Following the practice of Der and Deary (2006), Box-Cox transformation was used (Box & Cox, 1964) to normalise the distribution and to reduce the heteroscedasticity. This procedure selects a transformation parameter (λ) which is optimal for the model to be fitted. If λ is different from zero the data are transformed using the following equation: $y = (y^{\lambda} - 1)/\lambda$; if the selected λ equals zero, then a simple log transformation is performed. The values for λ were determined using the PROC TRANSREG separately for children/teenagers (< 18) and adults (aged 18 or over) and the whole sample; the selected values are presented in Table 3.2.

For graphs, predicted values from the final models were scored using PROC PLM, with number of errors, mean SRT and mean CRT set to their respective means, where necessary.

Results

Means and ISDs for SRT and CRT of males and females in three broad age groups (children/teenagers, young to middle-age adults, and older adults) are presented in Table 3.3. The RT means and ISDs are generally largest in the youngest group, intermediate in the oldest group, and smallest in the middle group. Inter-individual (between-subject) variability in mean RT and RT ISD followed a similar pattern. The observed differences in variances were tested for significance using Levene's test. Children and adolescents were more diverse in their RT means and ISDs than either of the adult groups (all ps $< .05$). The middle and oldest groups significantly differed only in the variances of their mean SRT scores ($p < .001$).

Means and standard deviations of the RT measures in narrower age bands (1-year bands up to the age of 14, and 5-year bands thereafter) are presented in Figure 3.3. RTs became markedly shorter and less variable with age in childhood and adolescence. Age-related increases in both mean and variability in adulthood were much less marked. Whereas SRT seemed to remain stable until the middle adulthood and showed a slow increase, CRT appeared to deteriorate throughout the adult age range. CRT IIV showed a slight increase throughout the adulthood and older ages, whereas there was little evidence of increased variability with advancing age in SRT. In all four RT measures there appeared to be a plateau in the deterioration, or even a slight improvement in the performance of the oldest individuals.

Modelling was performed on the transformed data (as described in the "statistical analyses" section above), including polynomial terms in age, sex and their interactions entered as potential predictors. For CRT variables the number of errors was also included. Final models were obtained by stepwise elimination of non-significant terms.

The results will be presented under three sections, starting with the examination of age and sex effects on reaction time through the lifespan, followed by the same investigation in children and teenagers (under 18), and then in adults.

Age and sex effects on reaction time through the lifespan

Full models of RT fitted to the entire sample, with the relevant statistics, are given in Table 3.4. Predicted values obtained from these models for SRT and CRT means, ISDs, CVs and ISDs adjusted for RT mean, have been back-transformed to their original units and are shown in Figure 3.4. All measures showed a steep decrease in childhood until early 20s followed by a much less rapid increase until old age. A small decrease in SRT, SRT ISD, CRT ISD and CRT CV was observed in the oldest group.

Predicted values of mean SRT were consistently higher for females than for males, although the difference was modest; for mean CRT, there was an interaction between the fifth-order polynomial of age and sex. Sex differences in CRTs showed an irregular pattern, with females appearing to have slightly slower CRTs, but only from the late teens onwards; throughout the adult age range the magnitude of the difference varied, with no clear differences at some ages.

There were significant sex by age interactions in both SRT ISD and CRT ISD: there was no clear sex difference in childhood and a slight male advantage only appeared in the early 20s. The increase in both SRT ISD and CRT ISD appeared to start earlier in women than in men, whose RT ISDs remain relatively stable well into the middle ages.

SRT CV and CRT CV had rather different patterns. The change in SRT CV with age was bi-phasic. A clear improvement was apparent until the 20s, as with other measures. Then there was a slight increase in SRT CV among the middle-aged subjects, followed by a small drop before an accelerated increase in the late 60s. The trend was similar for males and females, but men seemed to decline less and improve more than females. Girls appeared to have slightly lower SRT CVs than boys. CRT CV resembled the pattern of CRT ISD. The patterns were different for males and females: whereas females' CVs decreased until the 20s and increased after that age,

males' CVs seemed to continue decreasing until much later (early 40s) and only then decline. Females had slightly lower CRT CVs between the ages of about 10-20 and consistently higher CV from the early 20s on.

When RT intra-individual variability was modelled controlling for RT mean, age remained significant predictor for both SRT ISD and CRT ISD; however, the direction of the relationship was reversed. In other words, after the effect of RT mean was controlled, RT intra-individual variability actually decreased with age. For SRT ISD there was a main effect of sex, with females apparently less variable. There was a significant sex by age interaction in CRT ISD: females appeared to be less variable than males in the early years and teens and more variable from the early 20s onwards.

To directly test the hypothesis that sex differences will be apparent in the adulthood but not in the younger ages, the models were re-fitted separately on the younger (<18) and older (18 and above) portions of the sample.

Age and sex effects in childhood and adolescence

Details of the models fitted to RT data of children and teenagers are presented in Table 3.5. Figure 3.5 shows the observed means and standard errors of RT means and ISDs for each age group (left panel) and back-transformed predicted values from the models (right panel). In participants younger than 18, there was a curvilinear decrease in mean SRT, mean CRT, SRT ISD and CRT ISD, with accelerated decrease noted in the earliest ages. Mean SRT and CRT were a quadratic function of age, while SRT ISD and CRT ISD were best described by a cubic function of age. Mean CRT and CRT ISD were a function of the number of errors. The only significant gender effect was found in SRT mean, where males were faster than females.

Models of IIV adjusted for the mean RT are also summarized in Table 3.5 and predicted values from these models are shown in Figure 3.6. With the effects of SRT mean controlled (both by using CV and including mean SRT as a covariate), the effect of age on SRT variability was cubic. There was also an effect of gender on SRT IIV, with females slightly less variable than males. As far as CRT IIV is

concerned, females were again less variable using both methods of controlling for RT mean, and there was a quadratic effect of the number of errors. One notable difference between the two models of mean-adjusted CRT is that when mean RT was controlled by including it in the model, the age effect was not significant, while there was a cubic effect of age on CRT CV. Both SRT ISD and CRT ISD were cubic functions of their respective mean RTs.

Age and sex effects in adulthood

Final models of the adult RTs are summarised in Table 3.6. Figure 3.7 shows the means and standard errors by age group in 5-year bands (left panel) and predicted values transformed back to their original units (right panel). One main observation is that, probably due to the smaller adult sample, the graphs of mean values are less regular and standard errors are larger than was seen in children and teenagers.

In the adult sample, there was a linear increase in SRT with age and males were significantly faster than females. Contrary to expectations, SRT ISD was a function of only sex – females were more variable than males, but there was no change in SRT ISD with age in adults. Curvilinear slowing of mean CRT with age was apparent, whereas CRT ISD increased linearly throughout the adult age range. There was a significant sex difference in CRT ISD, with females more variable than males. This difference appeared to be especially marked in the middle ages; however, the age by sex interaction was not significant.

Predicted values of IIV adjusted for mean RT are presented in Figure 3.8 (see also Table 3.6 for the models statistics). SRT CV did not vary with age or sex, whereas CRT CV was a linear function of age and females were slightly more variable than males. The results were quite different when mean RT was included in the models of ISD. SRT ISD actually decreased with age (very slightly, but significantly); age was not a significant predictor of CRT ISD, which was greater among females and increased with the number of errors made. Both SRT ISD and CRT ISD increased linearly with their respective mean RTs.

Indices of Multiple Deprivation and reaction time measures

As mentioned before, Scottish IMDs are obtained differently to those in England and Wales. Therefore, in order to investigate potential associations between IMD and RT measures, sets of non-parametric correlations were run separately on the Scottish and English portions of the sample (see Table 3.7). There was no clear pattern of the relationships, probably due to the restriction of range of IMDs in the sample. For England, there were no associations between any of the IMD scores and RT measures. For Scotland there were two significant correlations: crime score with CRT ISD, rho=.064, $p<.05$; and crime score with CRT CV, rho=.084, $p<.01$. Both correlations indicated greater variability with greater crime score.

Discussion

The discussion of our findings will be presented in three sections: Sex differences in RT, Age effects, and RT intra-individual variability controlling for RT mean.

Sex differences in reaction time

The key finding of the present study is that, as hypothesised, greater IIV of females in SRT and CRT was observed in adults, but not in children. This finding builds on the work of Deary and Der (Deary & Der, 2005a; Der & Deary, 2006), who investigated sex differences in adults, by adding evidence from a single sample that included children, adolescents and adults.

The pattern of results, with no sex difference in RT variability in under 18s and more variable RTs of adult females, is in line with Deary and Der's (2005a) sex hormone hypothesis. It is plausible that sex hormones, which come into play at puberty and which have different effects in male and female brains (McEwen, 2001), lead to a differentiation of responses between the genders from early adulthood

onwards. However, this explanation remains tentative. It remains to be explored whether hormones are related to RT variability differences between males and females and if so, how their effects come about.

One approach to studying the effects of estrogen on cognition is to compare individuals who undertake hormone replacement therapy (HRT) with those who do not. However, one problem with this approach is that HRT and non-HRT groups may not be comparable and women choosing to undertake HRT may be doing so to alleviate various negative symptoms. Nonetheless, the results from such studies may shed some light on the issue in question. One study which investigated the effects of HRT on intra-individual variability in SRT and CRT (Low, Anstey, Jorm, Christensen, & Rodgers, 2006) was identified. In a representative sample of 1,154 postmenopausal women, HRT interacted with statin use and menopause type. Specifically, among statin users, SRT IIV was greatest in the HRT group and lowest in no-HRT group; the effect was not found among women who did not use statins. When menopause type (natural versus hysterectomy) was considered, differences between HRT groups in CRT variability were found only in women who had hysterectomies, with lower variability in HRT non-users than current users. It is of note that Low et al.'s study had sufficient power to detect even small effect sizes but, despite that, no main effects of HRT were found on SRT or CRT variability. Moreover, the effects that were found were not found consistently in the two RT tasks and they had opposite directions (protective effect of HRT in statin users and detrimental effect in women with hysterectomy). Taken together, these results do not provide a compelling evidence for the effect of estrogen on intra-individual variability or cognition more generally in older women.

Another study which considered differences in intra-individual variability in RT from a memory task between female estrogen users and nonusers also failed to find a significant effect (Wegesin & Stern, 2004). This study was much smaller than that of Low et al. (2006) (n = 15 for estrogen users and n = 16 for non-users) and had little power to produce significant findings.

Given the above, there is little evidence that estrogen supplementation in older women affects IIV in RT. However, it remains to be established whether endogenous levels of estrogen can affect RT variability at younger ages. There are at

least two mechanisms which could lead to sex differences in RT variability in adolescence. One obvious mechanism is the direct effect of circulating hormones on human brain functioning (see McEwen, 2001). However, given that structural changes (such as myelination, synaptogenesis, or synaptic pruning) occur in the brain through adolescence and even into adulthood (Blakemore & Choudhury, 2006), differential effects of sex hormones on these processes could lead to RT differences between the sexes. There is evidence of sex differences in developmental cerebral white and grey matter volume changes through adolescence (De Bellis et al., 2001; Lenroot et al., 2007), and these are thought to reflect myelination and changes in synaptic density. Therefore, brain dimorphism is another mechanism which may underlie sex differences in RT variability.

An alternative explanation could be in terms of cohort effects. Because this study was cross-sectional, it may be that the differences observed between younger and older participants are a reflection of genuine changes in the magnitude of sex differences in RT. A recent meta-analysis (Silverman, 2006) suggested that sex differences diminish over time. Perhaps the youngest participants in our study represent a generation in which sex differences in RT has reduced beyond the level of statistical significance. A longitudinal investigation, in which the same individuals are followed up from childhood through adolescence to adulthood, would shed more light on the issue.

With regard to average speed of responding, females had consistently slower SRTs across the age range, whereas there was no reliable sex difference in CRT mean. The former finding is consistent with a bulk of previous literature, which found men to have faster RTs both in childhood and in adulthood (Gilbert, 1894; Goodenough, 1935; Lahtela et al., 1985). Conversely, the lack of clear sex difference in CRT, especially in adults, opposes many previous results (notably Deary & Der, 2005a, Der & Deary, 2006). However, when our findings are considered in light of Landauer et al.'s (1980) findings, it is not entirely surprising that male advantage in SRT does not hold in CRT. In fact, Landauer et al. also reported no difference in CRT between adult males and females. The two RT tasks used in the present study (SRT and CRT) were very similar to each other in terms of the motor component; that is, the (minimal) movement distance and force required to press a response

button were the same for SRT and CRT tasks. However, the decision component differed between the tasks, with CRT requiring not only a decision *to* respond but a decision *how* to respond. Given that Landauer and colleagues found females to have faster decision times, it is possible that their advantage over males in decision time on the CRT task outweighed the male advantage in the speed of motor response. However, since the apparatus used in the present study did not fractionate RTs into movement and decision components, this explanation is only speculative and needs further evaluation. Moreover, it does not account for the difference in findings of the present study and those of Deary and Der (Deary & Der, 2005a; Der & Deary, 2006), all of which used a very similar apparatus.

It should be noted that effect sizes for the observed sex differences in both mean and variability were small. The largest effect, found for CRT ISD in adults aged 18 to 59, was -0.317 which, according to benchmarks proposed by Cohen (1992), is small. Given the relatively small number of reports of sex differences in RT intra-individual variability and that the sex differences that are reported are usually also small, it is possible that publication bias is present with null findings being under-reported. However, it should be noted that greater female variability in RT has now been found in a number of independent samples, for example, Twenty-07 (Deary & Der, 2005b), HALS (Der & Deary, 2006), and the Science Festivals sample in the present study (Dykiert et al., in revision). These are not accompanied by reports of the opposite effect, which should not be subject to publication bias. Therefore, the effect, although small is likely to be genuine.

Noble et al. (1964) have reported a finding of different ages at which CRT mean starts to decline in males and females, but this was not confirmed in the present study. There was no evidence of males and females beginning to slow at different ages, but it did appear that the increase in ISDs began earlier for females than for males.

Age effects on reaction time

The findings in terms of age effects on RT mean and intra-individual variability are in line with majority of evidence from the lifespan studies. Like Koga

and Morant (1923), Wilkinson and Allison (1989), Williams et al. (2007), and others, I found that RT trajectories are characterised by a rapid improvement in childhood and adolescence and a much slower deterioration through the adult age range. Our findings confirm Noble et al.'s (1964) observation that the developmental slope (from childhood to young adulthood) is much steeper than the decline (through adulthood and older ages). The work presented in this chapter added to the existing knowledge by comparing and contrasting, within a single study, different measures of RT intra-individual variability, including raw ISD, RT CV and RT ISDs modeled controlling for mean RT.

The shape of the age-RT relationships in adults was different from that reported by Der and Deary (2006), who found an accelerated increase in mean RT and in RT ISD in the older ages. In the present study, only one of the significant age effects in adults was quadratic (mean CRT), while the remaining effects were linear. This finding may not necessarily suggest that the two studies disagree on the shape of the curves but, rather, reflect the different ages of the participants. Der and Deary's sample comprised adults up to the age of 94, while the current investigation was carried out only up to the age of 75. It is quite possible that, if octogenarians were included in our sample, the quadratic increase among older adults would occur with the remaining measures as well. One other explanation for the difference between our studies may be the more self-selected nature of the present study's Science Festivals sample. However, if a quadratic relationship between mean CRT and age in adulthood reflects more accelerated slowing of CRT than SRT, then the present finding is consistent with that of Fozard et al. (1994), who reported larger age effects in a more complex task.

The increase in RT IIV in old age found in the present study was also less pronounced than that found by Williams et al. (2007), who used a similar sample (visitors to a science museum; age range 6-76). One explanation for this may be the difference in RT tasks between our studies. Williams et al. used a task in which some trials required suppressing irrelevant or even interfering information. Specifically, their participants had to ignore the location of the stimulus (left, right, or centre) and respond to the direction in which it was pointing (left or right). On some trials the location and the direction of the stimulus were incongruent (e.g., left pointing

stimulus presented on the right side of the screen). Group differences in RT variability are often reported to be greater on more complex task (e.g., West et al., 2002). Therefore, we might expect greater age effects to be observed in Williams et al.'s task than on SRT or CRT. However, Williams et al. found a marked increase in variability in the older adults even on neutral trials (with the stimulus presented in the centre of the screen). It is possible that the presence of incongruent trials interspersed throughout a test affects responses even on neutral trials. However, this possibility requires an empirical test.

Another finding that relates to the pattern of change in RT measures with age is that, in several RT measures, there was a plateau or an improvement apparent in the oldest age group. Such a finding is not only unexpected, but also unlikely to reflect the true lifetime pattern of change in RT. The effect is probably due to a mixture of the degree of sample selection and a healthy survivor effect. Firstly, older adults in the sample were usually grandparents of the children visitors to the Science Festivals. It is likely that those who were able to accompany their grandchildren to the exhibitions were relatively highly functioning. Secondly, it is quite probable that in the oldest end of the age range, survivor effects start playing a role. Since, as demonstrated by Shipley, Der, Taylor, & Deary (2006), both mean RT and RT variability predict all-cause mortality, with those having slower and more variable RTs more likely to die, it is not entirely surprising that the surviving older visitors to the Science Festivals have relatively shorter and more consistent RTs.

Due to relatively small numbers of adults, and in particular young adults, the current study did not allow us to examine the age when peak performance was reached or when the decline began for SRT and CRT measures. However, from the graphs of the predicted values for mean SRT and CRT, it appears that CRT declines steadily throughout the adult age range, while SRT remains relatively stable until middle adulthood, findings similar to those of Der and Deary (2006) and consistent with slightly larger cognitive load in CRT than in SRT.

Reaction time intra-individual variability controlling for reaction time mean

Sex effects were apparently affected by adjustment of ISDs for mean RT. However, the actual method of adjusting for mean RT did not change the results: the outcomes were the same when using CV as when controlling for RT mean by including it in the models. In children and adolescents, female variability was lower than male when adjusted for RT mean in both SRT and CRT; in the adult portion of the sample, females were more variable than males in CRT, whereas there was no significant effect of sex on SRT variability. Although this pattern appears surprising at first, it can be explained in terms of previously reported sex differences in speed and variability in responding. Considering children first, if females have slower SRTs and there is no sex difference in SRT ISD, then the ratio of variability to mean (and so, CV) for females will be lower than for males. Similarly for adults, if there is a significant difference in SRT mean favouring males and no difference in variability, then females will have lower variability when controlling for the RT mean. Conversely, if females have greater variability and there is no sex difference in mean (as was the case in CRT in the adult portion of the sample), then the variability to mean ratio will be greater for females than males – and so CRT CV and CRT ISD modelled with mean RT will remain greater among women.

Generally, the pattern of findings regarding sex effects on RT variability adjusted for RT mean among adults agree with those reported by Der and Deary (2006). Sex seems to have an effect on mean-adjusted variability in CRT but not in SRT. One explanation of these different outcomes in SRT and CRT tasks, in the present study, may be a significant sex difference found in mean SRT but not CRT. As there is no difference in mean CRT between the sexes, controlling for the mean, does not lead to findings other than those seen with raw CRT ISD. However, when SRT variability is adjusted for SRT mean, the greater female mean SRTs counteract greater female variability, resulting in no overall sex difference in SRT variability adjusted for mean SRT.

In terms of the relationship between age and RT IIV, the effect of controlling for mean RT depended on the method employed. In models with RT CV as an outcome there were significant age effects in both SRT CV and CRT CV in the younger group and in CRT CV among the adult participants. However, controlling for RT mean by including it in the RT ISD models had a considerable effect. In the younger portion of the sample, there was no age effect on CRT ISD modelled controlling for mean CRT but a significant age effect on mean-adjusted SRT ISD. Similarly, in adulthood age did not have a significant effect on CRT ISD when mean RT was included in the model, whereas there was a (negative) age effect on SRT ISD. These findings support, in part, those of Deary and Der (2005a) and Der and Deary (2006), who reported attenuation of age effects on RT ISD when controlling for mean RT. One important implication of these findings is that whether or not age effects on IIV are found depends on the method of controlling for RT mean one chooses to use. The simpler and more commonly used CV is related to age in a way similar to RT ISD. However, models of RT ISD with mean RT as a covariate lead to different findings than those from models of either RT CV or of unadjusted CRT ISD.

The differences in findings from the two methods of controlling for RT mean cast doubt as to the comparability of the two approaches. Since there is no consensus on which method is the best, a variety of methods are currently employed by the investigators in the field (which are outlined in Chapter 1). Some authors opt for a simple CV (Gorus, De Raedt, & Mets, 2006), others use mean RT as a covariate in their models (Shammi, Bosman & Stuss, 1998) and other still regress RT on age and other variables, such as trial number, prior to the calculation of RT ISDs (Hultsch et al., 2002; Williams et al., 2007). Given the clear indication that using RT CV and controlling for mean RT by including it in the models of RT ISD can produce quite different results, care needs to be taken when comparing findings drawn using different methods.

It should be noted that the relationship between RT mean and variability was different in the younger and older portions of our sample – it was linear in adults and cubic in under 18s. This finding supports claims of Schmiedek, Lövdén and Lindenberger (2009), who argued that the relationship is not linear and that it varies

with age. However, the difference noted here should be treated as tentative, because the power to detect non-linearity among the older participants in the Science Festivals sample was lower than in the larger, younger group. Consequently, the apparently different shapes of the relationship between RT mean and variability found in older and younger individuals in the current sample may be an artifact. This finding should be replicated on a sample with large and similar numbers of younger and older participants.

Indices of Multiple Deprivation and reaction time measures

No evidence of the effect of area deprivation on RT performance was found in the present study. The relationships between an overall IMD score and different RT measures were investigated, as well as those between separate domain scores and RT measures. It was expected that participants from more deprived areas would perform more poorly than those from less deprived backgrounds. There was no clear pattern apparent, however, and, given the number of correlations performed, the two significant correlations may be considered artifacts of multiple testing.

The null finding opposed the expectations as well as the existing evidence. For example, Johnson et al. (1985) found that males belonging to more privileged classes had faster RT than their lower class counterparts. A more recent investigation from the English Longitudinal Study of Ageing (ELSA), showed that area deprivation (assessed through IMD) had a significant and negative effect on cognitive function (Lang et al., 2008), although RT was not among the measures considered.

The null finding can be due to the very small numbers of participants from the most deprived areas. As has been shown before, about half the sample were from the least deprived backgrounds. It is likely that any effects of deprivation would be most marked in the extreme cases. Lang et al. (2008) found no difference in cognition between participants from the two least deprived quintiles, which constitute about ¾ of the sample used in the present study. The lack of a representative group from the most deprived quintile might have contributed lack of significant associations between deprivation and RT variables.

It should also be noted that IMDs were created in order to identify areas needing special attention in terms of policy making and funding and are not meant to be indices of deprivation of individuals. They are derived from various deprivation indicators of groups of people living in small areas, but IMDs do not necessarily reflect the amount of deprivation of an individual. On the one hand, not all people living in most deprived areas are themselves deprived and, on the other, not all deprived individuals live in deprived areas. Therefore, IMD is only an imperfect proxy for individual deprivation, with considerable measurement error.

Methodological issues

From a visual inspection of the plots, it becomes clear that patterns for CRT are smooth and with narrow standard errors, whereas SRT patterns are not. One reason for this may be that SRT is more prone to aberrant responses. For example, it is relatively easier to pre-press in SRT than CRT. In this study, individual trial data were available, so these could be removed prior to analyses. Pre-presses and very fast responses were much more common in SRT (1.5% of trials) than in CRT (0.1% of trials). Pre-presses were easily identifiable as RT of 0 ms. Attempts were made to exclude RTs from responses which were apparently invalid (i.e. anomalously fast or slow). One problem with SRT is that such trimming cannot remove all responses initiated before the appearance of the stimulus. One can, with a degree of confidence, deem any values below 100 or even 150 ms invalid, as it is extremely unlikely that a reaction including a motor response could be executed in a time shorter than this. In other words, recorded RT of below 100 ms reflects a response initiated before appearance of the stimulus. If a 200 ms response is initiated 150 ms before the appearance of the stimulus, and followed through, the apparatus would record a RT of 50 ms – clearly an invalid response. However, if a 200 ms response is initiated 25 ms before the stimulus, then the apparatus would record RT as 175ms, which would be considered a valid datum. Therefore, some of the fastest SRTs will invariantly be accidental or premature responses. CRT is somewhat protected from a similar issue because of the possibility of making an error. Premature responses will often result in an error—in the case of a 4-CRT—75% of the time. In the current analyses, any

individuals who made more than 20% errors were excluded and only CRTs from correct trials were considered, hence largely eliminating premature responses. The same is not possible for SRT, since a response cannot be wrong, and so, more accidental responses remain interspersed among the valid SRT responses. Therefore, the quality of data is likely to be better for CRT than SRT.

No solution to this discrepancy may be offered at this time. Due to their different nature, SRT and CRT will remain differentially vulnerable to pre-pressing. The practice of removing extremely fast responses reduces the problem of pre-presses and should be adopted whenever possible.

Strengths and limitations

The main strength of this study is that it provided an analysis of age and sex effects in a large sample, including both children and adults, and that it tested both SRT and CRT. In addition, access to trial-level data allowed better control of data quality than some earlier studies.

Although the sample was expected to consist of individuals who are from higher than average social classes and from more privileged backgrounds, this is unlikely to have affected within-sample comparisons. As shown, the IMD quintiles were similar for the two genders and for younger and older participants. Moreover, there were no differences in the main deprivation score for these groups. Therefore, although not representative of the wider population, the sample was nonetheless appropriate for the study of age and sex effects.

The sample was particularly well-suited for investigating effects with children. Previous studies which investigated RTs in children were often hindered by small sample sizes, narrow age ranges, or both. These shortcomings often prevented investigators from performing proper statistical analyses to test their hypotheses. With over 1,500 children and adolescents in the Science Festivals under 18s sample, the present study is among the largest out of the RT studies reviewed before, and the largest one on which sex differences in RT variability were investigated. The smaller number of adults, particularly of those aged over 60, compounded with the likelihood of them being relatively high functioning, make drawing firm conclusions from the

adult portion of the sample more difficult. Nevertheless, the adult sample was sufficient to show significantly greater female variability, which was not found in the childhood sample, in which there was more power to detect any differences.

The cross-sectional design of the present study is a limitation: any patterns found may be confounded by cohort effects. This limitation is particularly serious for studies examining a wide age-range, because a large gap between the youngest and the oldest subjects creates scope for difference in the circumstances and experiences of the individuals. Although the pattern of age-RT relationship can be easily determined, one cannot be certain that the same pattern will stand a few decades on, especially as some suggest sex difference in RT are diminishing with time (Silverman, 2006). Therefore, this line of research must be complemented by longitudinal investigations.

Another limitation of the present study is that young children may not be very proficient in reading numbers, which were the RT stimuli. The decline in this skill is not expected in the older adults, therefore the task may place more demands on the children than on the older participants. Older adults, on the other hand, are more likely to suffer from impaired vision. However, they are, at the same time, more likely to refuse to participate being aware of their problem. Consequently, inconsistency in responding of young children might have been overestimated. In order to be able meaningfully to compare the amount of RT IIV observed in children and older adults, stimuli easily identifiable by both groups should be used (e.g. coloured or spatially distributed lights).

Conclusion

The Science Festivals sample created a rare opportunity to investigate RTs across a wide age range, spanning from childhood to old age and including unusually large number of children. The key finding of the present study is that greater intra-individual variability of females in SRT and CRT was observed in adults, but not in children. This finding is in agreement with sex hormone hypothesis (Deary & Der, 2005a); however a direct investigation of the association between sex hormones levels and RT variability is required to verify the proposed mechanism. The disparity

of sex differences in younger and older individuals was not found in either SRT or CRT mean, which may suggest that different mechanisms bring about sex differences in mean and variability measures. The findings also confirmed the commonly reported pattern of increasing speed and decreasing variability in RT in childhood and adolescence, followed by a decrease in speed and increase in variability through adulthood and into old age. The more complex CRT task deteriorated with age faster than the simpler SRT task. The results also indicated that different methods of controlling for RT mean may lead to different findings regarding the effects of age, but not sex. Until more is known about the relationship between RT ISD and mean RT, various methods should be used in parallel to strengthen the findings.

Table 3.1. Sample size by age groups

Age group	N	Males/Females
4	31	12/19
5	80	41/39
6	145	63/82
7	185	84/101
8	224	88/136
9	249	100/149
10	224	87/137
11	179	80/99
12	153	57/96
13	97	54/43
14	38	18/20
15-19	49	18/31
20-24	22	10/12
25-29	36	11/25
30-34	24	11/13
35-39	61	29/32
40-44	75	29/46
45-49	55	21/34
50-54	18	9/9
55-59	14	8/6
60-64	17	9/8
65-69	10	7/3
70-75	8	2/6

Table 3.2. Transformation parameters (λ) used for Box-Cox transformation

RT measure	Whole sample	Under 18	18 and over
Mean SRT	-0.80	-0.71	-2.09
SRT ISD	-0.01	0.04	-0.28
Mean CRT	-0.47	-0.44	-0.28
CRT ISD	-0.13	-0.12	-0.12
SRT CV	0.00	0.05	-0.27
CRT CV	-0.03	0.04	-0.30
SRT ISD (controlling for SRT mean)	0.15	0.18	-0.18
CRT ISD (controlling for CRT mean)	0.03	0.05	-0.02

Note. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation.

Table 3.3. Means and standard deviations (in seconds) of reaction time measures for both genders in three broad age groups

RT measure	Children/teenagers (Under 18)			Adults (18-59)			Older adults (60+)		
	Male (n=697)	Female (n=950)	ES	Male (n=133)	Female (n=179)	ES	Male (n=18)	Female (n=17)	ES
Mean SRT	0.355 (0.097)	0.359 (0.092)	-0.042	0.268 (0.035)	0.281 (0.045)	-0.317	0.300 (0.059)	0.304 (0.065)	-0.065
SRT ISD	0.102 (0.064)	0.098 (0.059)	0.065	0.052 (0.026)	0.056 (0.025)	-0.157	0.058 (0.027)	0.065 (0.030)	-0.246
Mean CRT	0.765 (0.243)	0.760 (0.233)	0.021	0.512 (0.071)	0.523 (0.064)	-0.164	0.653 (0.089)	0.635 (0.092)	0.199
CRT ISD	0.184 (0.102)	0.177 (0.099)	0.070	0.084 (0.024)	0.093 (0.024)	-0.375	0.119 (0.030)	0.118 (0.025)	0.036

Note. ES=Effect size (Cohen's d for male-female difference). Values in parentheses are Standard Deviations.

Table 3.4. Summary statistics of reaction time models for the whole sample.

RT Measure	Parameter	Estimate	SE	F	p
SRT	Intercept	-2.325810621	0.02138409		
	Age	-0.050054748	0.00365492	187.56	<.001
	Age^2	0.005578373	0.00020423	746.06	<.001
	Age^3	-0.000239919	0.00002386	101.13	<.001
	Age^4	0.000004639	0.00000076	37.72	<.001
	Age^5	-0.000000033	0.00000001	22	<.001
	Sex (female)	0.059112535	0.01817287	10.58	.001
SRT ISD	Intercept	-2.992542991	0.02480925		
	Age	-0.043808486	0.00432772	98.23	<.001
	Age^2	0.004954511	0.00023707	436.76	<.001
	Age^3	-0.000224460	0.00002768	65.77	<.001
	Age^4	0.000004332	0.00000088	24.42	<.001
	Age^5	-0.000000030	0.00000001	13.48	<.001
	Sex (female)	0.001676503	0.02108135	0.01	.937
	Age*Sex (female)	0.003557388	0.00156618	5.16	.023
SRT CV	Intercept	2.966258689	0.01841788		
	Age	-0.027687810	0.00204220	194.79	<.001
	Age^2	0.002207625	0.00017986	150.66	<.001
	Age^3	-0.000062971	0.00000885	50.63	<.001
	Age^4	0.000000567	0.00000011	26.5	<.001
	Sex (female)	-0.022666674	0.01639067	1.91	.167
	Age*Sex (female)	0.002677585	0.00121682	4.84	.028
SRT ISD (adjusted for SRT mean)	Intercept	-0.760454004	0.07226390		
	Mean SRT	1.128449005	0.13956889	65.37	<.001
	Mean SRT^2	0.298425847	0.08523221	12.26	.001
	Mean SRT^3	0.046334618	0.01630603	8.07	.005
	Age	-0.002128628	0.00042470	25.12	<.001
	Sex (female)	-0.030437975	0.01043226	8.51	.004
CRT	Intercept	-0.709977579	0.01277593		
	Age	-0.047839219	0.00217952	943.23	<.001
	Age^2	0.003818190	0.00011701	2462.22	<.001
	Age^3	-0.000106547	0.00001393	209.55	<.001
	Age^4	1.2846E-06	0.00000043	55.15	<.001
	Age^5	-5.2E-09	0.00000000	24.83	<.001
	Sex (female)	0.015870654	0.01438271	1.22	.270
	Age*Sex (female)	0.008259948	0.00284619	8.42	.004
	Age^2*Sex (female)	0.000209431	0.00015692	1.78	.182
	Age^3*Sex (female)	-0.000055218	0.00001850	8.91	.003
	Age^4*Sex (female)	1.7803E-06	0.00000059	9.24	.002
	Age^5*Sex (female)	-1.64E-08	0.00000001	9.32	.002
	Errors	-0.047993135	0.00565535	72.02	<.001
	$Errors^2$	0.002775958	0.00087886	9.98	.002

RT Measure	Parameter	Estimate	SE	F	p
CRT ISD	Intercept	-2.739419654	0.02669256		
	Age	-0.070199407	0.00379760	451.88	<.001
	Age^2	0.006389895	0.00022706	1326.45	<.001
	Age^3	-0.000260074	0.00002129	155.55	<.001
	Age^4	0.000004952	0.00000067	54.11	<.001
	Age^5	-0.000000034	0.00000001	30.26	<.001
	Sex (female)	-0.048318712	0.03055601	2.5	.114
	Age*Sex (female)	0.001542613	0.00322179	0.23	.632
	Age^2*Sex (female)	0.000536999	0.00024614	4.76	.029
	Age^3*Sex (female)	-0.000010165	0.00000392	6.74	.010
	Errors	-0.019281368	0.00430464	20.06	<.001
CRT CV	Intercept	2.747160582	0.01530983		
	Age	-0.017298717	0.00200582	97.67	<.001
	Age^2	0.001484319	0.00011998	285.15	<.001
	Age^3	-0.000070811	0.00001125	42.62	<.001
	Age^4	0.000001522	0.00000036	18.31	<.001
	Age^5	-0.000000011	0.00000000	11.92	.001
	Sex (female)	-0.032342048	0.01614087	4.01	.045
	Age*Sex (female)	0.000499706	0.00170176	0.09	.769
	Age^2*Sex (female)	0.000294195	0.00013001	5.12	.024
	Age^3*Sex (female)	-0.000005033	0.00000207	5.92	.015
	Errors	0.022795327	0.00691718	10.86	.001
	$Errors^2$	-0.002136896	0.00107577	3.95	.047
CRT ISD (adjusted for CRT mean)	Intercept	-1.420524450	0.01161511		
	Mean CRT	1.478297299	0.02591503	3254.02	<.001
	Mean CRT^2	0.556858143	0.07057742	62.25	<.001
	Mean CRT^3	0.247756231	0.06003380	17.03	<.001
	Age	-0.003401441	0.00049553	33.9	<.001
	Sex (female)	-0.007199943	0.00830170	0.75	.386
	Age*Sex (female)	0.002525520	0.00061326	16.96	<.001
	Errors	0.041656001	0.00680799	37.44	<.001
	$Errors^2$	-0.003199509	0.00105048	9.28	.002

Note. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation.

Table 3.5. Summary statistics of reaction time models for ages below 18 years.

RT measure	Parameter	Estimate	SE	F	p
Mean SRT	Intercept	-1.703888702	0.01568759		
	Age	-0.133129629	0.00360536	1363.49	<.001
	Age^2	0.008645723	0.00111810	59.79	<.001
	Sex (female)	0.043532753	0.01820898	5.72	.017
SRT ISD	Intercept	-2.402987603	0.01331377		
	Age	-0.100375692	0.00701983	204.46	<.001
	Age^2	0.010225272	0.00140467	52.99	<.001
	Age^3	-0.000823129	0.00035882	5.26	.022
SRT CV	Intercept	3.463977410	0.01858016		
	Age	-0.055780257	0.00721472	59.78	<.001
	Age^2	0.006398262	0.00144426	19.63	<.001
	Age^3	-0.000917849	0.00036897	6.19	.013
	Sex (female)	-0.042932090	0.02126626	4.08	.044
SRT ISD (adjusted for SRT mean)	Intercept	-0.883810311	0.08155829		
	Mean SRT	0.955478509	0.16233484	34.64	<.001
	Mean SRT^2	0.252102650	0.10738551	5.51	.019
	Mean SRT^3	0.048016869	0.02250010	4.55	.033
	Age	0.003512423	0.00425316	0.68	.409
	Age^2	0.003256881	0.00081883	15.82	<.001
	Age^3	-0.000765769	0.00019540	15.36	<.001
	Sex (female)	-0.033465480	0.01082275	9.56	.002
Mean CRT	Intercept	-0.312502702	0.00898151		
	Age	-0.105057382	0.00152522	4744.49	<.001
	Age^2	0.006216372	0.00047116	174.07	<.001
	N errors	-0.040732910	0.00630167	41.78	<.001
	N $errors^2$	0.002006658	0.00096058	4.36	.037
CRT ISD	Intercept	-2.106115042	0.01623884		
	Age	-0.160083808	0.00596947	719.16	<.001
	Age^2	0.013453352	0.00119154	127.48	<.001
	Age^3	-0.001043640	0.00030434	11.76	.001
	N errors	-0.016038082	0.00454958	12.43	<.001
CRT CV	Intercept	3.219858897	0.01561241		
	Age	-0.044071092	0.00396572	123.50	<.001
	Age^2	0.005307086	0.00079178	44.93	<.001
	Age^3	-0.000770515	0.00020219	14.52	.001
	Sex (female)	-0.025935000	0.01166785	4.94	.026
	N errors	0.033123963	0.00958356	11.95	.001
	N $errors^2$	-0.003265680	0.00145991	5.00	.025
CRT ISD (adjusted for CRT mean)	Intercept	-1.378898222	0.01215834		
	Mean CRT	1.453126025	0.02563960	3212.06	<.001
	Mean CRT^2	0.532578324	0.07439869	51.24	<.001
	Mean CRT^3	0.225049765	0.07540139	8.91	.003
	Sex (female)	-0.020655362	0.00884757	5.45	.020
	N errors	0.043475335	0.00733225	35.16	<.001
	N $errors^2$	-0.003542471	0.00111079	10.17	.002

Note. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation.

Table 3.6. Summary statistics of reaction time models for ages 18 and above.

RT measure	Parameter	Estimate	SE	F	p
SRT	Intercept	-7.2462180	0.1657992		
	Age	0.0354655	0.0090814	15.25	<.001
	Sex (female)	0.4899546	0.2207139	4.93	.027
SRT ISD	Intercept	-4.8838300	0.0803724		
	Sex (female)	0.2222442	0.1069408	4.32	.038
SRT CV	Intercept	2.0026396	0.0086203	53971	<.001
SRT ISD (adjusted for SRT mean)	Intercept	-2.3531880	0.1059012		
	Mean SRT	0.2369009	0.0145822	263.93	<.001
	Age	-0.0052750	0.0025217	4.38	.037
CRT	Intercept	-0.6517050	0.0108219		
	Age	0.0076309	0.0005963	163.76	<.001
	Age^2	6.971E-05	3.373E-05	4.27	.040
	N errors	-0.0408690	0.0043306	89.06	<.001
CRT ISD	Intercept	-2.8305570	0.0330910		
	Age	0.0117774	0.0014949	62.07	<.001
	Sex (female)	0.1312265	0.0363016	13.07	<.001
	N errors	-0.0386970	0.0114354	11.45	.001
CRT CV	Intercept	1.8825397	0.0069889		
	Age	0.0008574	0.0003828	5.02	.026
	Sex (female)	0.0346750	0.0093037	13.89	<.001
CRT ISD (adjusted for CRT mean)	Intercept	-1.6727440	0.0486499		
	Mean CRT	1.2539405	0.0686302	333.83	<.001
	Sex (female)	0.0748501	0.0220551	11.52	.001
	N errors	0.0208817	0.0075432	7.66	.006

Note. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation.

Table 3.7. Correlations between reaction time measures and Index of Multiple Deprivation scores (overall and for each domain) by country

IMD	Mean SRT	SRT ISD	Mean CRT	CRT ISD	SRT CV	CRT CV
Scoltand						
IMD score	.004	.011	-.017	.002	.014	.025
Income	.005	.016	-.010	.006	.022	.030
Education	.030	.027	-.013	.005	.025	.029
Health	.038	.027	.012	.020	.021	.028
Employment	-.003	.007	-.014	-.005	.015	.012
Housing	.041	.038	-.020	-.013	.032	.006
Environment	-.048	-.026	-.027	-.024	-.015	-.017
Crime	.059	.017	.030	**.064**	-.007	**.084**
England						
IMD score	-.013	-.015	-.006	.004	-.018	.014
Income	-.017	-.037	.014	.018	-.040	.020
Education	.054	.014	.024	.029	-.012	.012
Health	.030	.004	.017	.014	-.016	.002
Employment	.002	-.025	.014	.003	-.038	-.008
Housing	.001	.012	.014	.016	.018	.023
Environment	-.068	-.037	-.041	-.029	-.017	-.003
Crime	-.061	-.062	-.072	-.071	-.059	-.040

Note. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation, IMD = Index of Multiple Deprivation.
Values in bold are significant at the .05 level.

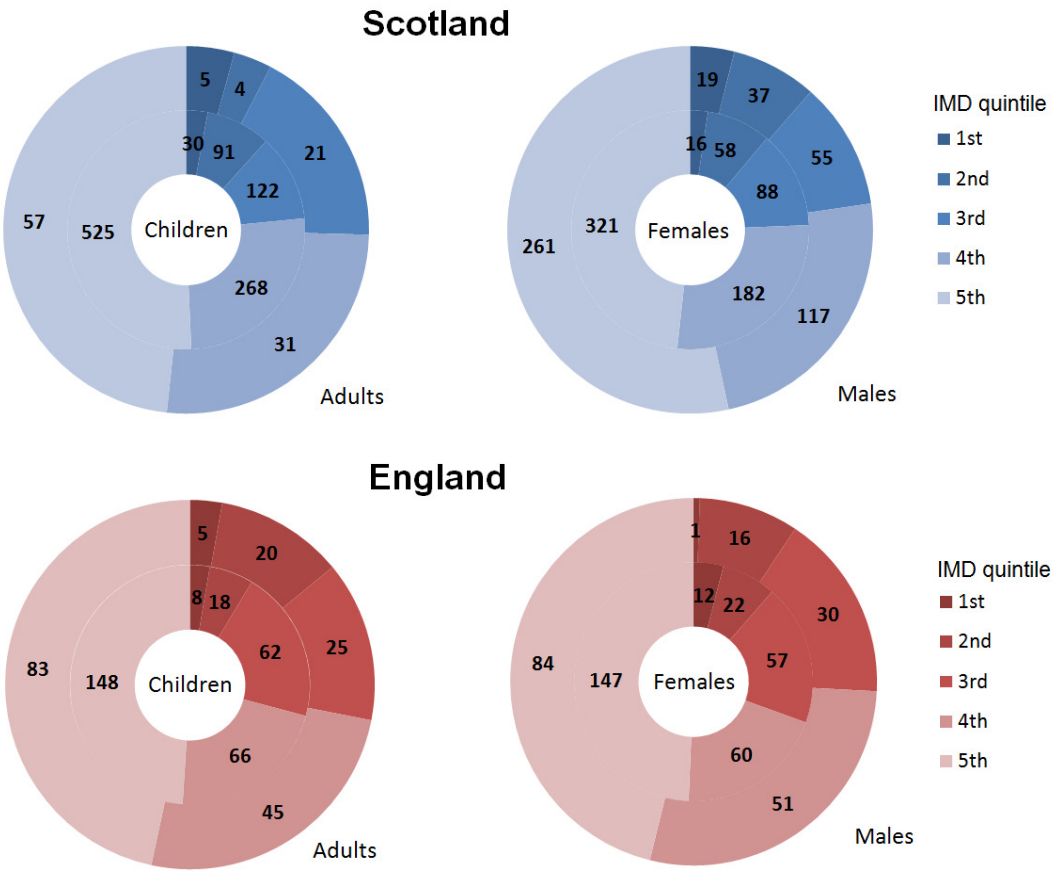

Figure 3.1. Number of representatives from each quintile of Index of Multiple Deprivation by country (blue=Scotland, red=England), age group (inner ring = children, outer ring = adults), and sex (inner ring = females, outer ring = males). Lowest quintiles were most deprived. The shaded areas represent the proportional representation of different quintiles within each group considered.

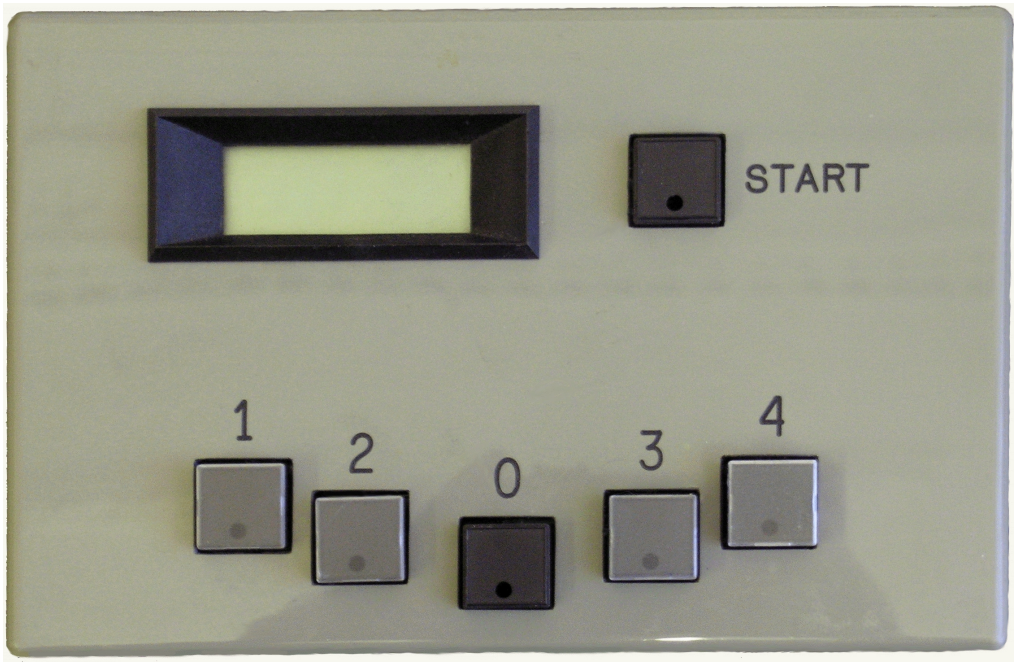

Figure 3.2. Reaction time apparatus.

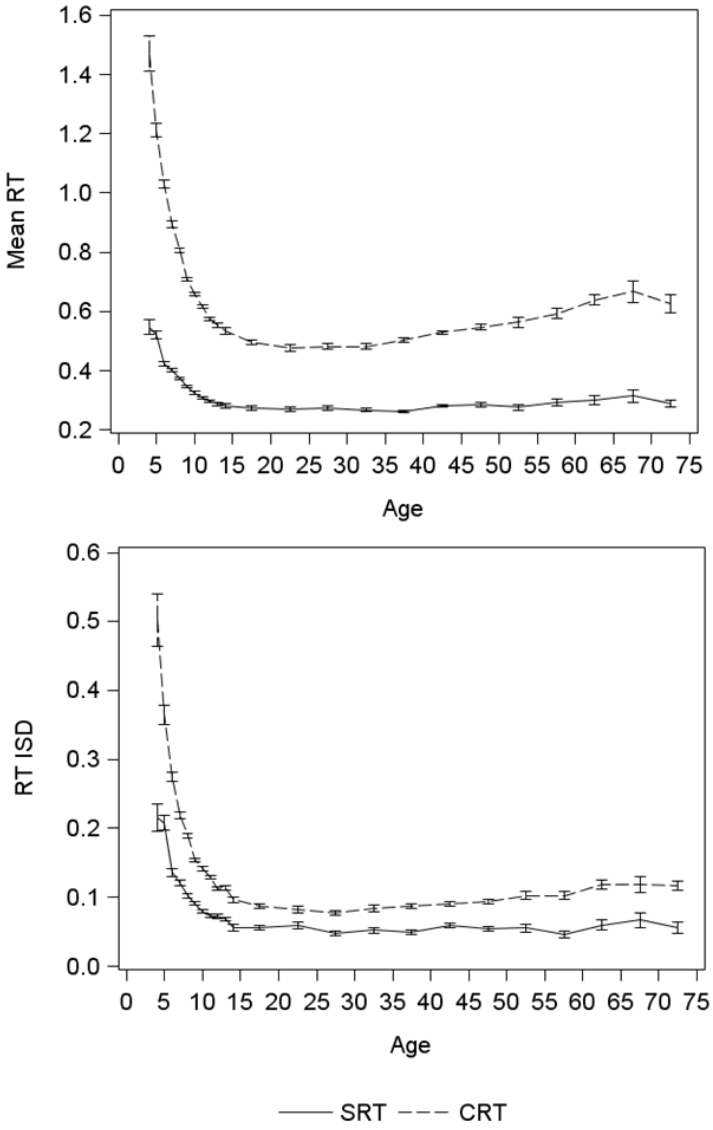

Figure 3.3. Means of reaction time measures across age (grouped in 1-year bands up to 14 and 5-year bands thereafter). Error bars represent group standard errors.

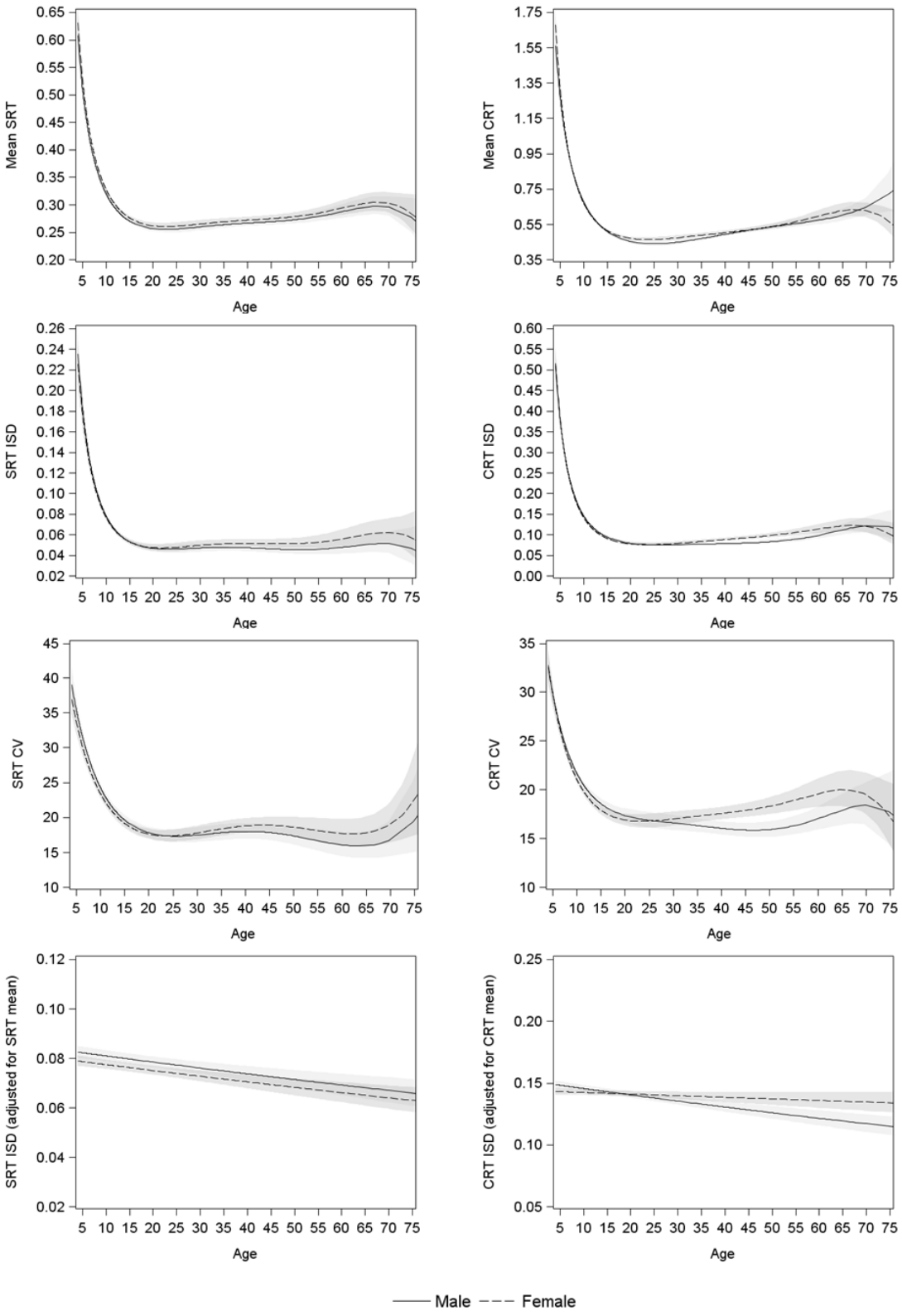

Figure 3.4. Predicted values and confidence bands for reaction time measures by age and sex.

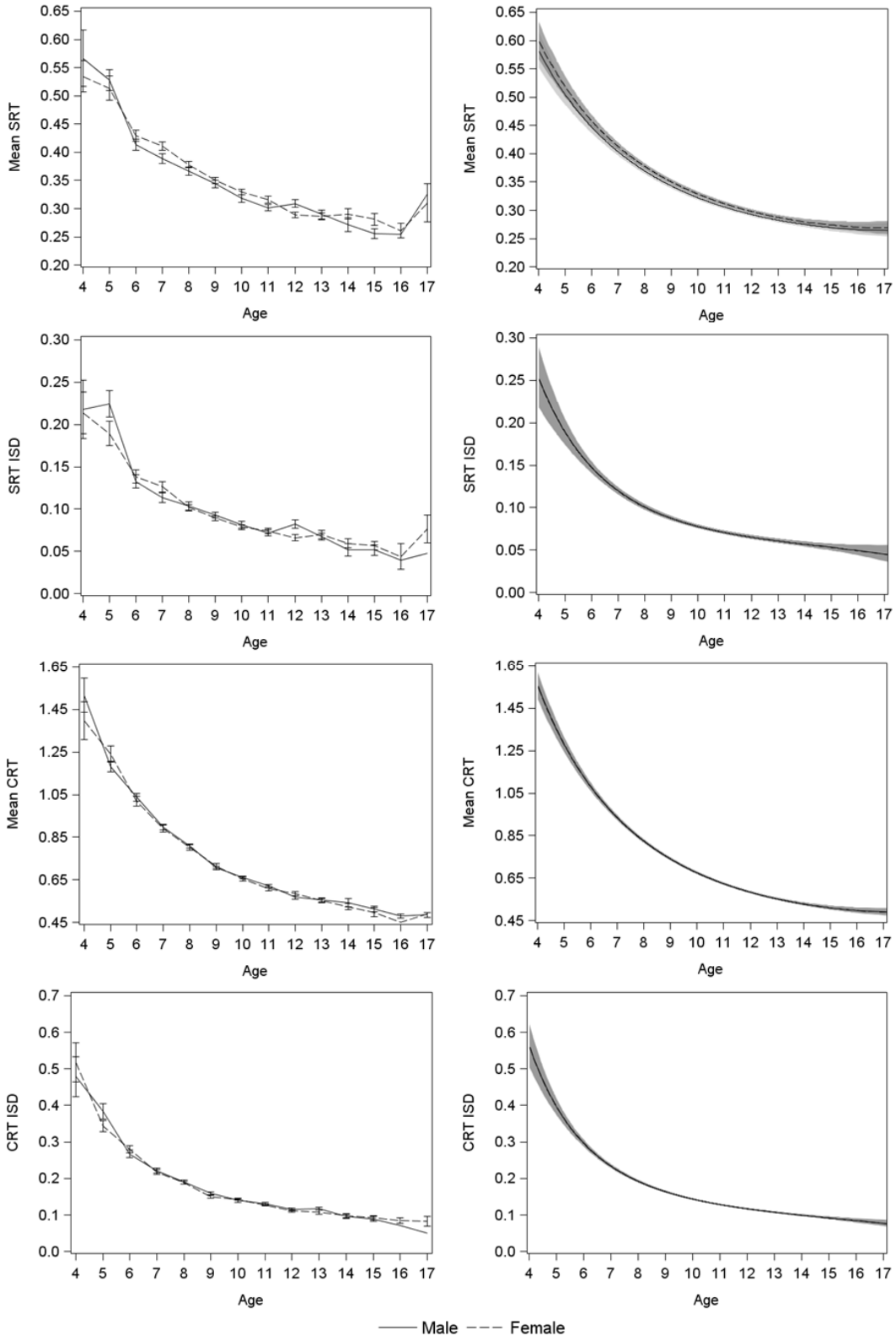

Figure 3.5. Observed and predicted vales of reaction time measures for males and females aged under 18. Left panel – means and standard errors (in 1-year age bands); right panel – predicted values and their corresponding 95% confidence bands.

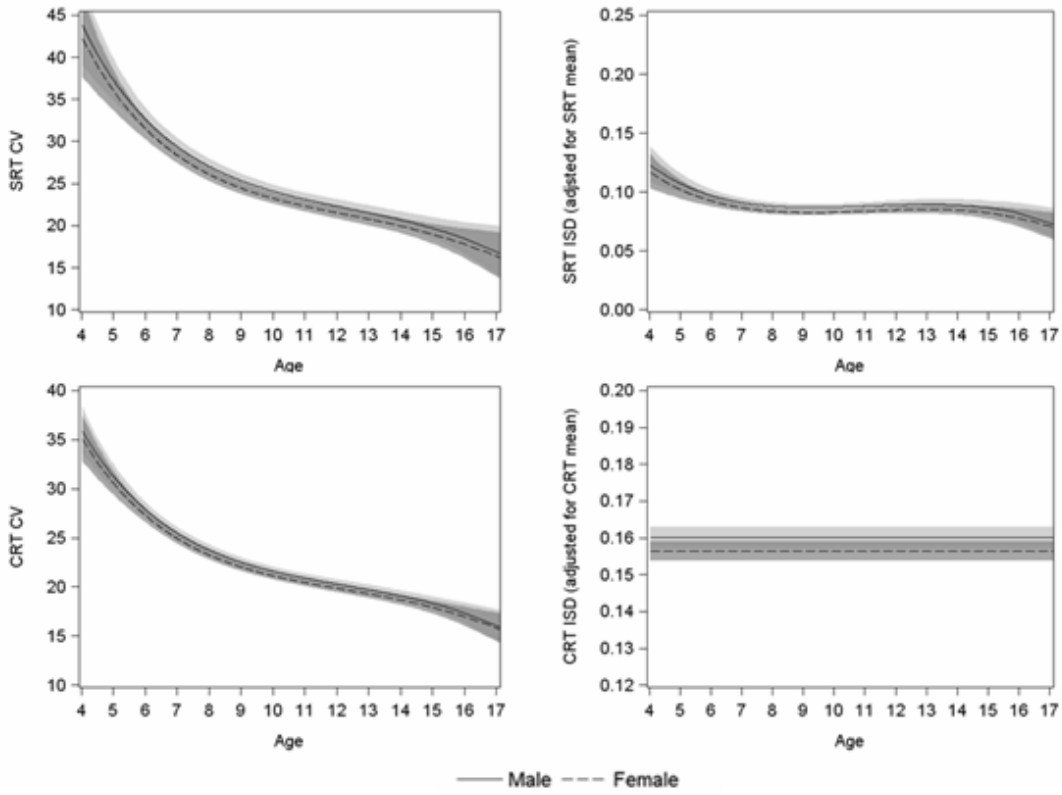

Figure 3.6 Predicted values and their corresponding 95% confidence bands of simple and choice reaction time coefficients of variability (SRT CV, CRT CV) and intra-individual standard deviation from models controlling for mean reaction time for participants aged under 18.

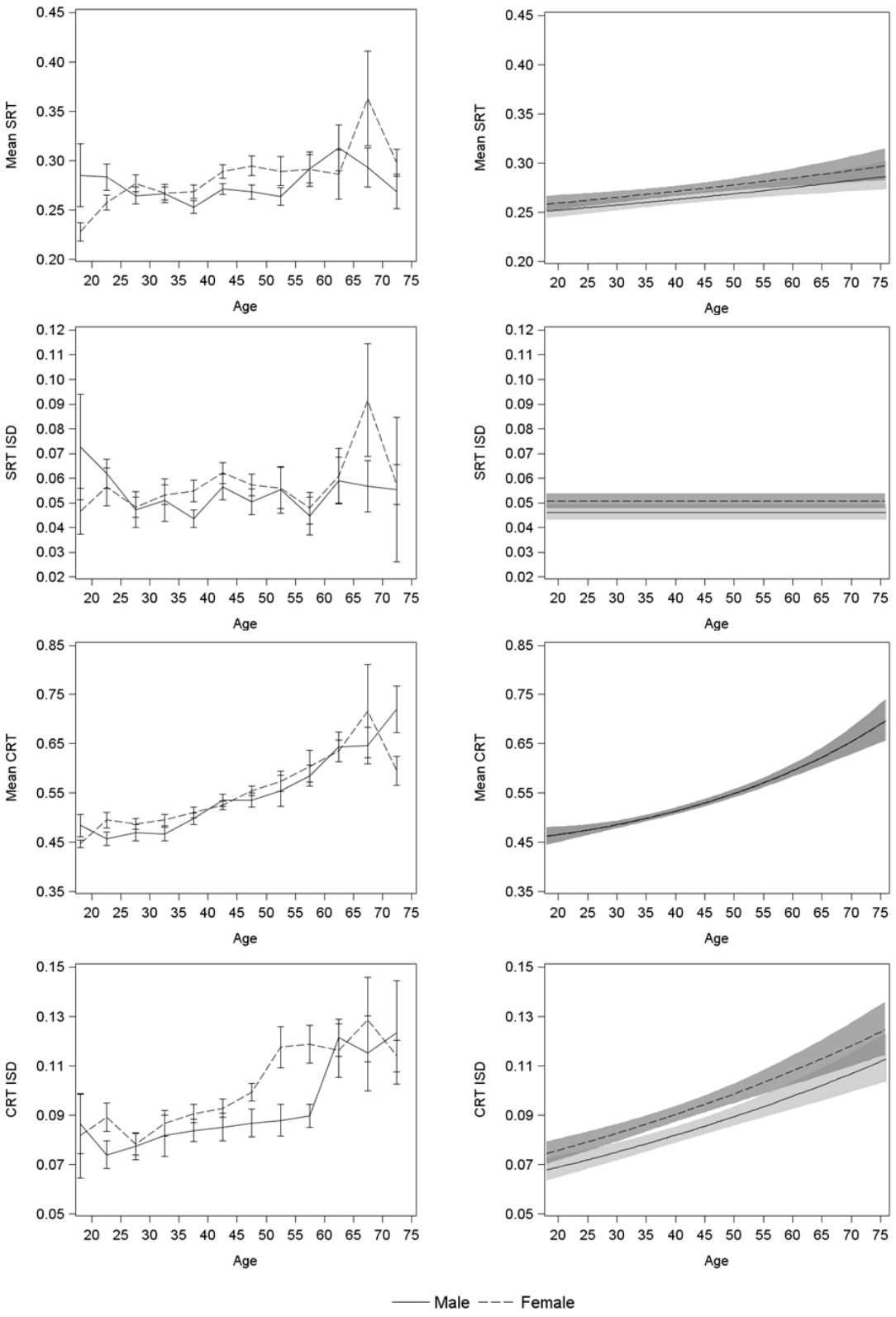

Figure 3.7. Observed and predicted values of reaction time measures for males and females aged 18 or over. Left panel – means and standard errors (the first age band includes ages 18 and 19, the remaining ones are 5-year bands); right panel – predicted values and their corresponding 95% confidence bands.

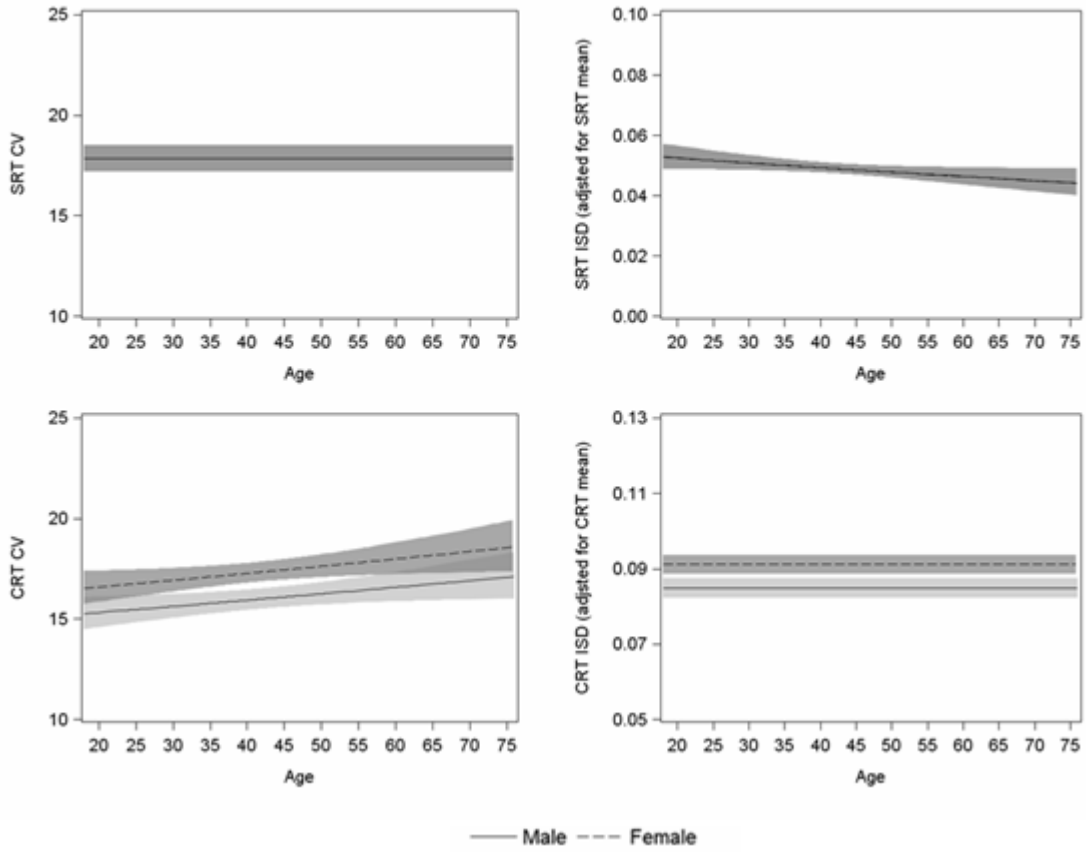

Figure 3.8. Predicted values and their corresponding 95% confidence bands of simple and choice reaction time coefficients of variability (SRT CV, CRT CV) and intra-individual standard deviation from models controlling for mean reaction time for participants aged 18 and over.

Chapter 4 -
Age and sex effects on reaction time variability: Insights from a longitudinal Twenty-07 study

As reviewed in Chapter 3, evidence from cross-sectional investigations suggests that reaction time intra-individual variability (RT IIV), in adults, is greater at older ages. The empirical data from the Science Festivals sample provided further support for this observation. However, cross-sectional studies are affected by a number of limitations which, especially when used to make inferences about age effects, may lead to erroneous conclusions. Crucially, cross-sectional studies do not provide information about *change* over time, but only about *differences* between age groups, which are confounded with cohort differences. Although findings from cross-sectional studies of age differences are often interpreted to be informative of age effects, Schaie (2005) suggests that:

> unless there is independent evidence to suggest that older cohorts performed at the same level as younger cohorts at equivalent ages, it would be most parsimonious to assume, at least in comparisons of adult samples, that cross-sectional age differences represent estimates of cohort differences that may be either inflated or reduced by maturational changes occurring over a specified age range (pp. 21-2).

Longitudinal studies, although subject to a host of additional problems, at least have the ability to test hypotheses about change with age. They are obviously more demanding in terms of time and effort than comparable cross-sectional studies. Therefore, they are rather scarce. In fact, only a few studies have investigated changes in simple or choice RT (SRT, CRT) variability over time.

Maddox and Douglass (1974) investigated change in RT variability on a 2-choice visual RT in participants from the Duke Longitudinal Studies. RT data were available from five out of six waves of testing, spanning approximately 13 years. The study conceptualised IIV as the standard deviation of mean decision times from four blocks of 10 RT trials. This measure showed no change over time among the 56

participants who completed all waves. The authors concluded that "stability in intra-individual variability in RT measures characterizes these subjects in late life" (p. 561).

Fozard, Vercruyssen, Reynolds, Hancock, and Quilter (1994) reported on data from another study which collected RT data across a number of waves, the Baltimore Longitudinal Study of Aging (BLSA). The BLSA included a snowball sample of family, friends and acquaintances of an initial group of participants, who were not representative of the Baltimore population. Nonetheless, the BLSA provides data on a relatively large sample of individuals with a wide age range. Fozard et al.'s report is based on data from 1,265 individuals aged between 17 and 96 years. As already mentioned in Chapters 2 and 3 of this thesis, cross-sectional data from Fozard et al.'s study showed greater variability in both SRT and a disjunctive (go/ no-go) RT task was in older age groups. The longitudinal findings from three waves across 4 years were less clear-cut, but the power to detect differences was reduced by the smaller longitudinal sample (580 individuals completed three waves of testing over 4 years). SRT IIV increased across the 4-year period but, whereas go/no-go RT showed slowing over-time, no significant change in variability was observed for this task. When SRT and go/no-go tasks were analysed as two factors of RT task, a significant age by sex by task by visit interaction was found. This interaction was not explored further. Fozard et al. also investigated longitudinal change over 8 years, and reported that the findings were similar to those obtained from cross-sectional and 4-year longitudinal investigations. However, the sample was further reduced to 279 participants for that analysis and many of the comparisons did not reach statistical significance.

MacDonald, Hultsch and Dixon (2003) considered both cross-sectional and longitudinal age effects on SRT and CRT variability in three groups of participant in the Victoria Longitudinal Study. The age groups considered were: young-old (aged 55 to 64 years), mid-old (aged 65 to 74), and old-old (aged 75 to 89). MacDonald et al.'s investigation is based on participants who completed three testing waves across a 6-year period, with a total of 446 individuals. They used a measure of variability independent of age group, sex, or trial effects and, for longitudinal analyses, the effect of occasion was also partialled out. Specifically, they calculated the intra-

individual standard deviation (ISD) from residuals purified of these effects. In cross-sectional comparisons, the oldest groups had more variable SRT than the middle group which, in turn, was more variable than the youngest group. Variability in CRT was greater in the old-old group than either mid-old or young-old groups, but the latter two did not differ significantly from each other. Longitudinal findings did not confirm these effects and CRT IIV actually decreased across the 6 years, whereas SRT IIV showed an increase but only in the oldest group (and, arguably, only at the last wave).

Deary and Der (2005a) analysed data from the first two waves with RT data from the West of Scotland Twenty-07 study. Twenty-07 is a prospective longitudinal study, which followed up three narrow age cohorts, aged about 15, 35 and 55 on the first wave of testing, for 20 years. The initial sample was representative of the population from which it was drawn, and so Twenty-07 provides one of the most representative samples from which longitudinal RT data are available. Almost 2,000 people had complete data from the two waves and were included in Deary and Der's analysis. SRT ISD increased over the 8-year follow-up across the cohorts; CRT ISD decreased in the youngest and increased in the oldest cohort, with not much change noted in the middle cohort. Deary and Der also found sex differences, with females being more variable. When mean RT was controlled, sex differences remained, but age effects were largely removed.

The longitudinal data available to date do not provide compelling evidence for longitudinal age effects on inconsistency in RT. The lack of clear pattern of age-related increase in RT IIV over time may be because the evidence to date is based on small or unrepresentative samples (Fozard et al., 1994; Maddox & Douglass, 1974) and relatively short follow-up periods (Deary & Der, 2005a; MacDonald et al., 2003). The present study aimed to enhance the available information by investigating longitudinal changes in RT IIV in a large and representative sample (Twenty-07) followed up for a relatively long period of time of 20 years.

Method

Participants

Participants were from the West of Scotland Twenty-07 study (MacIntyre et al., 1989). It is a longitudinal study, which started in 1987, aiming to investigate health inequalities in individuals living in Glasgow and the surrounding areas. It included three very narrow age cohorts born around 20 years apart. The three cohorts were aged about 15 years (cohort 1), 35 year (cohort 2), and 55 years (cohort 3) at the first wave of testing in 1987/88 (see Figure 4.1 for an illustration of the narrow age ranges in each cohort at wave 1). They were followed up about every 4 years for 20 years. Therefore, at the last wave, in 2007/08, the youngest cohort was the same age as the middle cohort at the start of the study, and the middle cohort was the same age as the oldest cohort at the start of the study. Taken together, these three cohorts cover 60 years. The design of the Twenty-07 study allows the investigation of lifespan changes across 60 years, longitudinal changes over 20 years within individuals, as well as cross-sectional differences between the three cohorts.

The initial sample (n = 4,510) comprised two sub-samples: *regional* (from the Central Clydeside Conurbation) and *locality* (including two contrasting areas of Glasgow). The regional sample recruitment was performed in two stages (see Ecob, 1987, for a detailed description of sampling procedures). First, local areas were stratified by rates of unemployment and proportion of groups on low income and 52 postcode sectors were systematically selected from these. The probability of each sector being selected was proportional to the size of its population. Second, individuals were randomly selected from within the chosen sectors. Der (1998) compared the Twenty-07 sample with the 1991 census data on the relevant population from which the sample was drawn. He found that they were similar in terms of four variables considered: sex, social class, household tenure and car ownership. Therefore, it can be concluded that the Twenty-07 sample is generally representative of its population.

The analyses in the present study are based on the regional sub-sample of the Twenty-07 study (n = 3,036), described above. This sample was chosen because it

was representative of its population and had RT data collected at four waves. There were 1,204 participants with complete RT data from the four waves. Out of these, 23 were removed because of low accuracy on CRT task (8 or more errors) and one had SRT ISD recorded as 0 ms. These exclusions left a working sample of 1,180 participants (532 males, 648 females), with 365 people in cohort 1, 487 in cohort 2 and 328 in cohort 3.

Reaction time measures

SRT and CRT data were collected at waves 1, 3, 4, and 5 (no RT data were collected at wave 2), using a small portable apparatus, which was originally designed for and used in the Health and Lifestyle Survey (HALS; Cox, Huppert, & Whichelow, 1993). The device is described in Chapter 3 (see also Figure 3.2 for an illustration). The RT tests were administered in participants' homes by trained nurses.

In the course of the study, three versions of the apparatus were used. All testing at wave 1 was performed on version 1; waves 3 and 4 used a mixture of version 1 and version 2 devices, which were made by the same company to the same specification. A new version (version 3) of the apparatus was introduced at wave 5 and all the data were collected using this upgraded apparatus. Version 3 of the apparatus was an upgraded model with additional functionality – it was re-designed to record individual trial data as well as summaries, which were the only data available from the earlier two versions. The final version of the RT testing device was developed by the same company as the earlier two versions and attempts were made to replicate all the features as far as possible.

The variables used in this study were mean latency of SRT and CRT (mean SRT, mean CRT), raw IIV (SRT ISD and CRT ISD), and mean-adjusted IIV (SRT CV and CRT CV, calculated after the trimming described below). CRT measures were obtained only from correct responses.

Data preparation

The RT data were initially cleaned after each collection wave. The screening of the full dataset revealed that there were more very large mean RTs and RT ISDs at wave 5 than in the remaining waves, especially for the middle and older cohorts. Because of long intervals between the data collection waves, a few different researchers performed the initial data cleaning. Therefore, the differences observed might be a result of the changes in RT testing devices used at different waves, the potential variations in the trimming rules applied at different waves, or a mixture of both. Raw data from the earlier waves were no longer available, nor were the rules used to clean them. Because of this, the data were trimmed based on preliminary models fitted to them. The procedure was based on the technique proposed by Ecob and Der (2003). In brief, this procedure trimmed outliers, which were defined as residuals falling beyond 5 SDs from the line of best fit. Identified data points were replaced with a value predicted from the model plus or minus mean of the residuals calculated for a given individual across the four waves of testing. The sign (plus or minus) depended on whether the data point identified as an outlier fell above or below the predicted value.

Overall, 11 mean SRTs were trimmed and replaced using this method. One trimmed data point was from wave 1, two from wave 3, two from wave 4, and six from wave 5. Four were from cohort 2 and seven from cohort 3. Sixteen SRT ISDs were trimmed and replaced. Most of the trimmed data points were from wave 5 (four from cohort 2 and 11 from cohort 3); one outlying SRT ISD was from wave 3 (cohort 3). There were 9 mean CRTs that needed trimming and replacing. Eight were from wave 5 and one from wave 1; two were from cohort 2 and seven from cohort 3. Finally, 12 CRT ISDs were trimmed. All these were from wave 5 and from the two older cohorts: four from cohort 2 and eight from cohort 3.

In total, 48 individual data points across the four RT measures were trimmed and replaced. The majority of them were from the last wave (wave 5) and from the oldest cohort (cohort 3). It is clear that outliers were not randomly distributed across cohorts and waves. However, because greater variability in older cohorts is expected, removing more outliers from that group relative to the middle and youngest ones

would act to reduce the expected age effect. Therefore, an age effect indicating greater variability among older rather than younger individuals would be underestimated rather than overestimated by the trimming procedure adopted.

Statistical analysis

RT IIV data were modelled using mixed models analysis (PROC MIXED, SAS, version 9.2). The effects of cohort (1, 2, and 3), wave (1, 3, 4, and 5), sex (male, female), and their interactions on each of the four RT IIV measures of interest (SRT ISD, CRT ISD, SRT CV, and CRT CV) were tested. Because the age ranges in each cohort were very narrow, wave was used as a proxy for age, with the caveat that it represents a mixture of ageing and wave of testing effects (including practice).

Because RT data are often positively skewed, residuals were examined after fitting initial models. The residuals from all initial models were positively skewed and heteroscedastic; therefore, the data were transformed using Box-Cox procedure (Box & Cox, 1964, as described previously in Chapter 3 of this thesis). Subsequently, all models were re-fitted using the transformed data.

Results

SRT IIV

Cohort 3 had the highest mean SRT ISD, followed by cohort 2, and cohort 1 (see Table 4.1 and Figure 4.2). Females had greater SRT ISD than males. In each cohort, the difference between males and females appeared larger on waves 3 and 4 than on waves 1 or 5. SRT ISD increased between waves 1 and 3 and decreased between waves 3 and 4 for all cohorts. Cohorts 1 and 2 showed a decrease in SRT ISD from wave 4 to 5, but this trend was not observed in cohort 3. Mixed models analysis revealed that there were significant effects of cohort $F(2, 1176) = 62.58$, $p < .001$; wave, $F(3, 3528) = 76.47$, $p < .001$; and sex, $F(1, 1176) = 14.58$, $p < .001$. There were two significant interactions: cohort by wave $F(6, 3528) = 19.63$, $p < .001$,

and sex by wave, F(3, 3528) = 3.64, p < .012). All pairwise comparisons between cohorts were significant. SRT ISD increased between waves 1 and 3 and decreased on the subsequent waves. The cohort by wave interaction appeared to reflect the decline in SRT ISD from wave 4 to 5 in cohorts 1 and 2, but not in cohort 3. The interaction between sex and wave was due to females having greater SRT ISD than males at waves 3, 4 and 5, but the sexes not being significantly different at wave 1.

SRT CV showed a similar pattern of means, with greatest variability in cohort 3, intermediate in cohort 2 and least in cohort 1, but the difference between the two younger cohorts was very small (see Table 4.1 and Figure 4.2). Females had slightly higher mean SRT CV than males. All cohorts showed an improvement between waves 1 and 3; at the subsequent waves, a decrease in SRT CV could be observed, except for males in the oldest cohort, who showed a linear increase across waves. There were significant effects of cohort, F(2, 1176) = 20.19, p < .001; wave, F(3, 3528) = 78.01, p < .001 and sex, F(1, 1176) = 11.26, p < .001. Significant interactions were found between cohort and wave, F (6, 3528) = 19.54, p < .001 and between sex and wave, F (3, 3528) = 3.36, p < .018. Pairwise comparisons confirmed significant differences between cohorts 1 and 5 and between cohorts 3 and 5; cohorts 1 and 3 were not significantly different from each other. SRT CV increased from wave 1 to wave 3 and decreased on the subsequent waves. Cohorts 1 and 2 became more variable between waves 1 and 3 and less variable on subsequent occasions. Cohort 3 also became more variable between waves 1 and 3, but there were no significant changes in SRT CV on subsequent waves. Sex differences were found at waves 3 and 4 but not at waves 1 or 5.

CRT IIV

CRT ISD was smallest in cohort 1, largest in cohort 3 and intermediate in cohort 2; females had more variable CRTs than males (see Table 4.1 and Figure 4.2). Wave appeared to have different effect across cohorts: CRT ISDs decreased from wave to wave in cohort 1, and increased with each wave in cohorts 3 and remained relatively stable in cohort 3. The linear mixed model revealed significant effects of cohort, F(2, 1174) = 251.29, p < .001; wave, F(3, 3522) = 2.67, p = .046; and sex

F(1,1174) = 57.75, p < .001. There was a significant cohort by wave interaction, F(6, 3522) = 60.92, p < .001, and a three-way interaction between cohort, wave and sex, F(6, 3522) = 2.90, p = .008. All pairwise comparisons between cohorts were significant. The effect of wave seemed to be driven by the smaller CRT ISD at wave 1 than any other wave. CRT ISD in cohort 1 decreased across the four waves, whereas in cohort 2 it increased from wave 1 to wave 3 and remained stable thereafter. In cohort 3, CRT ISD increased steadily from wave to wave. In cohort 1 there was no sex difference at wave 1, whereas females had greater CRT ISD than males at waves 3, 4, and 5. In cohort 2, females were consistently more variable than males across the four waves. In cohort 3, there was no significant sex difference at wave 4, but at all other waves, females demonstrated greater variability in CRT than males.

For CRT CV the differences between cohorts were relatively small, but the pattern remained similar to CRT ISD; that is, variability was smallest in the youngest cohort, greatest in the oldest cohort and intermediate in the middle cohort (see Table 4.1 and Figure 4.2). Females were more variable than males. Again, the effect of wave appeared different for different cohorts. The general trend was for CRT CV to decrease across waves in cohort 1, to increase and then stabilise in cohort 2, and to increase from wave to wave in cohort 3. It appeared that whereas females showed no decrease in CRT CV on waves 1-3, males' CRT CVs showed a linear decrease across the four waves. Statistical analysis revealed significant effects for cohort F(2, 1174) = 4.17, p = .016; wave F(3, 3522) = 8.20, p < .001; and sex, F(1, 1174) = 82.96, p < .001. There was a significant cohort by wave interaction F(6, 3522) = 25.25, p < .001, and a three-way cohort by sex by wave interaction, F(6, 3522) = 3.08, p = .005. The only significant difference between cohorts was between the youngest and the oldest one. CRT CV was significantly higher at wave 3 than at wave 1, with no significant differences between waves 3 and 4. Across the three cohorts, CRT CV decreased between waves 4 and 5, but this effect was only significant in cohorts 1 and 2. There was no significant sex difference in the youngest cohort at wave 1, but from wave 3 onwards, females demonstrated greater CRT CV than males. In addition to that, whereas males' CRT CV decreased across the four waves, CRT CV of

females did not change until the last testing wave. Sex differences were significant at all waves in cohort 2 and all but wave 4 in cohort 3.

Discussion

Cross-sectional and longitudinal effects of age

For all four RT variability measures considered in this study (SRT ISD, SRT CV, CRT ISD, and CRT CV), variability tended to be greatest in the oldest cohort, smallest in the youngest cohort and intermediate in the middle cohort. This finding is in line with numerous earlier cross-sectional reports of greater RT IIV with increasing age in adults (Der & Deary, 2006; Fozard, et al., 1994; Hultsch et al., 2002; see also Chapter 3 of this thesis).

For both RT tasks considered here, there were significant longitudinal effects of wave, providing some indication that RT variability changes within-people with age. However, the patterns were different in SRT and in CRT. Moreover, wave interacted with cohort and with sex in models of SRT IIV and a three-way interaction was found in models of CRT IIV. The effects of wave varied between cohorts and sexes.

Wave effects, especially on measures of SRT IIV, were non-linear. Also, for SRT IIV measures the effects were somewhat counterintuitive. All cohorts showed an increase in SRT IIV between waves 1 and 3 (as found previously by Deary & Der, 2005a), and a decrease between waves 3 and 4. In cohorts 1 and 2, there was also a pronounced drop in variability from wave 4 to 5, beyond the level of performance observed at the first wave of testing. Although some decrease in SRT IIV over time could be expected in the youngest cohort (e.g., Der & Deary, 2006, found that SRT ISD continue to decrease with age in young adulthood) the effect in the middle-aged cohort is less readily explained. Because the analyses reported here were based only on participants who attended all four waves of testing, the effect cannot be attributed

to selective drop-out of individuals with poorer (i.e. slower and more variable) performance.

A possible explanation for the unexpected pattern of SRT IIV is the change in the RT apparatus used across waves. As already mentioned, three versions of the device were used in the course of the study. At wave 1 all data were collected using the first version of the device. At wave 3 a new device was introduced and used alongside the first version. The same mixture of the two versions of devices was used again at wave 4. Wave 5, however, used a newly-designed version of the testing device, and did not make use of either of the previous versions. Consequently, Waves 1 and 3 had some devices in common, waves 3 and 4 had all devices in common, whereas wave 5 had no common devices with any of the other waves. Therefore, the trends in variability observed across the four waves could be attributed to a mixture of age, practice, and equipment box effects. For example, the increase in variability between waves 1 and 3 could be explained as the addition of the new testing device. Because the overall trend is upwards, it could be assumed that version 2 of the apparatus overestimated variability relative to version 1. The decrease between waves 3 and 4 could be attributed to practice effects (given that the same devices were used). Finally, the large decrease in variability from wave 4 to 5 could be due to a mixture of the new box and practice. The larger increases and smaller decreases in variability in cohorts 3 and 5 relative to cohort 1 could indicate that ageing effects, which increase variability, play some part in these groups. However, this explanation is very tentative and given insufficient data to test any of the predictions empirically, future investigations are required to shed light on this issue. It needs to be pointed out that despite the unexpected pattern of results obtained from each cohort across the four waves, which may not reflect the actual change in performance from wave to wave, cross-sectional differences at any one wave remain as would be expected. In other words, there seem to have been some systematic effects at different waves, but the effects were observed in all cohorts and the cross-sectional differences between cohorts are preserved. It is also of note that the effect of wave, and especially the difference between waves 4 and 5, is less pronounced in CRT than in SRT measures. Therefore, the proposed effects of RT device could be specific to the SRT test.

In CRT IIV, the youngest cohort showed a decline in CRT variability across waves, the middle cohort increased somewhat and stabilised, while the oldest cohort increased across the four waves. The increase in CRT ISD across waves was accelerated in the oldest cohort. This finding is consistent with the trends reported by Deary and Der (2005a) based on the first two waves. Maddox & Douglass (1974) found no evidence for an increase in CRT IIV in their study, but their sample was small and the measure of IIV was different to the ones usually used. MacDonald et al. (2003) also failed to detect an increase in CRT variability over time, but their measure was purified of mean RT differences between the age groups. Given that Deary and Der found that the age effects on RT IIV were attenuated by controlling for RT mean, it may be that longitudinal increases in RT IIV can be accounted for by increases in mean RT (i.e. general slowing of RTs). This issue is further discussed below.

Sex differences

Sex differences were found in all four measures of RT IIV (SRT ISD, SRT CV, CRT ISD, and CRT CV). Females were consistently more variable than males, which is in agreement with various previous reports (Deary & Der, 2005a; Der & Deary, 2006; Reimers & Maylor, 2005; Chapter 3 in this thesis). In all three cohorts sex differences in SRT ISD and in SRT CV varied with wave. They were larger at waves 3 or 4 and smaller at waves 1 or 5 which, again, could be explained by different versions of the RT testing device used between waves 1, 3-4, and 5. For both CRT IIV measures, there was an interaction between cohort, sex and wave. Sex differences on average increased across waves in the youngest cohort, but not the middle or older cohorts.

There was no significant sex difference in CRT IIV at age of about 15 (which was reported in an earlier study using the first two waves of RT data; Deary and Der, 2005a), whereas females were almost uniformly more variable than males at later ages. This pattern of findings is consistent with Deary & Der's proposition that sex differences in CRT IIV may occur only some time after puberty. It also confirms

findings from cross-sectional data reported in Chapter 3 and in Dykiert, Der, Starr, & Deary, in revision).

Raw versus mean-adjusted measures of RT variability

The effects found for raw and mean-adjusted measures were similar within tasks. In other words, although the obtained models were slightly different for SRT than CRT, the same patterns of effects and interactions were found for ISD as for CV in each task. Cohort (but not sex) effects appeared to be attenuated when mean-adjusted IIV measures were used in both SRT and CRT. This was manifested in the reduction of the number of significant between-cohort differences. Specifically, in both SRT and CRT all cohorts were significantly different from each other on ISD (with the youngest cohort least variable and the oldest cohort most variable). However, when CV was used instead of ISD, the two younger cohorts were no longer significantly different from each other in SRT variability and only the youngest and older cohort differed in CRT variability. This finding is consistent with the previous reports of attenuation of (cross-sectional) age effects on RT IIV when mean RT is controlled (e.g. Der & Deary, 2006; Chapters 2 and 3 of this thesis). However, contrary to findings of Shammi, Bosman, and Stuss (1998), there was still a significant difference in RT inconsistency between the youngest and oldest groups when mean RT was controlled. The pattern observed in the first two waves by Deary and Der (2005a), whereby the middle cohort had lowest mean-adjusted variability in SRT, was not replicated when all four waves were considered in the present investigation.

The findings were slightly different when longitudinal effects of age (i.e., wave) were considered. In the present study, wave can be considered a proxy for age, because the cohorts had narrow age ranges and, within each cohort, chronological age was highly correlated with wave number. The effect of wave on SRT IIV was not attenuated by using CV instead of ISD. All pairwise comparisons remained significant and the effect sizes for wave-to-wave differences obtained using ISD and CV were similar (see Table 4.2). For CRT, the effect was somewhat more complex.

Variability increased from wave 1 to wave 3 when CRT ISD was considered, and this was also the case for CRT CV. However, when the CV was used, there was a decrease in variability from wave 4 to 5, which was not found with CRT ISD. Consistent with this observation, effect sizes obtained from CRT ISD and CRT CV were similar for differences between the first 3 waves, but different in magnitude and direction for difference between waves 4 and 5 (see table 4.2). It is of note that the significance value for the decrease in CRT CV between waves 4 and 5 was small and would not be considered significant if correction for multiple comparisons was applied. In addition, from the visual inspection of the means (Figure 4.2), it seems that the only notable attenuation of variability in CRT by using a mean-adjusted measure, was at the last wave in the oldest cohort. On the whole, it appears that using a mean-adjusted measure of RT IIV does not attenuate the effects of age when it is considered longitudinally, until the age of about 75.

Strengths and limitations

A clear strength of the study was its representative initial samples and a relatively long follow-up period of 20 years, and a cross-cohort span of 60 years. The design allowed an investigation of both cross-sectional and longitudinal effects of age, and a comparison of these effects within the same group of people.

However, the longitudinal nature of the study as well as the length of the follow-up pose a set of limitations at the same time. For example, practice is one of the major problems in longitudinal studies even when intervals between measurements are large (e.g. Rabbitt, Diggle, Smith, Holland, & McInnes, 2001). Improving performance with each testing occasion may mask ageing effects. Therefore, it has been suggested that practice and age should be modelled separately to disentangle their respective effects (e.g. Ferrer, Salthouse, Stewart, & Schwartz, 2004). However, in this study, age and wave were very highly correlated, and it was not possible to separate the effect of practice from the effects of age. Consequently, the age effects obtained in the oldest cohort may be underestimated. On the other hand, the apparent decrease in CRT ISD in the youngest cohort may reflect practice rather than genuine decrease in variability with age in that group. Finally, the lack of

longitudinal change in the middle cohort may result from different combinations of practice and ageing. For example, it may be that CRT ISD increases in mid-adulthood but is off-set by practice effects, or it may be that there is no age-related increase observed in that age group and neither are practice effects. Without separately modelling practice and age effects, it is unclear which scenario is more plausible.

As is the case with all longitudinal studies, participant drop-out was a problem. Two different approaches may be used to deal with longitudinal data affected by drop out, each with different issues. If only participants with complete data are considered, as was the case in the present study, attrition acts to decrease the sample size and to reduce its initial level of representativeness. If all participants are used but not all contribute to all waves, the interpretation of trends over time becomes questionable. It is known that participants do not drop out at random. For example, in an earlier investigation of RT from the waves 1 and 3 of the Twenty-07 study, Deary and Der (2005a) found that RT (from wave 1) of participants who took part in wave 3 were faster than those of individuals who dropped out. MacDonald et al. (2003) showed that drop-outs had more variable SRT than those who completed all waves, and that the differences in inconsistency between drop-outs and returnees (in both SRT and CRT) increased with age. Returnees for subsequent waves of testing in longitudinal studies are also more likely to be females, to come from higher occupational class and to have higher intelligence than people who do not (Dykiert, Gale, & Deary, 2009). Dykiert et al. demonstrated the practical importance of this differential drop-out by showing that it may lead to biased estimates of sex differences in general intelligence. Therefore, the data available from later waves of longitudinal studies will be from the most selected sample (over-representing higher social class, higher levels of ability and females), which may be no longer comparable with the initial sample. If age differences are of interest, any that are observed will be affected (decline tempered and improvement exaggerated) in particular by the selective drop-out of participants with lower levels of ability. In this study, only participants with complete data from all four waves were considered, which allows valid comparisons between waves to be made, but reduces the degree of sample representativeness. Consequently, although the initial sample was

representative of the population from which it was drawn, the effects observed with the restricted sample of completers, may not be generaliseable to this population. However, as discussed before, the observed age effects are likely to be underestimated rather than overestimated.

Another problem associated with the longitudinal design in this particular study was the change in the RT apparatus between waves 1 and 3 and between waves 4 and 5. This was the most likely contributor to an increase in SRT variability between the first two waves and a drop between the last two waves in cohorts 1 and 2. Informed by the apparently greater effect of the RT testing device version for SRT rather than CRT, one idea was that the way pre-presses (i.e. responses made prior to the presentation of the stimulus) were dealt with differed between the RT apparatus versions. Pre-presses occur more commonly in SRT than in CRT tasks. Moreover, many of the pre-presses that do occur in the CRT task would be removed. This is because pre-presses will have chance-level accuracy, so in a 4-CRT 75% of them will be errors. These would not be included in the calculation of mean CRT or CRT ISD, which only use correct RTs. In this study it was established that the two versions of the box that were still available (2 and 3), did differ in the way they treated pre-presses. Whereas version 3 records them as RT of 0 ms, version 2 does not seem to allow pre-presses to occur. It should be noted that at wave 5 RTs of 0 ms were removed prior to the calculation of RT summaries. Therefore pre-presses would not affect the data directly. However, when they occurred, they could have affected subsequent responses which were included in the RT variables used in the analyses. It is not possible to determine whether this contributed to the peculiar decrease in SRT mean and SRT ISD between waves 4 and 5. However, it has been noted that CRT data seemed virtually unaffected by the issue. Therefore, although longitudinal trends in SRT have to be interpreted with caution, CRT data appear to be uncompromised.

Conclusion

In this study both cross-sectional and longitudinal effects of age on variability in SRT and CRT were investigated in a large and representative sample. Cross-

sectional findings were consistent with the majority of the available evidence: older adults demonstrate greater variability in both SRT and CRT than younger adults and adult females have greater SRT and CRT variability than males. Longitudinal trends, showed an increase in variability between that first two waves and, for the two younger cohorts, a marked decrease between the last two. This unexpected finding was explained in terms of RT apparatus changes between the waves. CRT data appeared less affected by the apparatus change. CRT ISD had a different longitudinal trajectory in each cohort: it decreased in the youngest, remained relatively stable in the middle one and increased in the oldest cohort. The trends were similar for CRT CV, but the increase in the oldest cohort was less marked than with the raw measure of CRT variability. The main limitation of the study was the inability to separate practice from age effects. Consequently, the magnitude of the age effect found is likely to be an underestimation.

Table 4.1. Means and standard deviations of reaction time (in milliseconds) for males and females, different cohorts and at different waves.

Group`	N	SRT mean		SRT ISD		SRT CV		CRT mean		CRT ISD		CRT CV	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Sex													
Male	2128	310.0	86.0	74.4	51.6	23.3	12.5	625.2	113.4	113.2	37.9	18.1	4.7
Female	2592	319.8	91.1	82.2	58.6	24.7	13.7	634.5	118.8	124.5	44.1	19.5	5.0
Cohort													
1 (youngest)	1460	287.0	61.1	68.5	46.0	23.1	12.4	545.6	69.9	101.5	29.9	18.5	4.4
2 (middle)	1948	314.7	86.2	75.8	51.7	23.3	12.4	624.5	90.7	117.6	35.2	18.8	4.7
3 (oldest)	1312	348.1	106.4	94.2	66.8	26.2	14.9	733.1	112.1	142.1	50.8	19.3	5.6
Wave[a]													
1	1180	310.7	77.7	70.5	45.2	22.2	11.1	627.2	95.5	115.0	31.1	18.4	4.1
3	1180	326.2	97.3	90.7	60.0	27.1	14.2	626.3	109.1	119.2	35.1	19.0	4.4
4	1180	316.8	93.6	82.8	55.3	25.4	13.3	622.6	113.4	119.3	36.5	19.1	4.5
5	1180	308.0	85.0	70.5	58.5	21.7	13.2	645.0	141.9	124.3	58.6	18.9	6.2

[a] Reaction time data were not collected at wave 2.
Note. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation

Table 4.2. Effect sizes (Cohen's d) for wave-to-wave differences in reaction time intra-individual variability measures.

Comparison	Intra-individual variability measure			
	SRT ISD	SRT CV	CRT ISD	CRT CV
Wave 1 - 3	-0.380	-0.384	-0.129	-0.158
Wave 3 - 4	0.137	0.125	-0.002	-0.019
Wave 4 - 5	0.217	0.278	-0.103	0.031

Note. Positive numbers indicate greater mean in the first wave of comparison relative to the second one;
SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation,
CV = coefficient of variation.

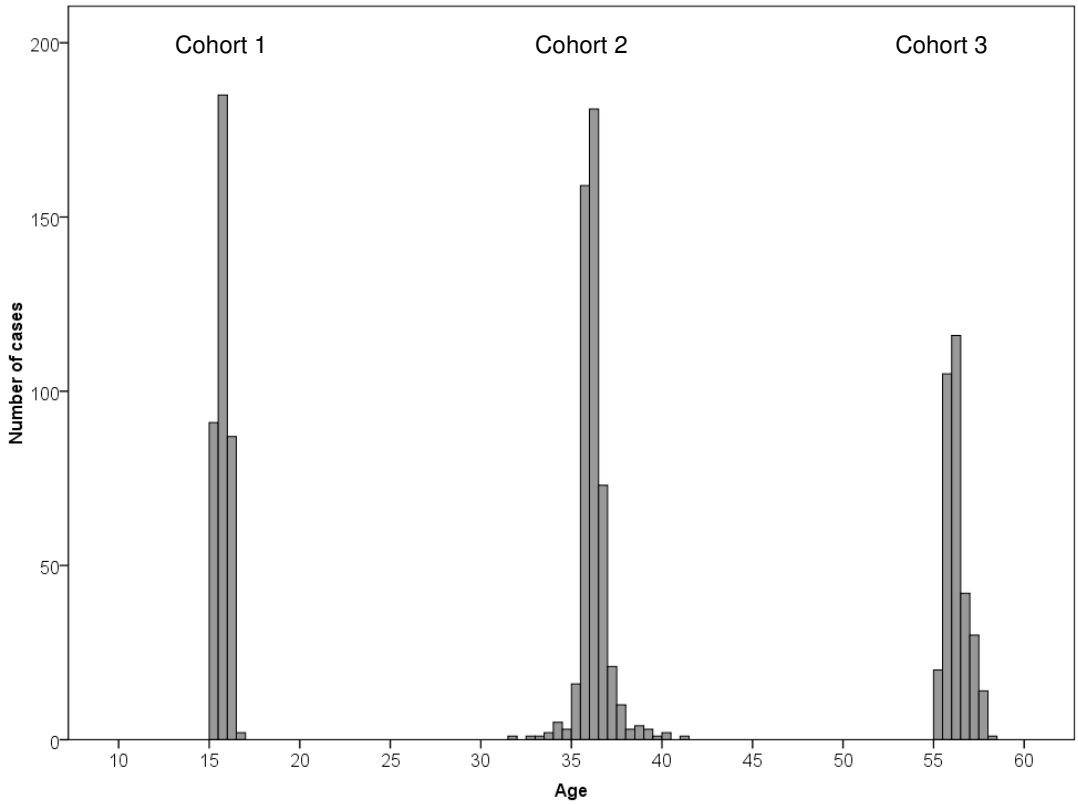

Figure 4.1. Distribution of ages (in 0.5-year bins) in each cohort at the first wave of testing.

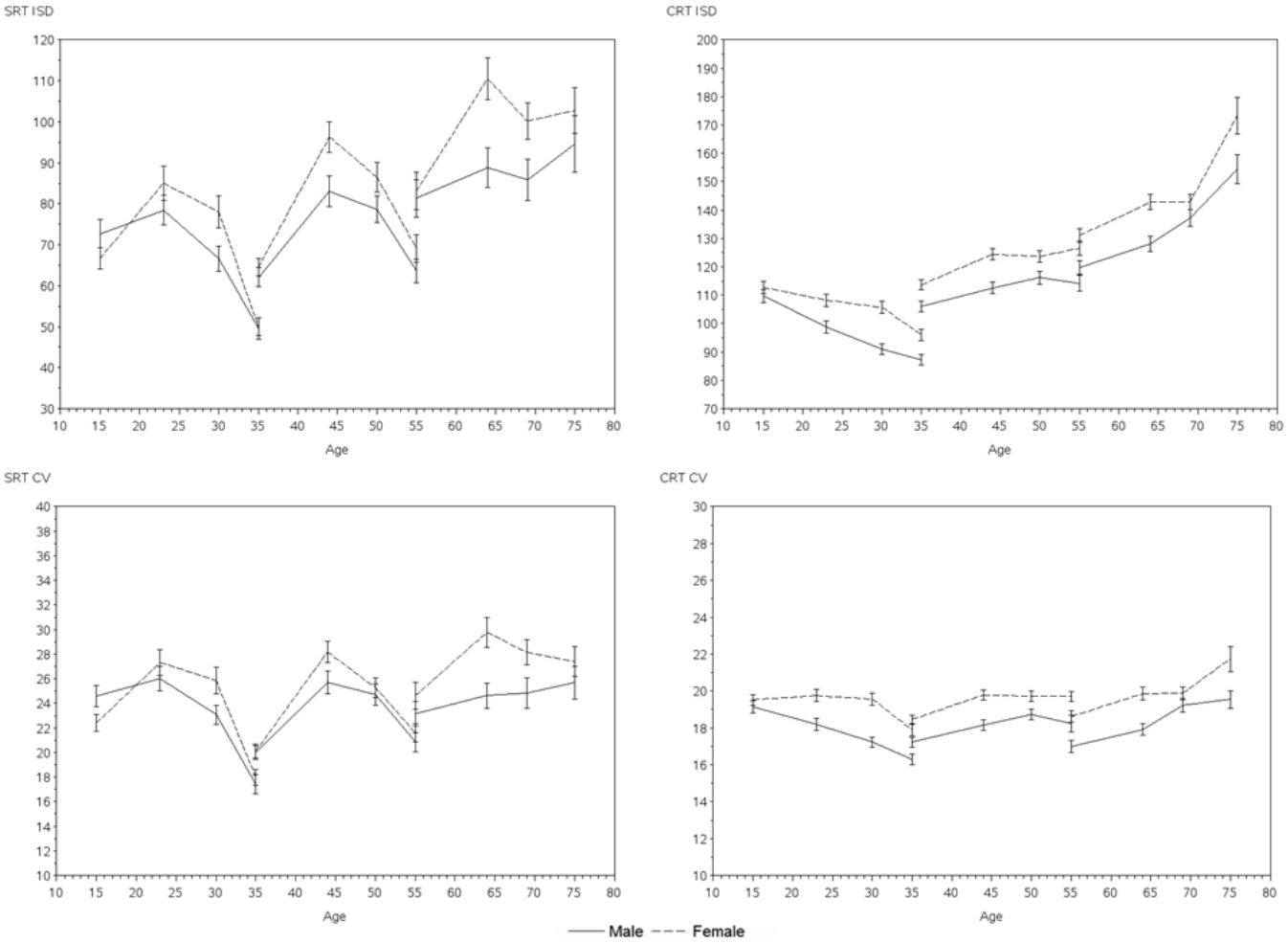

Figure 4.2. Means of reaction time intra-individual variability measures for males and females. SRT = simple reaction time, CRT = choice reaction time, ISD = intra-individual standard deviation, CV = coefficient of variation. Errors bars represent standard error of the mean; ISDs are in milliseconds.

Chapter 5 -
Sex differences in reaction time trial by trial

In Chapter 3 it was shown that reaction times (RTs) of adult females are characterised by greater intra-individual variability (IIV) than RTs of adult males, whereas the same is not true for participants younger than 18 years. As already discussed in Chapter 3, one explanation that has been proposed for this pattern of findings, is in terms of estrogens' effects on the brain (Deary & Der, 2005a). Although this hypothesis has not been tested directly, it is based on a plausible mechanism bringing about sex differences in post pubescent participants. However, an alternative explanation for greater RT IIV in females has been suggested by Reimers and Maylor (2006) and this shall be addressed in the present chapter.

Reimers and Maylor (2006) suggested that RT performances of males and females differ in the way they change over a number of trials. They reported findings from a large sample (n = 5,137) of participants aged 15 to 66 years, who completed an on-line gender classification task. The tasks involved 12 trials on which one of four faces (two male and two female; for each sex one face had a happy expression and one had a sad expression) was briefly presented in the middle of a computer screen. The faces remained on the screen for approximately 175 ms. In the course of the task each face was presented three times, in random order. The participants were required to respond by pressing one of two keys depending on whether the face was male or female (the emotion expressed by the faces was not relevant to this task and was to be ignored).

In their study, Reimers and Maylor (2006) found no sex difference in mean RT, but females had significantly greater RT intra-individual standard deviation (ISD) than males. They went on to investigate RT performance across the 12 trials of the task and found a significant sex by trial number interaction, indicating that sex differences changed across trials. The difference between the sexes was largest on

the first trial, where females had longer RT than males; there was little difference apparent on trial 2, and from trial 3, females actually seemed to have slightly shorter RTs than their male counterparts (although this observation was not accompanied by a significance test). From this pattern of sex differences Reimers and Maylor proposed that RTs of females are not intrinsically more variable, but rather that they are disproportionately slower at the first trial, which inflates the overall variability of females. When Reimers and Maylor re-analysed the RT IIV data having excluded the initial two trials (i.e. used RT ISD calculated from trials 3-12 only), there was no longer a significant sex difference or sex by trial interaction. This demonstrated that greater IIV of females in Reimers and Maylor's sample was due to greater change (decrease) in RT of females in the initial trials relative to males.

However, there were a number of major differences between the tasks used by Reimers and Maylor and those used by Deary and Der (2005a) and Der and Deary (2006). The task differed in the stimuli used (faces vs. digits), practice (no practice vs. 8 trials) and the mode of testing (internet vs. face-to-face) to name a few. Therefore, the findings reported by Reimers and Maylor may not be applicable to a choice reaction time (CRT) task used by Deary and Der. Individual trial data were not available from datasets used in the earlier studies that investigated sex differences in RT (Deary & Der, 2005a; Der & Deary, 2006). The apparatus used to collect RT data in Twenty-07 (Deary & Der, 2005a) and Health and Lifestyle Survey (HALS; Der & Deary, 2006) provided only summaries of each participant's RTs (means and SDs, and the number of errors) across all trials in the test. Therefore, the hypothesis proposed by Reimers and Maylor (2006) could not be tested in those data sets. However, the RT apparatus used in the HALS and Twenty-07 studies has since been upgraded, and the new version records and stores RTs from each trial, thus allowing an investigation of trial-to-trial effects.

The current study attempted to replicate Reimer and Maylor's (2006) findings from a gender classification task on a CRT task, in which there were no sex difference in mean performance, whereas females demonstrated greater IIV than males. The data presented here were collected during UK Science Festivals, using an upgraded version of the RT apparatus, which was originally designed for and used in the HALS study.

Method

Participants

The participants in this study were a subgroup of the sample described in Chapter 3; they were recruited from among visitors to Science Festivals held in Cambridge, Edinburgh and Cheltenham in 2008, and Edinburgh in 2009. Of the 2,392 records, 398 were excluded because of missing data, high error rate or following data trimming described in detail in Chapter 3. Out of the 1,994 remaining participants (age range 4 to 75 years), all those aged below 15 or over 66 years were excluded, for consistency with Reimers and Maylor (2006). Because the Science Festivals were aimed primarily at children, this age group constituted a bulk of the sample. Therefore, the exclusion of individuals below the age of 15, eliminated a large proportion of the original sample. The resulting working sample included 374 participants, including 157 males and 217 females.

Reaction time measures

Four-choice RT was measured using a specially-designed portable RT apparatus, which was an upgraded version of that used in the UK HALS study (Cox, Huppert, & Whichelow, 1993). Full description of the device and of the CRT task is provided in Chapter 3 (see also Figure 3.2 for an illustration of the RT apparatus).

In the previous analysis (Chapter 3) only correct responses were used for the calculation of mean CRT and ISD. In the present investigation, for comparability with Reimers and Maylor, the summaries were calculated from all trials, including those on which errors were made (but, those with aberrantly fast or slow responses were excluded as described in Chapter 3).

Statistical analysis

All analyses were performed using SAS software (version 9.2, SAS Institute). Summary measures of CRT—that is, CRT means and CRT ISDs—were analysed using between-subjects ANOVAs (PROC GLM). Because of the RT trimming, a proportion of the sample had missing data on one or more individual trials. Therefore, to analyse trial-by-trial effects, a mixed models approach (PROC MIXED) was used, which can deal with missing data better than can GLM procedures.

Results

Results relating to the effects of age and sex on CRT mean and ISD have been reported elsewhere (Chapter 3; Dykiert et al., in revision). For the purpose of this study I verified whether sex and age effects, as reported previously, apply to the restricted age range (15 – 66 years) and to CRT summary measures calculated using both correct and error responses. Means and SDs of mean CRT and CRT ISD for each sex and age group are presented in Table 5.1. Two between-subjects ANOVAs were used to tests the effects of age group (15-19, 20-29, 30-39 40-66), sex (male, female) and their interaction on CRT mean and CRT ISD. For mean CRT there was a main effect of age group, $F(3,370) = 28.93$, $p < .001$, with the oldest group slower than the two middle groups (no other pairwise comparisons were significant); there was no significant sex effect or sex by age interaction. CRT ISD was a function of age group, $F(3,369) = 8.33$, $p < .001$, and sex, $F(1,369) = 9.71$, $p = .002$; there was no interaction between age group and sex. The oldest group was significantly more variable than all younger groups, but the other groups were not significantly different from each other. Females had more variable CRTs than males.

Mean CRTs of males and females on each trial of the task are shown in Figure 5.1. The overall pattern of RTs was irregular, with notable (and systematic) variations in RTs from trial to trial. Given that the sequence of stimuli and of the PIs were identical for each participant, these variations are likely to be due to the specific

combinations of stimuli and PIs on each trial. It is apparent that the variations were consistent between sexes: both males and females showed a similar pattern of differences in CRTs from trial to trial. For example, for both sexes, there is a drop in RTs from trial 1 to trial 3, followed by an increase in RT on trial 4. Local maxima (e.g. trial 11) were also observed in both sexes. On the whole, CRTs of females were either similar to or larger than CRTs of males, with few trials on which females had shorter responses. Despite the large variations from trial to trial, it is clear from Figure 5.1 that the responses of females were around 40 ms slower at the beginning as well as at the end of the block; there is little evidence of a decrease in the difference between males and females across trials.

To test these observations, a linear mixed model was fitted to the CRT data. If, as suggested by Reimers and Maylor (2006) males and females differ in the performance across trials then there should be an interaction between sex and trial number. The initial model included sex, age, trial number and their interactions, as fixed effects (trial was entered as a linear term rather than a factor, to allow modelling of trends across trials rather than of trial-to-trial differences). Also, the following covariates were included: stimulus, stimulus change (same or different to the stimulus on the previous trial), errors (was a response on a given trial correct or erroneous) and a combination of the two immediately preceding PIs (indicating change in PIs between previous and current trial). The model was reduced by eliminating non-significant terms in a backward stepwise manner. The resultant model included sex, age, trial number and age by sex interaction, as well as error, stimulus, stimulus change and PI combination. Because residuals from the preliminary models were positively skewed, a normalising transformation was determined using the Box-Cox procedure (Box & Cox, 1964; see also Chapter 3). The data were transformed prior to final analysis.

When the model was re-fitted on the transformed data, the pattern of fixed effects remained unchanged; that is, CRT was longer for older participants and increased with trial number. CRT was shorter on trials with errors and on trials on which the same stimulus was repeated. Previous-current PI combination and stimulus also had significant effect on CRT. Because this study was not designed to test the effects of different stimuli, PIs and their sequential organisation, not all combinations

of these were represented in the CRT task. These variables were included in the models to control for their effects, but it was not deemed appropriate to attempt to investigate their independent contributions to RTs. It was expected that inferences made based on post-hoc analyses could be biased and therefore none were performed. However, some observations made are mentioned in the discussion below.

Once the fixed effects structure was established, random effects were introduced (intercept and slope on trial), which reflect individual differences in the initial level of performance and in the slope of change across the block of trials. Information criteria suggested that a model with unstructured variance-covariance matrix fitted the data best. When random effects were included in the model, two of the fixed effects were no longer significant (sex and age by sex interaction). These were removed to obtain unbiased estimates of the remaining effects, which are presented in Table 5.2. In terms of random effects, there was significant variance in intercepts, $p < .001$. The key finding of this analysis is that there is no sex by trial number interaction, which suggests that trial-by-trial performance on a CRT task is similar for males and females.

Discussion

Main findings

The finding of no sex difference on mean CRT and a greater female variability found previously (Chapter 3) were also observed in the narrower age range (15-66 years). Contrary to the finding of Reimers and Maylor (2006), this study has not revealed a pattern of reversed sex difference in the initial trials. There was no interaction between sex and trial number, suggesting that male–female difference remains relatively stable across a RT testing block. This finding, together with the observation of higher female means at the beginning *and* at the end, is in opposition to the findings of Reimers and Maylor, who found that females responded more slowly than males on the first trial but, on average, faster than males from trial

3 until the end of the block (trial 12). To reiterate, although females demonstrated greater degree of variability across trials, this was not due to their disproportionately longer RT on the initial trials.

In search for explanations of the discrepancy in the present and Reimers and Maylor's (2006) findings, a few alternative explanations in terms of procedural differences between the two studies have been considered. First, Reimers and Maylor's (2006) participants were given no practice and the first trial in the data was the very first test trial the participants encountered. On the other hand, in the present study as well as in Deary and Der (2005a) and Der and Deary (2006), participants were given eight practice trials before the data collection commenced. The data from practice trials were not available for analysis, as they were not recorded by the apparatus. It is possible that if the first trial of practice was available, we would observe a pattern like Reimers and Maylor; that is, a large difference in favour of males (females slower) on the first trial, which diminishes rapidly over the first few trials. However, it needs to be noted that greater CRT ISD among females after practice was observed in the present study, suggesting that even if the mechanism observed by Reimers and Maylor is operating, it does not explain the sex difference in CRT variability in the Science Festival sample or in the studies carried out by Deary and Der, which used the same apparatus and procedure (Deary & Der, 2005a; Der & Deary, 2006). If one considers that the 40 trials in the tests followed the 8 practice trials, the data from the present study could be envisaged as a block of 48 trials with 8 first trials removed. It then follows that if the effect observed by Reimers and Maylor was present in the data obtained using our apparatus, sex differences in IIV should not be found after "removing" the initial trials. But Deary and Der consistently found greater female variability (after practice) from three independent samples (Deary & Der, 2005a; Deary & Der, 2006; Dykiert, et al., in revision; see also Chapter 3). Therefore, the explanation for the discrepancy in our findings and those of Reimers and Maylor is likely to lie elsewhere.

Second, the stimuli were different in the present study and that of Reimers and Maylor (2006). The CRT test was relatively simple: it involved four numerical stimuli and each stimulus had its assigned (and compatible) response button. The RT test chosen by Reimers and Maylor, used faces as stimuli and the task was to

categorise them according to gender. There were four stimuli and only two response keys, one for each sex. The fact that the stimuli were faces may be of importance given the established sex differences in face recognition. Females have been reported to have better accuracy at face recognition, especially for female faces (Lewin & Herlitz, 2002; McBain, Norton, & Chen, 2009), and shorter RTs (Godard & Fiori, 2010). The latter could explain Reimers and Maylor's observation that females had, on average, slightly shorter RT on trials other than 1 and 2. However, their finding is not in line with most literature on gender differences in mean RT which usually reports shorter RTs in males (e.g. Der & Deary, 2006; Lahtela, Niemi, & Kuusela, 1985). Although the nature of the task used by Reimers and Maylor is different than gender recognition, the fact that the same faces were repeated within a block of trials, some of them could be recognised at subsequent presentation (with greater success in females than males). Therefore, it is possible that the effect observed by Reimers and Maylor is specific to a test with face stimuli and therefore not replicated in CRT.

Third, the RT tasks used in the two studies differed in the stimulus presentation procedure. In the CRT task used in the present study, the stimulus remained visible on the screen until a response was made. By contrast, Reimers and Maylor (2006) presented their stimuli for a brief, pre-determined period of time (175 ms). Because of this, Reimers and Maylor's task alluded to an inspection time (IT) paradigm, in which the stimulus presentation time is tightly controlled. This could explain why their results were not replicated in the CRT task, in which there were no limitations on the length of stimulus presentation. However, it is difficult to attempt to envisage whether this brief presentation of the stimuli would benefit males or females, because there is little consensus in the literature on the direction, or even the existence, of sex differences in IT. Indeed, both male and female advantage have been reported (Bonney et al., 2006; Codorniu-Raga & Vigil-Colet, 2003, Luciano et al., 2001; Pesta, Bertsch, Poznanski, & Bommer, 2008). However, only one study examined the issue in a representative sample, and that one found no significant difference between the sexes (Burns & Nettelbeck, 2005). However, even if there were sex differences in average performance in IT paradigms, this would still not lend explanation to the greater female variability.

It should be noted that the length of stimulus presentation used in Reimers and Maylor (175ms), would almost certainly be sufficient to produce 100% accuracy in an IT task and, although face stimuli are much more complex than the stimuli used in classical IT tasks (i.e., two vertical lines of different lengths), this procedure would not have necessarily affected the accuracy of classifications. In fact, low overall error rates in Reimers and Maylor's study suggest that it did not. However, one possibility is that males and females tend to use different strategies in RT tasks. Reimers and Maylor themselves suggest that perhaps males and females differ in the speed-accuracy trade-off, with females initially slightly slower and more accurate than males. If females are inclined to adopt a strategy which is less optimal for a test with limited presentation time than males, it may need to be adjusted at subsequent trials. If this is the case, then females would show a more marked decrease in RT, compared to males.

Finally, Reimers and Maylor (2006) collected their data via the Internet, whereas the Science Festival data collection was carried out face-to-face. This is another possible source of the differences in the findings of the two studies. The presence of a person administering a RT test ensures that each participant is ready before the task is initiated, which is not guaranteed with a remote data collection. Another weakness of data collection via the Internet is lack of control of the participants' environment during the task. In addition, there is no way of verifying the genuineness of the information provided and the integrity of the data. Reimers and Maylor adopted some commonly used techniques for minimising these problems, for example they removed multiple entries from the same IP address, or cases with ages that were likely to be false. These procedures improve the overall quality of the remotely collected data, but even with these in place the comparability of on-line and off-line tests may not be assured. There are examples in the literature which suggest that the same tests presented on-line and off-line do not produce equivalent results (e.g. Buchanan et al., 2005). Among the reasons for the differences are not only lack of experimental control and decreased social desirability in web-based testing, but also different sample make-up (see Buchanan, 2002; Buchanan & Smith, 1999). The latter could have a major role in explaining the discrepancies between the present study and Reimers and Maylor's study.

Other findings

The effect of age on CRT has been discussed in Chapter 3 and only other findings which pertain to trial-by-trial effects will be considered here in detail. Firstly, the CRT data followed the well-established trends in terms of accuracy and stimulus repetition (for reviews see Welford, 1980; and Kirby, 1980)—they were shorter on trials with errors and on trials on which the stimulus from the previous trial was repeated. There was no practice effect apparent over the 40 trials. Instead, CRT showed a slight *increase* over the block of trials, perhaps suggesting increasing fatigue or decreasing motivation with time on task. There was no trial by age or trial by sex interactions, so the effect was similar for different ages and both genders.

There were large variations in CRT from trial to trial, which were similar for both sexes and ages. These could be largely attributed to specific combinations of PIs and stimuli, which were different from trial to trial but the same for all participants. For example, it is well-established that PIs have a considerable effect on the RTs (see review in Jensen, 2006). Although stimulus and the combination of PI were controlled in the analyses, their effects could not be properly investigated, because not all combinations were represented in the CRT task used. For example, CRT appeared to be longest in response to stimulus '3', a response to which was made with right index finger. Since the majority of the participants are likely to be right handed, responses made with the right index finger should be expected to be among the fastest. However, on closer inspection it became apparent the effect was due to a combination of PIs and stimuli repetition occurring for different stimuli. For the same reason, combination of PI could not be fully investigated. However, from a visual inspection of CRTs at different PI combination in Figure 5.2, it appears that CRT was longer on trials with a large previous-current PI decrease (i.e. 3 s followed by 1 s) compared to other trials. There was also an apparent decrease in CRT on a trial with repeated PI of 2 s, however on this trial the stimulus was also the same as on the previous trial, which added to the effect.

Concluding remarks

In the present study, the pattern of initially slower female responses followed by slightly faster responses on subsequent trials reported by Reimers and Maylor (2006), was not replicated. Moreover, sex difference in CRT IIV was confirmed even in the absence of the mentioned pattern, demonstrating that greater female variability is maintained across trials. Reimers and Maylor's finding, which may reflect task-specific (gender classification) or medium-specific (internet-based research) effects, is nonetheless interesting and warrants further research.

Table 5.1. Group means and standard deviations of choice reaction time measures (in seconds) for males and females and different age groups.

Group	N	Mean CRT		CRT ISD	
		Mean	SD	Mean	SD
Sex					
Male	157	0.520	0.078	0.087	0.025
Female	217	0.525	0.071	0.094	0.025
Age group					
15-19	49	0.495	0.057	0.089	0.024
20-29	58	0.480	0.053	0.081	0.021
30-39	85	0.497	0.059	0.087	0.023
40-66	182	0.555	0.076	0.097	0.027

Note. CRT=choice reaction time, ISD=intra-individual standard deviation

Table 5.2. Summary statistics of the final linear mixed model

Parameter	Estimate	SE	F	*p*
Intercept	-0.9505	0.02340		
Age	0.005661	0.000548	106.63	<.001
Trial number	0.001912	0.000185	107.08	<.001
PI Combination			52.63	<.001
30	0.1202	0.01284		<.001
1.0→1.5	-0.06826	0.009822		<.001
1.0→2.5	-0.01782	0.009019		0.050
1.0→3.0	0.02250	0.009996		0.024
1.5→1.0	-0.02545	0.009004		0.005
1.5→1.5	-0.00221	0.009022		0.807
1.5→2.5	0.005976	0.009492		0.529
2.0→1.0	0.03725	0.009920		<.001
2.0→2.0	-0.1111	0.009556		<.001
2.0→3.0	0.01638	0.008626		0.058
2.5→1.5	-0.01624	0.01075		0.131
2.5→2.0	-0.00908	0.008557		0.289
2.5→3.0	0.03094	0.01167		0.008
3.0→1.0	0.08919	0.009787		<.001
3.0→1.5	0.02178	0.009675		0.024
3.0→2.0	0.02304	0.01001		0.021
3.0→2.5 (Reference)				
Error (error made)	-0.09909	0.007420	178.34	<.001
Stimulus			185.51	<.001
1	-0.06778	0.005107		<.001
2	-0.02188	0.005555		<.001
3	0.07037	0.005706		<.001
4 (Reference)				
Stimulus change (Stimulus repeated)	-0.07880	0.004201	351.82	<.001

Note. PI = preparatory interval, values are in seconds

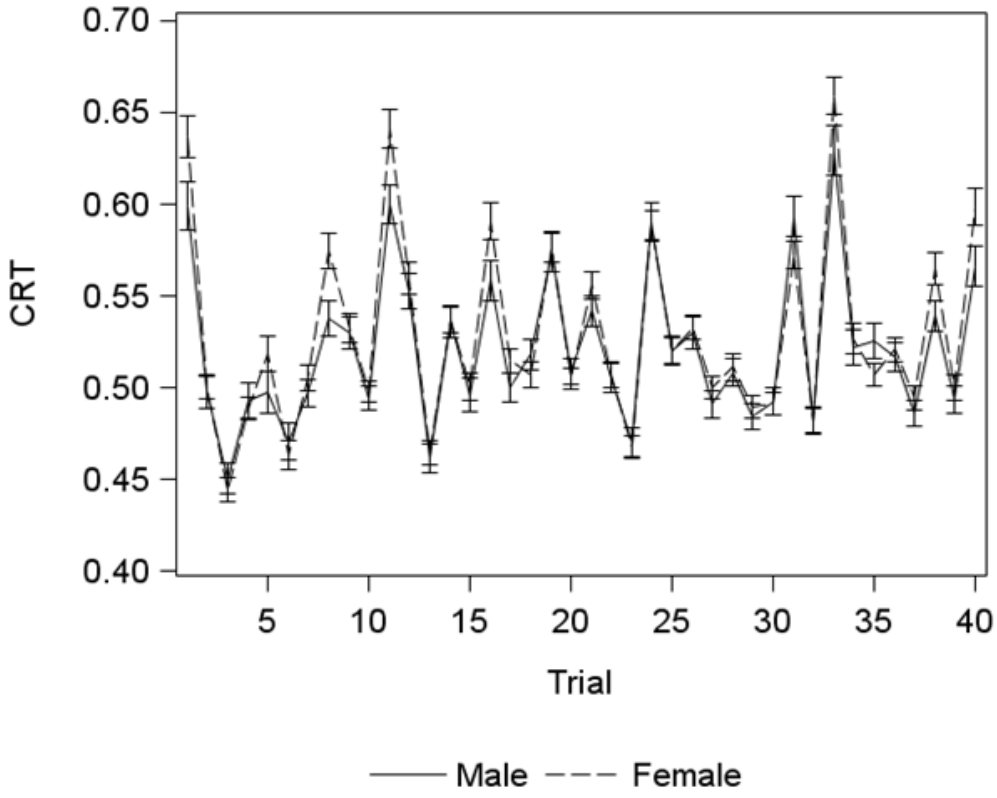

Figure 5.1. Mean choice reaction times (CRT) in seconds at each test trial for males and females. Error bars represent standard errors.

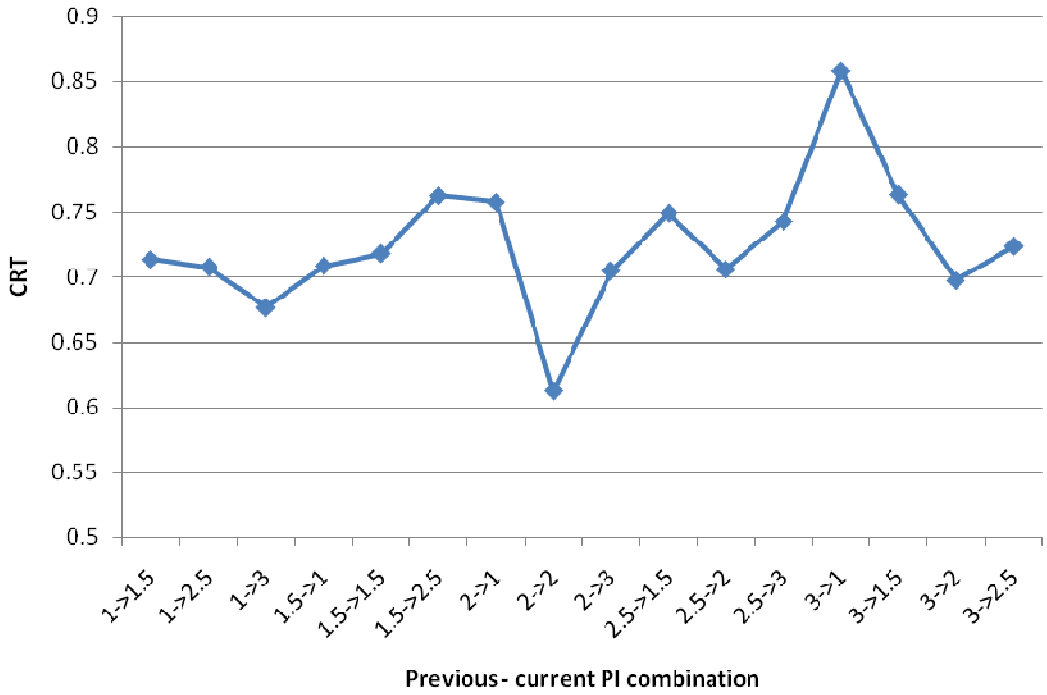

Figure 5.2. Mean choice reaction time (CRT) in seconds for each previous-current preparatory interval (PI) combination. PI values are in seconds

Chapter 6 -
The effects of high altitude on mean and intra-individual variability in choice reaction time

The hypoxic, hypobaric environment at high altitudes can lead to a spectrum of pathophysiological changes, including acute mountain sickness (AMS) and, in extreme cases, to high altitude pulmonary or cerebral oedema (HAPE and HACE respectively; West, Schoene, & Milledge, 2007). The earliest documented scientific investigations of the effects of altitude on human physiology and cognition were conducted by McFarland (1937a, b). Since then, numerous studies into the cognitive impairment associated with altitude have been carried out, including changes in psychomotor performance, perception, learning and memory (Bonnardel & Liberson, 1933; Regard & Landis, 1991; Virués-Ortega, Buela-Casal, Garrido, & Alcázar, 2004).

As mentioned in Chapter 1, reaction time (RT) is a well-established measure of cognitive functioning. Both mean RT and intra-individual variability (IIV) are (negatively) correlated with general intelligence (Deary, Der, & Ford, 2001; Jensen, 1992) as well as with performance on various specific cognitive tasks. Because the measurement of RT is non-invasive, fast, and relatively simple, it is very useful in the study of the effects of high altitude on cognition. Changes in RT in response to altitude have been investigated both in decompression chambers which control the oxygen fraction of inspired air (FiO_2) (Abraini, Bouquet, Joulia, Nicolas, & Kriem, 1998; Ledwith, 1970), and on actual expeditions to high altitude (Kramer, Coyne, & Strayer, 1993; Mackintosh, Thomas, Olive, Chesner, & Knight, 1988).

Increases in mean RT at altitude were originally believed to be the result of AMS. For example, Macintosh et al. (1988) showed that there was a positive correlation between RT and altitude, but that it was only significant in individuals

with severe symptoms of AMS. However, Macintosh et al. suggested that an increase in RT with altitude is a direct effect of hypoxia rather than a secondary effect of AMS. Furthermore, it has been suggested that there might be a threshold altitude above which the RT impairment starts to manifest itself. O'Carroll and MacLeod (1997) examined RT at relatively low altitudes (1,000 and 1,600 m) and found no evidence of worsening of RT following the ascent. Kramer et al. (1993) also failed to detect evidence of slowing of CRT during an ascent to 4,360 m. However, other investigators, who measured RT at higher altitudes, tended to find significant effects above a certain level. For example, McFarland (1937a) investigated simple and five-choice RT following a rapid ascent to 4,540 m by train or car. SRT remained unaffected at 4,540 m, while CRT was significantly slower at that altitude, compared to the performance at sea level (Cohen's d = 1.12). When ascent was more gradual, allowing for acclimatisation, CRT was significantly slowed at 5,340 m (Cohen's d = 0.70), whereas SRT remained relatively intact and only became significantly longer at 6,140 m (Cohen's d = 1.37), relative to their respective sea level baselines (McFarland, 1937b). McFarland's findings suggest that the speed of ascent may moderate the effects of altitude on RT.

It should be noted that there is considerable heterogeneity in design between the above-mentioned studies, which impedes the drawing of clear conclusions from their results. For example, in Mackintosh et al.'s (1988) study, all participants who performed the choice RT task were given either acetazolamide or methazolamide for AMS, and there was no placebo control group. Hence, the results may not be generalised to un-medicated climbers. The sample used in O'Carroll and MacLeod's (1997) study was also unusual. It consisted of professional rugby players, who may not be readily comparable to samples used in other studies. Moreover, the go/no-go RT task was different from those used commonly in other studies and required responses only to targets and inhibition of responses to non-target stimuli.

Evidence from laboratory studies using hypobaric chambers is no more consistent. Denison, Ledwith, and Poulton (1966) reported longer RTs in a complex spatial orientation task at 1,525 and 2,440 m relative to the sea level (Cohen's d = 4.31), but only when the participants were not familiar with the task. If the testing at simulated altitude was performed after a session at sea level, there was no significant

difference in RT performance between the sea level and altitude. This finding was not replicated by Paul and Fraser (1994), who used the same spatial orientation task. However, Paul and Fraser did find that RT in the spatial orientation task increased with altitude. They further found, using a 5-choice RT task, that CRT had a quadratic relationship with altitude, increasing at first (at 2,440 m) and then falling to the baseline level at 3,050 and 3,660 m, a finding similar to that reported by Ledwith (1970). In another study, using an auditory go/no-go paradigm, simulated altitude (4,300 m) was found to be associated with longer RTs in the morning and afternoon but not in the evening (Wesensten et al., 1993). It should be noted that, given the design of Wesensten et al.'s study, the finding could be due to acclimatisation effects or diurnal changes in performance or, perhaps, both. A study by Gerard et al. (2000) investigated RT obtained from a psychomotor vigilance task and found that higher simulated altitude was associated not only with an increase in mean RT, but also with an increase in the number of lapses, i.e. RTs slower than 500 ms. Finally, when using a 2-choice RT, at a gradual "ascent" to a simulated extreme altitude, no differences were found between performance at 8,848 m and the sea level (Abraini et al., 1998). This finding could be taken to suggest lack of effect of hypoxia on CRT. However, it could also be that the relatively simple two-choice RT was too easy a task to be affected (see Kobrick & Dusek, 1970, for evidence of significant detrimental effect of simulated 4,000 - 5,200 m altitude on a 48-choice RT task performance). In another study, based on the same data, Bolmont, Thullier, and Abraini (2000) suggested that RT performance may be influenced by mood states. They demonstrated that RT was positively related to state anxiety and negatively with tension, hostility and confusion. One limitation of the latter study was that it used simulated altitude with participants living in the hypobaric chambers over 31 days. Such a long confinement could result in mood states different to those occurring during a real high altitude expedition.

Both laboratory and field studies of the effect of altitude on RT have their unique strengths, but also weaknesses. Decompression chambers enable good experimental control, with precise manipulation of variables of interest and development of a true control condition. However, results obtained from such studies may not be applicable to real altitude. Firstly, often the only variable meant to mimic

effects of altitude in hypobaric chambers is the FiO_2, which is altered to reflect that expected at a given altitude. During an actual climb to a high altitude, individuals are exposed to a number of other variables, such as changes in temperature, physical fatigue, or even change in psychological state, such as excitement levels, all of which can have an impact on RT performance as well as the level of oxygen. Secondly, experiments in hypobaric chambers require participants to remain in a confined space for a period of time, sometimes over a number of days or weeks. Such confinement may have a detrimental effect on performance (decrease in fitness levels, psychological discomfort; West, 1984) unrelated to hypoxia. Therefore, although well-controlled experimental studies of hypoxia on cognitive functioning are needed, field studies are also essential in order to investigate the changes occurring at high altitude. Therefore the present study examined the effects of altitude during an actual climb.

Whereas previous studies have mainly investigated the effect of altitude on RT mean, RT IIV is of interest in its own right. Only three studies were identified that investigated the effect of altitude on CRT IIV. Bonnardel and Liberson (1933) reported no systematic differences in either mean or IIV of visual or auditory RTs measured in Paris and in Jungfraujoch (3,457 m). However, their conclusions were based on data from only two subjects and, as such, may not be replicable. The second study was carried out by McFarland (1937a). Six participants performed a 5-choice RT task at sea level and after a rapid ascent to 4,540 m. Apart from an increase in mean RT (as discussed above), a 50% increase in RT IIV was noted between the two locations (Cohen's $d = 1.31$). Finally, one study examined the effects of altitude in ten acclimatised individuals (McFarland, 1937b). In this study the participants spent four months gradually climbing to higher altitudes. IIV in SRT and CRT was significantly affected at 5,340 m, relative to the sea-level performance, with large effect sizes (Cohen's *d*s: 3.60 and 7.49, respectively). Since then, no further evidence from an investigation of CRT IIV has become available.

One problem common to all these three studies is the reliance on simple pairwise comparisons between the sea-level performance and measurements taken at an altitude. This approach is simplistic and open to misleading interpretations. If multiple comparisons are performed without correcting for the increasing probability

of Type I error, as was done in McFarland (1937b), for example, the significance of differences may be questionable. Moreover, such approach does not enable a precise determination of thresholds at which the effects start occurring. To do so, an approach which allows for modelling the effects of altitude, including at heights at which measurements were not taken, needs to be employed.

Given the lack of consensus regarding the effects of altitude on mean RT and the shortage of good quality research examining its effects on RT IIV, the present study was carried out. It aimed to investigate the effect of high altitude on CRT mean and CRT IIV in a homogenous group of healthy lowlanders climbing to 5,565 m. In addition, given that RT IIV is highly correlated with mean RT, it is thought to have independent sources of variance as well as those shared with RT mean (Jensen, 1992). Therefore, the extent to which altitude effects on RT IIV are independent of RT mean is also investigated here.

Method

Participants

Ten healthy participants (6 males, 4 females) were recruited to the study. They were aged between 19 and 24 years (M = 21.5, SD = 1.4) and were members of the Edinburgh Altitude Research Expedition 2008 to the Western Himalayas. The participants were in good general health and did not take any prophylactic medication for AMS such as acetazolamide (Diamox). Nine participants completed all 25 testing sessions, while one individual was not fit for testing on one occasion and provided 24 sets of data.

The study was undertaken with the ethical approval of the University of Edinburgh College of Medicine and Veterinary Medicine Committee for the Use of Student Volunteers for Experimental Work, with the written informed consent of each subject, and in accordance with the principles of the Declaration of Helsinki.

Measures

Reaction time

Four-choice RT was measured using a specially-designed portable CRT box, (Eagle CRT box, Eagle Designs, Edinburgh, UK). The illustration of the device is given in Figure 6.1. The CRT box had an LCD display at the top, under which there were four small red lights. At the bottom there was a row of four pushbutton switches corresponding to the lights. At each trial, one of the four lights lit up and participants were required to extinguish it by pressing the corresponding button as quickly and as accurately as possible. The stimulus lights lit up and remained lit until a response had been made. Pressing a button terminated the stimulus display and started the next trial. The subsequent stimulus was presented after a randomly varied interval of between 1 and 3 s.

Each testing session consisted of 8 practice trials, followed immediately by 40 test trials. The sequence of stimuli was random for practice trials, whereas in the test trials a fixed quasi-random sequence was used. RT latency (time between stimulus presentation and response, in ms) and accuracy (number of correct trials) were recorded. Mean RT and intra-individual standard deviation (RT ISD) were calculated for correct RT latencies for each individual at every testing session.

On a few occasions mechanical malfunction of a response button prevented the recording of actual RTs. Such occurrences were noted after each test. The anomalous responses were removed and RT means and standard deviations re-calculated for the remaining responses. The problem occurred at 12 testing occasions with only one at a peak (Phitse La); overall, 21 single responses were removed. Otherwise, outliers were not removed as it was believed that an increase in the number of anomalously long responses could be one mechanism of increasing intra-individual variability in RT (Gerard et al., 2000; Rabbitt, Osman, Moore, & Stollery, 2001).

RT ISD data were positively skewed. Therefore, for the purpose of the subsequent analyses, log base 10 transformation was performed to normalise the distribution.

Vital signs

Peripheral blood oxygen saturation (SpO_2, %) and pulse rate (beats/min) were obtained using a fingertip pulse oximeter (Shenzhen Creative Industry Co. Ltd., Shenzhen, China). Respiratory rate (breaths/min) was counted manually. Body temperature (ºC) was measured orally using a digital thermometer (Prestige Medical, Blackburn, UK).

Lake Louise score

The Lake Louise AMS scoring system (Roach, Bärtsch, Oelz, & Hackett, 1993) was used to assess physiological disturbance associated with altitude. This comprises five self-report questions (assessing headache, gastrointestinal symptoms, fatigue, dizziness and poor sleep) and 3 clinical assessments (change in mental status, heel-toe walking and peripheral oedema). Each self-report element was scored on a scale from 0 (no symptom) to 3 (severe and incapacitating). Change in mental status was scored from 0 (no change) to 3 (semiconscious; defined as <14 out of 15 points on the Revised Glasgow Coma Scale), heel-toe walking from 0 (no problem) to 4 (unable to stand) and if peripheral oedema was present, 1 or 2 points were given (depending whether it was present in one or more locations). AMS is deemed to be present if four criteria are fulfilled: a Lake Louise AMS score of 5 or more, the presence of headache, the presence of one other symptom described above, and a recent gain in altitude.

AMS had a low prevalence in the sample, with only seven individual instances of AMS recorded throughout the expedition. Given the low frequency of the syndrome, it was not included in the analyses. Instead, the Lake Louise Score was used as a surrogate measure of physiological disturbance at altitude.

Altitude

Altitude measurements were obtained by two independent global positioning system (GPS) devices (eTrex Venture HC and Foretrex 101, Garmin Ltd.,

Southampton, UK) and verified by cartographic spot height data where appropriate (Olizane, ISBN: 2880863171).

Time

Measurement occasion rather than elapsed time was used as a time variable because it was expected that the number of exposures to the RT task was the main factor affecting the learning of the task. Moreover, the testing sessions have been relatively evenly spaced with most measurements carried out daily, approximately 24 hours apart. In the course of the expedition, there were only three days on which testing was performed twice (on summit or high pass days: at peak and camp) and three days when no testing was carried out. The first four measurements were taken in Edinburgh, UK, before the start of the expedition; these were not included in the main analyses.

Procedure

At the start of the expedition, all participants remained in the base camp in Manali (1,992m) for two days to acclimatise. They then travelled by bus to the trail head at 3,271m, whence the trek commenced. Subjects were climbing on foot, over rough terrain (including snow and ice), carrying approximately 8-10 kg of personal equipment. Frequent rest stops were taken during climbing whenever required. The speed of ascent varied throughout the expedition, and on ascent days (days when there was a net gain in altitude relative to the previous day), between 70 and 1,036 m were climbed (mean = 543 m). Approximate duration of each day's climb was 6 hours, so an average speed of ascent across the ascent days could be estimated at 1.5m/min.

RTs, vital signs and Lake Louise AMS score were recorded for each subject throughout the expedition at a range of altitudes (1,992 [PaO_2 79% of the sea level] – 5,565 m [PaO_2 51%]), as illustrated in Figure 6.2. Data were recorded on four separate occasions in Edinburgh (76-86 m above sea level) in order to allow participants to familiarise themselves with the RT task and minimise practice effects.

Throughout the expedition phase, test sessions took place at camp after the completion of each day's climbing, following at least 2 hours' rest. Where possible, this took place in a mess tent. Test sessions took place at the same time each day (late afternoon) in order to minimise any diurnal variation. Test sessions at the three highest altitudes were exceptions to these conditions, being conducted away from camp on the mountain, after a minimum of only 5 min rest. On those days, measurements were also taken back at camp later that day, at the usual time. Body temperature was not measured at any of the peaks (measurement occasions 11, 15, or 22) and respiratory rate was not tested at the last peak, (Friendship Peak; measurement occasion 22).

Statistical analyses

The preliminary analyses were performed using SPSS 16.0 (SPSS Inc., Chicago, IL). For the purpose of simple comparisons, means and standard deviations for variables were calculated separately for base camp (averaged across the four measurements taken at 1,992 m) and high altitudes (average for the three peaks). As mentioned before, data from the familiarisation sessions were not included in the analyses. This was because the sea-level measurements were expected to differ systematically from the data gathered during expedition. Firstly, they were collected for different reasons (familiarisation with the test vs. collection of data for analysis); secondly, the sea-level measurement were only taken prior to the expedition (hence were exposed to less practice than data collected later on); thirdly, a large time interval separated the two measurement bouts.

Correlations between the variables of interest were tested using Pearson's correlation coefficient. Variables which correlated with RT mean and log-transformed RT ISD were selected for inclusion in the relevant subsequent respective models. Partial correlations between log transformed RT ISD and other variables controlling for RT mean were also performed.

Linear mixed modelling (PROC MIXED) was used to model the effects of altitude and other variables on RT mean and on RT ISD (Der & Everitt, 2009; Singer, 1998). The modelling was performed using SAS 9.2 software (SAS Institute

Inc., Cary, NC). A set of linear mixed models was fitted separately for RT mean and RT IIV. Bearing in mind the correlation between the two variables, the effects of predictors on RT IIV were modelled both with and without controlling for RT mean. The method used for estimation was maximum likelihood, to enable comparisons of nested models.

Results

Correlations between the variables

As expected, respiratory rate was positively correlated with altitude ($r = .538$, $p < .001$). At the base camp mean number of breaths per minute was 12.35 (SD = 1.86) which increased more than two-fold at the peaks (M = 28.00, SD = 10.42). This increase in breathing rate did not prevent SpO_2 from dropping, however. SpO_2 was highly negatively correlated with altitude ($r = -.719$, $p < .001$) and decreased from 96.20% (SD = 1.54) at the base camp, to 83.41% (SD=6.47) at the peaks. Pulse was also correlated with altitude ($r = .299$, $p < .001$), and the mean increased from the baseline level of 74.63 (SD = 14.47) to 94.59 (SD = 22.02) at the highest altitude. As expected, Lake Louise score increased with altitude from 0.33 (SD = 0.86) at the base camp to 1.69 (SD = 2.00) at the highest altitudes ($r = .219$, $p < .001$). Greater SD at the peaks reflects greater inter-individual variability in Lake Louise scores: whereas some individuals reported a number of symptoms, others remained virtually unaffected even at the highest altitudes.

RT mean and ISD correlated positively with altitude ($r = .193$, $p < .01$ and $r = .234$, $p < .001$, respectively). RT mean increased from 260.85 (SD = 34.62) to 288.76 ms (SD = 32.70) between the base camp altitude and the three peaks. RT ISD averaged across participants also increased from 39.78 (SD = 16.87) at the camp to 61.45 ms (SD = 33.98) at the peaks. As expected, there was also a positive relationship between the two RT measures ($r = .295$, $p < .001$). All correlations between the variables are given in Table 6.1.

Mean reaction time

Mean RT averaged across the participants with superimposed altitude at which the test took place are shown in the upper panel of Figure 6.3. There was an apparent decrease in group mean RT over the 21 measurement occasions. Mean RT for the whole group on the first day of the expedition (measurement occasion 5) was 286.80 ms (SD = 33.45) whereas, on the last day, it was considerably lower, at 240.40 ms (SD = 28.35). RT mean tended to decrease throughout the expedition; however, there are three clear local points at which the longest RTs appear, which exactly correspond to the three peaks climbed by the participants (measurement occasions 11, 15 and 22).

The time dimension

Despite the apparent variability in mean RT from occasion to occasion, an overall downward trend could be observed in individual RT change trajectories in all 10 participants. Examples of the trajectories across occasions observed in three participants, with superimposed regression lines, are given in Figure 6.4. The average response times shortened throughout the expedition.

The uniform decrease in RT mean with time across the participants suggests a continuing practice effect. To test the relationship between measurement occasion and RT formally, a simple linear mixed model with RT mean as the outcome and measurement occasion and its quadratic term as fixed effects (Model 1.1; see Table 6.2) was fitted. Adding random slopes in measurement occasion did not significantly improve model fit over that obtained with random intercepts; therefore, individuals' intercepts were allowed to vary, but separate slopes for each participant were not modelled. In other words, the model permitted differences in mean RT level between individuals, but assumed that change trajectories would be the same. In this simple model, measurement occasion showed a curvilinear relationship with mean RT. To facilitate interpretation, predicted RT values from Model 1.1 have been plotted against measurement occasions (see Figure 6.5). It appears that RTs became faster with practice (i.e. the number of times the test was taken) and that the changes in RT mean tended to be larger in the later testing sessions. In terms of random effects,

there was significant variability in individual intercepts ($p < .05$), indicating that participants differed reliably in their initial mean RT level.

Altitude and other variables

Variables which had significant correlations with mean RT (see Table 6.1) were included in the initial linear mixed model (Table 6.2) as fixed effects. These were: number of errors and its quadratic term, measurement occasion and its quadratic term, altitude with its quadratic and cubic terms, altitude gain, Lake Louise score, SpO_2, and pulse. In addition to the main effects, two interactions were added to the model: namely, measurement occasion by altitude, and measurement occasion by altitude gain. These interactions were of theoretical interest as they could indicate whether acclimatisation to the effect of altitude and altitude gain had occurred.

To account for the intra-individual correlation, I tested for random intercepts (individual differences in mean RT) and slopes in measurement occasion (differences between individuals in the rate of change of their RTs). Neither variance of the slopes nor covariance between the random intercepts and slopes were significantly different from zero, therefore only random intercepts were retained for the subsequent analyses.

Fixed effects which were not significant at $p < .05$ were removed from the model. This was done in steps with model being refitted after each non-significant variable had been removed. This procedure resulted in the Model 1.2 summarised in Table 6.2. Mean RT decreased by 1.50 ms per error made, and by 2.16 ms per measurement occasion, suggesting that participants who had faster RTs also made more mistakes and that RTs became faster with practice. Lake Louise score had a negative effect on RT, with mean RT increasing by 1.97 ms with each point increase in Lake Louise score; in other words, individuals with more severe physiological disturbances responded more slowly than those less affected. The relationship between altitude and RT was nonlinear (cubic), and its representation is illustrated in the upper panel of Figure 6.6. There was also significant variance in individual intercepts ($p < .05$), indicating that there were reliable individual differences in mean RT level between the participants.

Because there were no measurements taken between 1,992 m (base camp) and 3,000 m, the initial rise and fall of RTs below 4,000 m could be an artefact of the cubic function in Model 1.2. Therefore, the altitude-RT relationship was further investigated by fitting a two-part model (1.3), with a linear function of altitude below or equal to 4,000 m and a quadratic one above 4,000 m, together with covariates as before (i.e. number of errors, measurement occasion and Lake Louise score). In this model, altitude below 4,000 m was not a significant predictor of mean RT and was therefore removed form the model. The resulting Model 1.3 is summarised in Table 6.2. There were only minimal changes in the effects of number of errors, measurement occasion and Lake Louise score, relative to the model, with a cubic function of altitude (Model 1.2). The effect of altitude above 4,000 m was a quadratic function and the predicted values for RT at different altitudes based on this model are illustrated in the lower panel of Figure 6.6. Models 1.2 and 1.3 fit the data equally well, as fit statistics were very similar for both (-2 Log Likelihood: 1,691.9 and 1,961.5, respectively).

Reaction time variability

RT ISD averaged across the participants with superimposed altitude at which the test took place are shown in the lower panel of Figure 6.3. Mean RT ISD on the first day of the expedition was 45.50 ms (SD = 19.46), and on the last day was 37.90 ms (SD = 11.23). As can be seen from the plot, RT ISD does not appear to change with time beyond occasion-to-occasion variability. However, as was the case with mean RT, some local rises in IIV can be observed, especially coinciding with the first peak, intermediate at the second peak and somewhat less marked for the third one.

The time dimension

The downward trends observed in each individual's mean RT data is not readily noticeable in RT ISD. Here, only some individuals demonstrated a decrease across measurement occasions (and a modest one) whereas, for others, the intra-

individual variability in RT remained relatively stable or even increased (see Figure 6.7 for examples of individual trajectories from three participants).

The nature of the relationship between RT variability and measurement occasion was not clear from examination of either the individual or group trajectories. A simple time-dimension model (Model 2.1, see Table 6.3) was fitted to test whether the two variables were related. In this model, log-transformed RT ISD was the outcome and measurement occasion and its quadratic term were entered as fixed factors. I tested for random intercepts and slopes in measurement occasion. Adding a random effect of measurement occasion did not significantly improve the fit of the model; therefore, a model with random intercepts only was selected. Measurement occasion showed a curvilinear relationship with RT variability. To illustrate this relationship graphically, predicted values for RT ISDs (log-transformed) derived from Model 2.1 have been plotted against measurement occasions (see Figure 6.8). Initially, RT ISD increases and then decreases to a level somewhat lower than that at the beginning. In terms of random effects, there was only a trend towards variability in individual intercepts ($p = .072$), suggesting that participants did not differ in the level of RT variability.

Altitude and other variables

Variables significantly correlated with log-transformed RT ISD (see Table 6.1) were included in the linear mixed model as fixed effects. These were: altitude with its quadratic and cubic terms, atmospheric temperature, Lake Louise score, and SpO_2. Measurement occasion and number of errors were also included, together with their quadratic terms. Respiratory rate was not included in the model as it was not measured on one of the peaks. Measurement occasion by altitude interaction was also entered to test for potential acclimatisation effects. Because RT mean and RT variability are related, two models for log RT ISD were fitted. One contained the variables as described above (model 2.1), and one contained an additional fixed effect of mean RT (model 2.2).

As with mean RT, random effects of intercept (individual differences in the RT ISD) and slope on measurement occasion (differences between the individuals in

the rate of change in their RT ISD over time) were included. For both Models 2.2 and 2.3 there was no significant variance of random slopes or covariance between intercepts and slopes. Hence, only random intercepts were specified in the subsequent models.

The modelling was carried out in a backwards stepwise manner, as described before. The summary statistics for final versions of models 2.2 and 2.3 are given in Table 6.3. For ease of interpretation, the results have been back-transformed to the original values using an anti-logarithm and could then be expressed as a percent change in RT ISD per unit change in a predictor. When not controlling for the mean RT (model 2.1), RT ISD increased by approximately 4.2 % with a point increase in Lake Louise score, and decreased by 1.7 % per degree Celsius increase in temperature. In other words, more severe physiological disturbances were associated with more variable RTs, and higher temperature was associated with a decrease in the RT ISD. In this model, the effect of random intercepts, i.e. individual differences in the level of intra-individual variability, approached significance ($p = .068$).

When mean RT was included in the model (model 2.3), it was a significant predictor of RT IIV, with RT ISD increasing by approximately 1.2 % per millisecond increase in mean RT. Measurement occasion was also a significant predictor of intra-individual variability in RT, with RT ISD increasing by 2.7 % per occasion tested. In other words, when practice effects are controlled for, slower RTs were also more variable; with mean RT kept constant, IIV seemed to increase each time an individual was tested. There was also a significant variability in individual intercepts ($p < .05$); that is, there were reliable differences between the participants in their level of RT IIV.

Discussion

The main findings of the present study were as follows. Altitude had little effect on RT mean below 4,000 m, whereas a marked increase in mean RT was observed above that threshold. Altitude was not a significant predictor of RT ISD. However, when not controlling for RT mean, two parameters closely related to

altitude, namely temperature and Lake Louise score, were significant predictors of RT variability. RT ISD increased when more physiological disturbance (as assessed by Lake Louise score) was present and decreased with a rise in temperature. When the effects of RT mean were controlled for, the altitude related predictors were no longer significant. Instead, RT intra-individual variability became greater with slower mean RTs and RTs became more variable with each measurement occasion. The results of the present study shall be discussed in more detail under three sections: correlations between the variables, reaction time mean, and reaction time variability.

Correlations between the variables

This study found that altitude was related to vital signs and to physiological disturbance, as assessed by the Lake Louise score. Both RT mean and ISD were weakly but significantly positively correlated with altitude and with each other. Mean RT was positively correlated with measurement occasion, altitude, altitude gain, Lake Louise score and pulse. Number of errors and SpO_2 were negatively related to RT mean. Bivariate correlations showed that RT intra-individual variability was greater when altitude, respiratory rate and Lake Louise scores were higher, whereas temperature and SpO_2 were negatively related to RT ISD. Because, as mentioned above, RT mean and RT intra-individual variability were correlated, it was expected that a proportion of variance between them and their common correlates is shared. Therefore, partial correlations between log transformed RT ISD and other variables were investigated, when controlling for mean RT. When the effects of mean RT were partialled out, all correlations except for one were slightly attenuated, with two reduced to below the conventional significance level (SpO_2 and Lake Louise score). This finding suggests that SpO_2 and Lake Louise score share little unique variance with RT ISD beyond that explained by mean RT. Altitude and respiratory rate remained significantly positively related to RT IIV, and temperature had a negative correlation. One correlation, between RT variability and number of errors, which was not significant when not controlling for mean RT, became larger and conventionally statistically significant.

Reaction time mean

There were clearly-visible practice effects in mean RTs of all participants, as RTs showed a continuous decrease throughout the expedition. The shape of the relationship between RT mean and measurement occasion was contrary to the expectations. There was no evidence of plateau in the more advanced stages of the expedition, which would have suggested that the effect of practice diminished with time. To the contrary, the first part of the curve is relatively flat and the second part is noticeably steeper, suggesting that practice effects were relatively small at the initial measurements and increased with the number of tests taken. This finding opposes both the pre-study expectations and the evidence from previous literature, which found practice effects to diminish (Abraini et al., 1998; Denison et al., 1966) or disappear with time (Parkin, Kerr, & Hindmarch, 1997; Rabbitt & Goward, 1994). One possible explanation of the finding is that shape of the curve representing the relationship between RT and measurement occasion is influenced by the effects of altitude, not modelled in the time dimension model (Model 1.1). If RT is slowed by altitude, and it is accepted, for simplicity of argument, that the first part of the expedition involved mainly ascent and the second part mainly descent, then the following explanation might be offered: any practice effects occurring in the first part of the expedition would be masked or offset by an increase in RT associated with the ascent to a higher altitude. In the second part of the expedition, on the other hand, the descent leads to shortening of RTs adding to the practice effects, thus creating an impression of an accelerated decrease in mean RTs.

Another explanation of this finding might be in terms of the fixed stimulus sequence. Although the fact that the same sequence is used in each test was not made known to the participants, it is possible that it was detected, either implicitly or explicitly, and led to a continued learning, and consequently, faster RTs. It would be, therefore, beneficial to investigate practice effect in RT across a number of occasions with both set and random or quasi-random sequences to establish whether implicit learning is taking place when the sequences remain the same. Another possibility is that the finding is sample-specific; the sample used in this study comprised

undergraduate students, who could differ from the general population in task and strategy learning.

In the full model, altitude exhibited an effect on mean RT after controlling for measurement occasion, number of errors and Lake Louise score. From investigating the plot of predicted values of RT mean (Figure 6.6), it appears that there was little change in RT at the altitudes below 4,000 m, but above that threshold there was a steep increase. Further investigation revealed that there was no effect of altitude at or below 4,000 m, while a quadratic relationship was apparent above 4,000 m. This threshold is lower than that suggested by McFarland (1937b). However, the estimate in the present study has been obtained using more sophisticated modelling methods, taking into account various altitudes rather than relying on pairwise comparisons of performance at certain altitudes with that at sea level. As such, it allows detection of a more precise threshold altitude.

As expected, RT was positively related to number of errors, which is a well-established phenomenon in easy discrimination tasks requiring fast responses (Luce, 1986). Lake Louise score was also related to mean RT, indicating that physiological disturbances that are not accounted for by altitude affect response times. Interestingly, after accounting for variance brought about by altitude, physiological disturbances and the number of errors, measurement occasion was linearly related to mean RT, implying that there were continuous practice effects.

Reaction time intra-individual variability

From the visual inspection of the RT variability there were no apparent practice effects. However, a simple time dimension model fitted showed that measurement occasion was a significant predictor of RT ISD. More precisely, a quadratic relationship was revealed with the curve rising first and then falling to the level slightly below that at the start of the expedition. It is difficult to imagine that the curve suggested by the time dimension model reflects genuine shape of practice effects in RT ISD. One possibility is that, as suggested before with regards to the mean RT model, measurement occasion may be tapping into altitude effects which are not modelled in the time dimension model. It is possible that the rising arm of the

curve reflects rising altitude, which leads to an increase in RT ISD and, similarly, the falling arm represents the effect of the descent to lower altitudes, which is expected to be associated with less variable RTs.

Lake Louise score and atmospheric temperature were significant predictors of RT ISD, with no effect of altitude. As expected, higher Lake Louise scores were associated with greater intra-individual variability in RT. The effect of temperature was negative, with less variability apparent at higher temperatures. The lack of association between altitude and RT intra-individual variability was unexpected. Although Bonnardel and Liberson (1933) found no reliable changes in RT intra-individual variability at high altitude, as mentioned previously, their conclusions were drawn from only two participants and so they cannot be accepted without due reservations. It is quite possible that their null finding is due to low power. Somewhat more convincing findings were provided by McFarland (1937a, b), and are not corroborated by our results. The gradual ascent to a high altitude in our study makes it similar to the expedition described by McFarland (1937b), who noted an increase in RT ISD at altitudes above 5,340 m. The lower threshold observed in McFarland (1937a) was most likely attributable to an exaggerated effect of altitude with a rapid ascent, and so not expected to be replicated in our study. Given that only 1 out of 21 measurement occasions took place at an altitude above the 5,340 threshold, lack of association between RT ISD and altitude in our data is not entirely surprising. Future expeditions setting out to test the effects of altitude on RT intra-individual variability should attempt to increase the number of measurements taken above the suggested threshold to enable better investigation of the effects of the highest altitudes.

The mechanism linking atmospheric temperature with RT ISD is not clear. It may be that lower temperature leads to cooling of extremities, which in turn impairs manual dexterity. If that were the case, however, it is surprising that the effect of atmospheric temperature did not affect mean RT.

Because altitude and temperature are related, it was expected that a portion of variance in RT ISD would be accounted for by both these variables. However, since altitude did not remain in the model whereas temperature did, it may be that altitude does not contribute unique variance to intra-individual variability in RT above that of

temperature. Needless to say, the existence of a link between temperature and intra-individual variability in RT needs to be investigated further. If, however, one accepts that the effects of altitude on RT ISD are mediated by temperature and Lake Louise score, then the fact that both these effects are removed when RT mean is included in the model suggests that the changes in RT variability related to altitude are secondary to changes in RT mean.

Given the above, it appears that altitude-related effects influence intra-individual variability in RT via mean RT. As such, it is in line with the reasoning of some (e.g. Shammi et al., 1998), who argue that RT ISD is secondary to mean RT. The correlation between RT mean and ISD could then be explained in terms of variability increasing as a function of mean. Others, however, suggest that it is the changes in RT ISD that drive the mean change. For example, Rabbitt, Osman et al. (2001) proposed that an increase in RT mean may be due to an increase in the number of very slow responses (i.e. increased RT ISD). It is not possible to imply causality from a simple correlation; however, using more sophisticated statistical methods can help shed light on the issue. For example, in the present study, RT mean was controlled to obtain independent estimates. The need for controlling for the RT mean when considering effects on RT IIV has been acknowledged by many researchers concerned with RT (e.g. Hultsch et al. 2002; Der & Deary, 2006). Without it, one cannot be certain that the observed effects are not brought about by RT mean. In the present study, after excluding the effects of RT mean on RT ISD, there remained no significant effect of altitude on RT ISD. Therefore, it appears that in this case, changes in RT mean are primary to those in RT variability; that is, altitude and altitude-related variables leads to slower RTs, which in turn, give rise to greater variability in responding.

Strengths and limitations

The main strength of the present study is that the measurements were taken frequently throughout the expedition, allowing for trajectories of the RT mean and IIV changes to be investigated. Moreover, the present study used a sophisticated method to model the relationships of RT mean and variability with altitude and other

variables. The statistical methods used in this study were not limited to pairwise comparisons between two RT measures, which was a common practice in the existing literature. Instead, the changes in RT and RT ISD at all altitudes were modelled, providing a more reliable way of investigating the effects of altitude on RT.

Finally, this study has implemented the practice of controlling for effects of mean RT when investigating RT ISD for the first time in this area of research. This procedure has not been used by any of the previous studies that investigated RT IIV at high altitude. Therefore, the present study has added to the existing knowledge about the effect of altitude on RT variability by demonstrating that the altitude-relevant effects on RT ISD are secondary to those on mean RT.

The limitations of the study are several. Firstly, the use of a set sequence of stimuli could have given rise to implicit learning and resulted in steady practice effects which did not plateau off. Secondly, only one peak was above the threshold altitude suggested by McFarland (1937b) which might have led to the null finding regarding the effects of altitude on RT ISD. Thirdly, respiratory rate was not measured at the peaks. This limitation has prevented the investigation of whether increased respiration alleviated the adverse effects of altitude on RT mean and ISD. Fourthly, the conditions of testing at the three peaks were different from those taken in camps. Testing at camps followed a rest period of at least 2 hours and the measurements were taken in a tent; at peaks, the rest period was considerably shorter (approximately 5 minutes) and the tests were carried out in the open air. This discrepancy in conditions could be one of the reasons why prolonged RTs were observed at the very highest altitudes. However, this issue is a more serious problem in studies that compare measures taken at high altitudes with baseline measures. In those circumstances, the less favourable conditions at the peaks could bias the RTs making them slower or more variable than they would be otherwise and, hence, artificially inflate the difference from the baseline measures. This study is less susceptible to the same bias, since measurements were made at various altitudes throughout the expedition and incorporated in an overall statistical model. Finally, the sample size used in this study was small; however, there has been no larger study to date to investigate the effect of altitude on RT IIV.

Future research

Following the limitations outlined above, it would be advisable to attempt to replicate the findings of the present study using a variable stimulus sequence and on a larger sample. Moreover, future expeditions should attempt to increase the number of peaks above 5,340 m to provide more data from altitudes above the suggested threshold for RT ISD and to measure respiratory rate at the peaks. Another useful addition to the design would be a control group performing the tests at the sea-level in parallel with the expedition group, which would give an idea of the expected amount of learning of the task not affected by altitude. Finally, future studies could broaden the existing research by investigating the relationship between altitude, RT measures and other factors such changes in mood states or physical fatigue.

Conclusion

This study has demonstrated that altitude affects RT mean at high altitudes, with a quadratic increase in RT observable above 4,000m. RT intra-individual variability appears to be affected by atmospheric temperature and physiological disturbance more than altitude per se. However, the altitude-related effects on RT ISD seem to be mediated by the RT mean, and not independently related to intra-individual variability. Continuous practice effects found in RT mean warrant further investigation.

Table 6.1. Correlations between the variables.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Mean RT	1	.295***	-.388***	-.488***	.193**	.372***	-0.107	0.114	.359***	-.298***	.277***	0.13
2. Log RT ISD (controlling for mean RT)		1	.021 (.154*)	-.073 (-.085)	.234*** (.189**)	.131 (-.023)	-.236*** (-.215**)	.018 (-.016)	.208**- .115	-.137* -.053)	.105 (-.026)	.188** (.158*)
3. RT errors			1	.330***	.059	-.126	-.192**	-.140	-.117	.041	-.214**	-.035
4. Measurement occasion				1	.032	-.543***	-.289***	-0.105	-.379***	.184**	-.114	-.007
5. Altitude					1	.125	-.516***	-.239**	.219**	-.719***	.299***	.538***
6. Altitude gain						1	.02	.017	.225**	-.091	.122	.096
7. Atmospheric Temperature							1	.247***	-.135	.292***	-.219**	-.556***
8. Body Temperature								1	.02	.042	.039	-.14
9. Lake Louise Score									1	-.167*	.229***	.261***
10. SpO_2										1	-.320***	-.342***
11. Pulse											1	.322***
12. Respiratory Rate												1

Note. RT=reaction time, RT ISD=reaction time intra-individual standard deviation, SpO_2=peripheral blood oxygen saturation, *N*s range from 180 to 209.
* $p<.05$, ** $p<.01$, ***$p<.001$

Table 6.2. Summary statistics of fixed effects in models of mean reaction time

Model	Parameter	Estimate	*SE*	*p*
Model 1.1	Intercept	283.35	9.68	<.001
	Measurement occasion	0.98	1.01	.333
	Measurement occasion^2	-0.12	0.03	<.001
Model 1.2	Intercept	175.67	42.96	.003
	Number of errors	-1.50	0.56	.009
	Measurement occasion	-2.16	0.17	<.001
	Lake Louise score	1.97	0.64	.003
	Altitude (per 1,000m)	121.21	38.90	.002
	Altitude^2	-38.27	10.99	<.001
	Altitude^3	3.90	0.98	<.001
Model 1.3	Intercept	568.19	83.79	<.001
	Number of errors	-1.48	0.56	.009
	Measurement occasion	-2.17	0.17	<.001
	Lake Louise score	1.96	0.64	.003
	Altitude > 4,000m (per 1,000m)	-134.26	36.62	<.001
	Altitude^2 > 4,000m	16.64	3.98	<.001

Table 6.3. Summary statistics of fixed effects in models of reaction time intra-individual standard deviation

Model	Parameter	Estimate [a]	*SE*	*p*
Model 2.1	Intercept	1.49000 (30.903)	0.068	<.001
	Measurement occasion	0.02500 (1.059)	0.0097	.011
	Measurement occasion^2	-0.00090 (0.998)	0.00032	.005
Model 2.2	Intercept	1.74406 (55.470)	0.04436	<.001
	Lake Louise score	0.01801 (1.042)	0.00669	.007
	Temperature	-0.00754 (0.983)	0.00223	<.001
Model 2.3	Intercept	0.03764 (1.091)	0.16575	.825
	Mean RT	0.00530 (1.012)	0.00052	<.001
	Measurement occasion	0.01142 (1.027)	0.00192	<.001

[a] Values in parentheses represent estimates back-transformed using anti-logarithm

Figure 6.1. Portable reaction time testing device used in the present study.

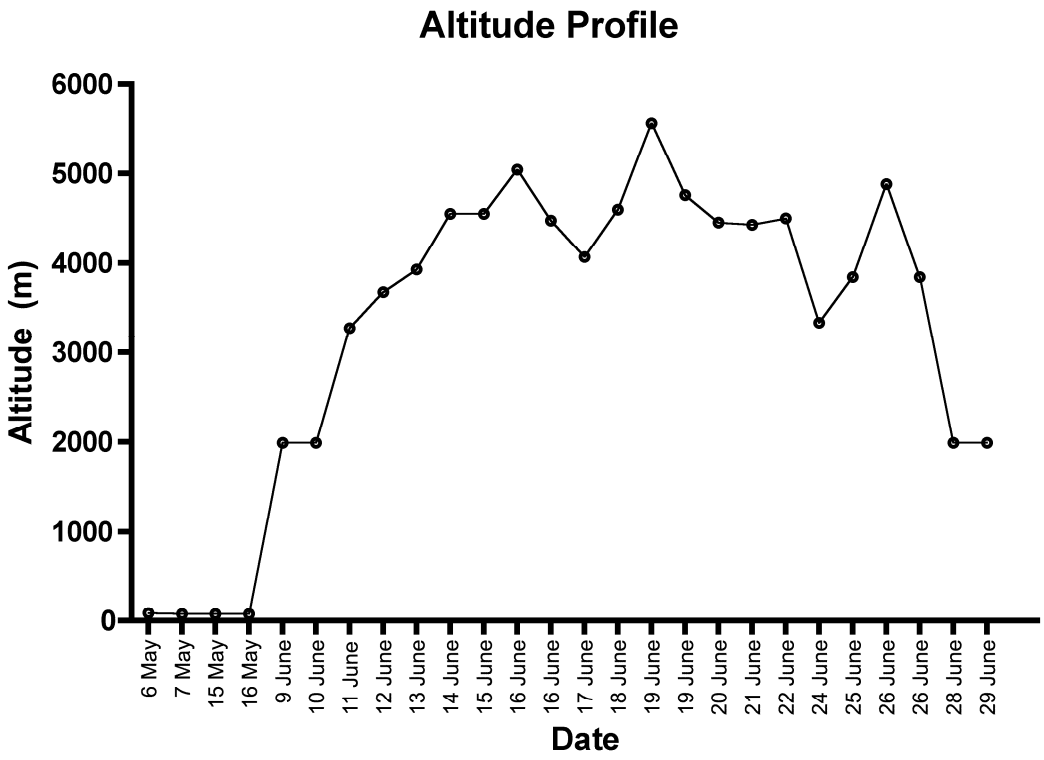

Figure 6.2. Altitude profile of the measurement occasions. Dates relate to when tests were performed. The first four occasions of measurement were familiarisation sessions performed in Edinburgh before the start of the expedition; data collected on those days were not included in the analyses.

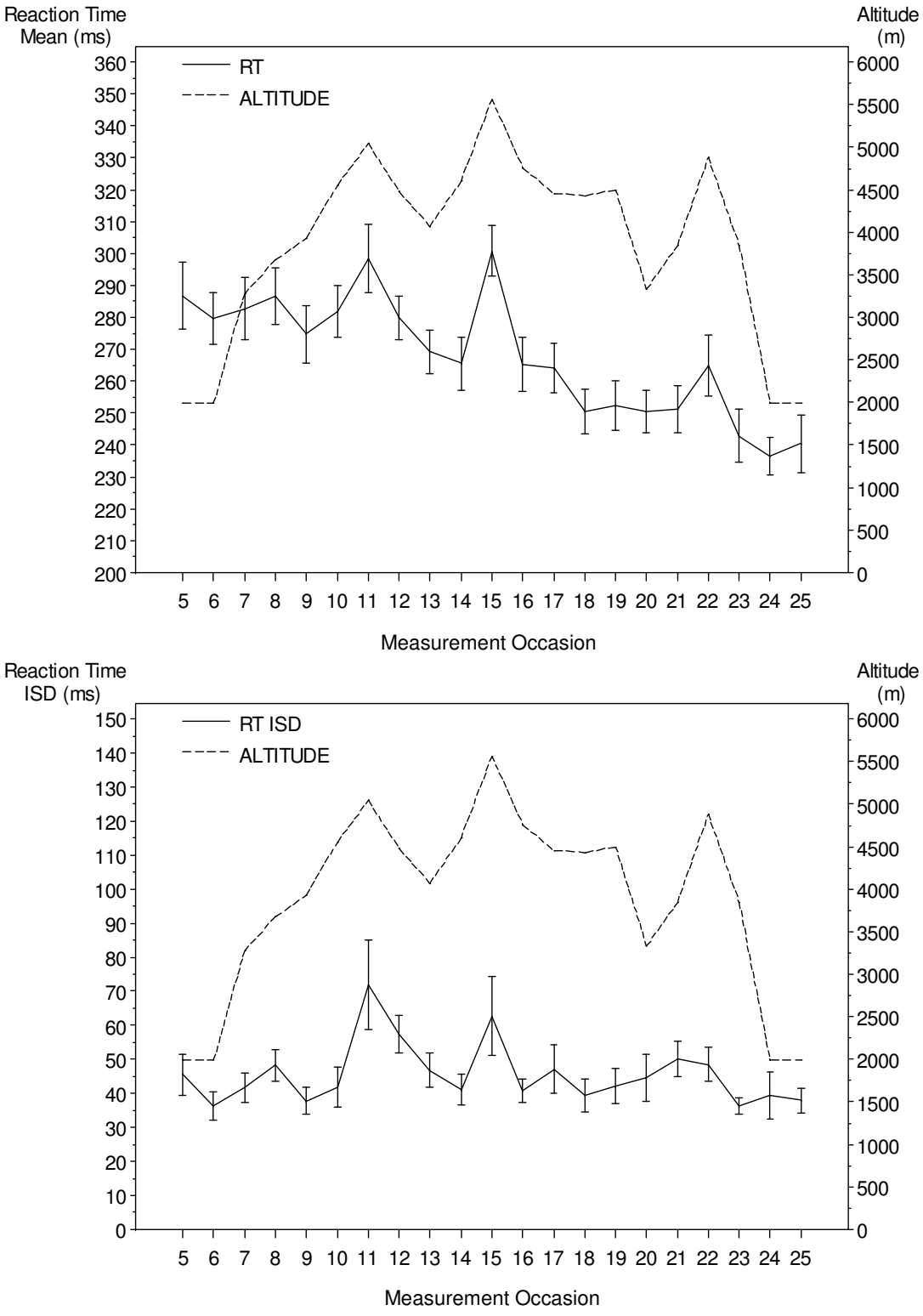

Figure 6.3. Group average mean reaction time (RT; upper panel) and intra-individual standard deviation (RT ISD; lower panel) across the measurement occasions with superimposed altitude at which the tests were taken. Bars represent standard error of the mean.

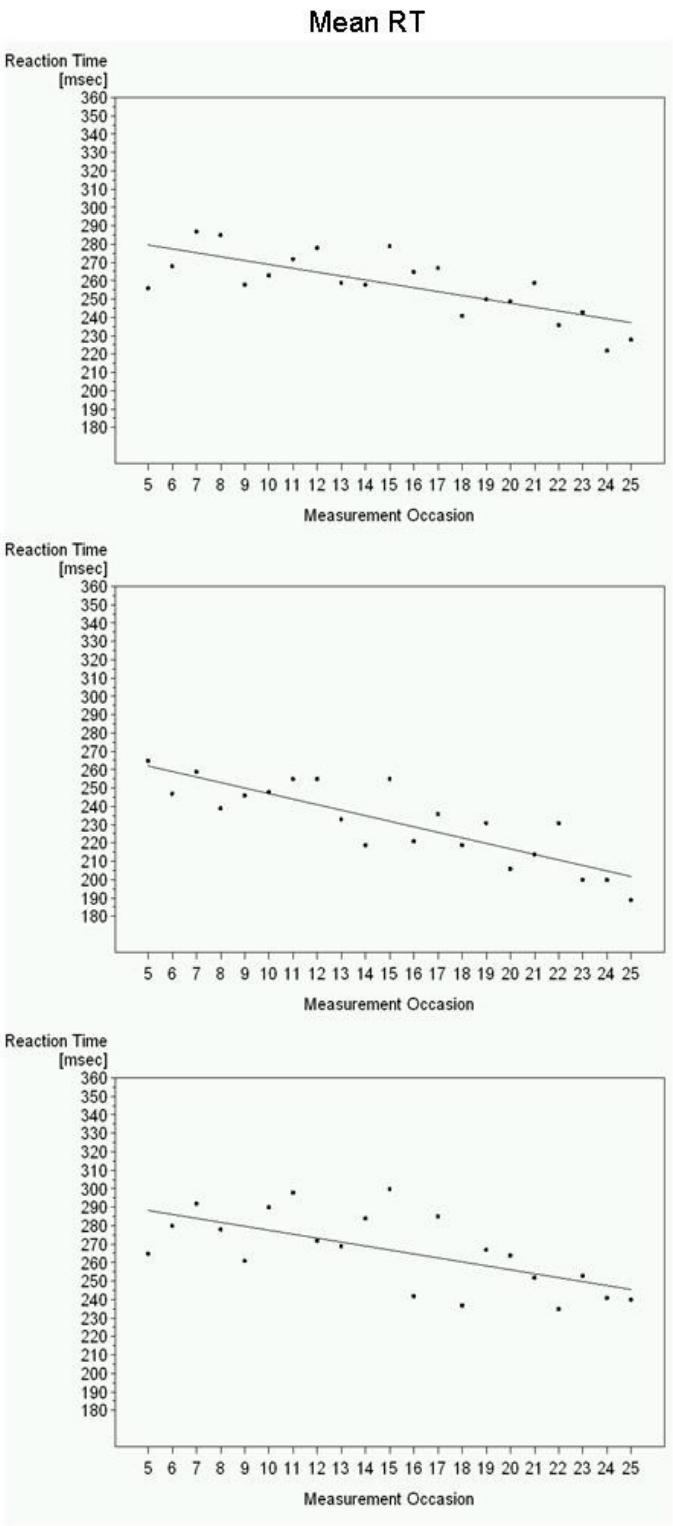

Figure 6.4. Individual trajectories of mean RT change across measurement occasions with fitted regression lines. Practice measurements 1 to 4 have been removed.

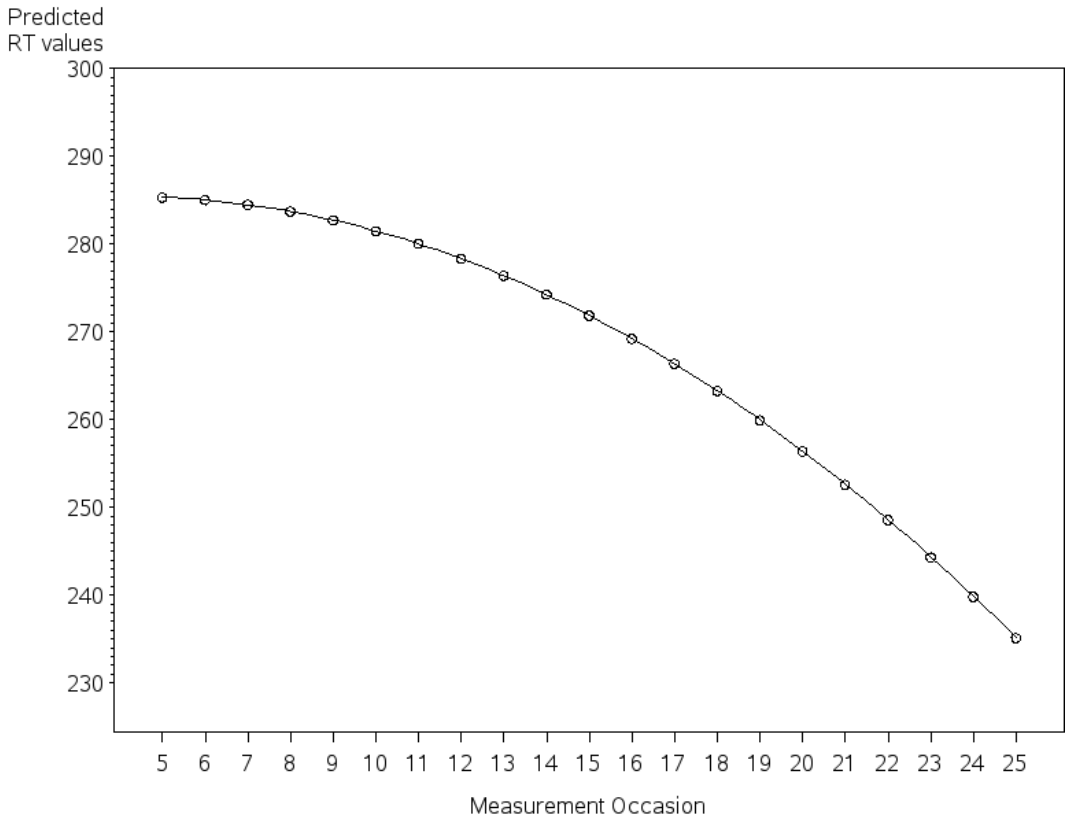

Figure 6.5. Predicted mean values of mean reaction time (RT) based on Model 1.1.

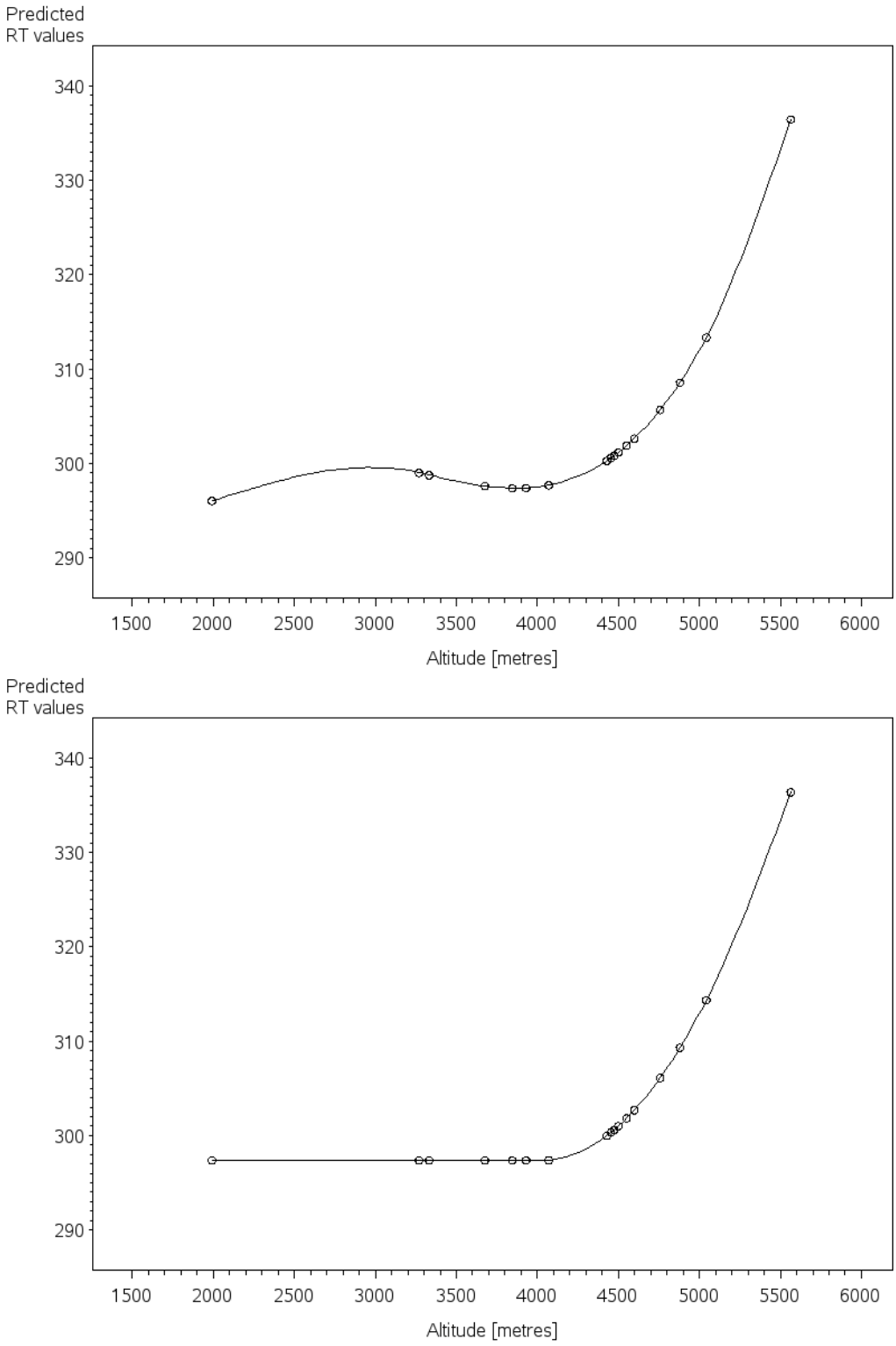

Figure 6.6. Predicted mean reaction time (RT) at different altitudes. The upper panel represents values based on Model 1.2 (a cubic function of altitude). The lower panel depicts the predicted values based on Model 1.3 (no effect of altitude below 4,000m and a quadratic relationship between altitude and RT above 4,000m). Points represent the values predicted for the altitudes at which measurements were taken. The solid line is the curve best fitting the data.

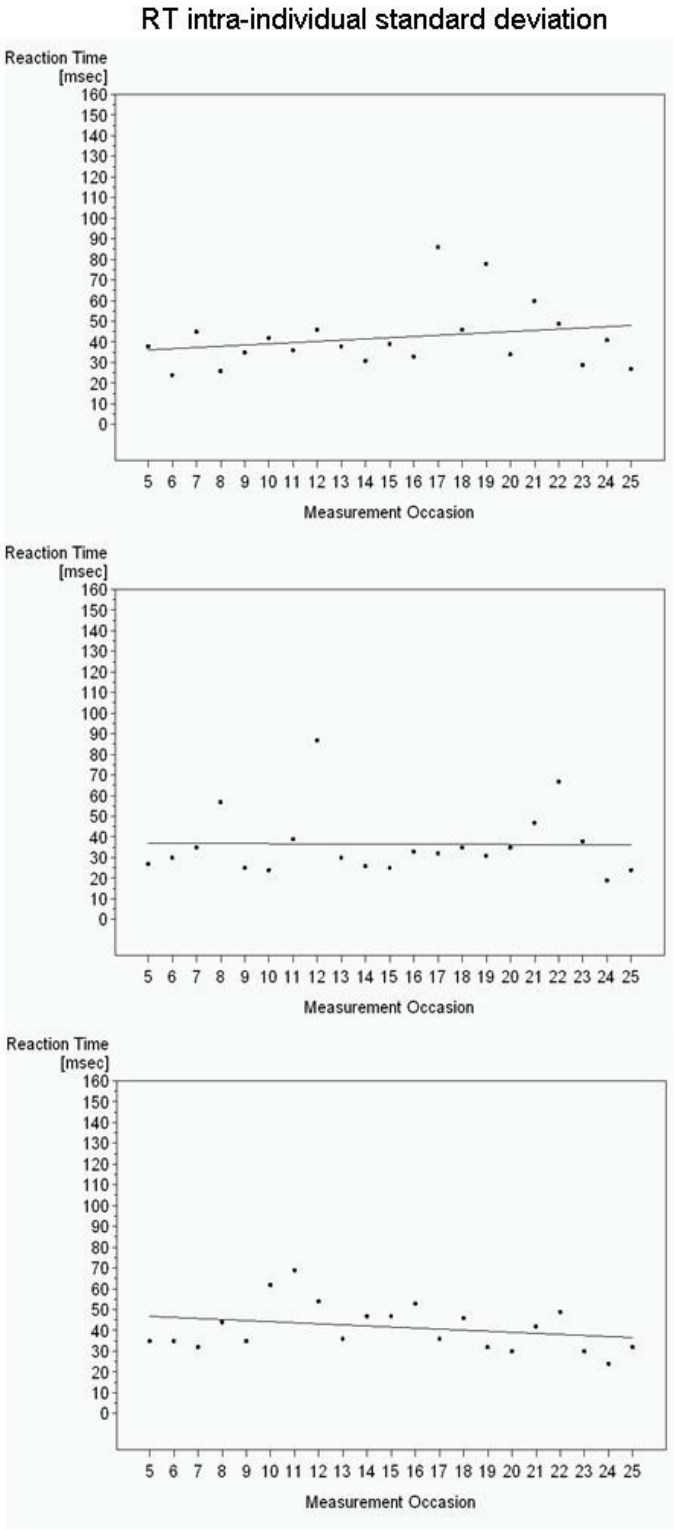

Figure 6.7. Individual trajectories of RT ISD change across measurement occasions with fitted regression lines. Practice measurements 1 to 4 have been removed.

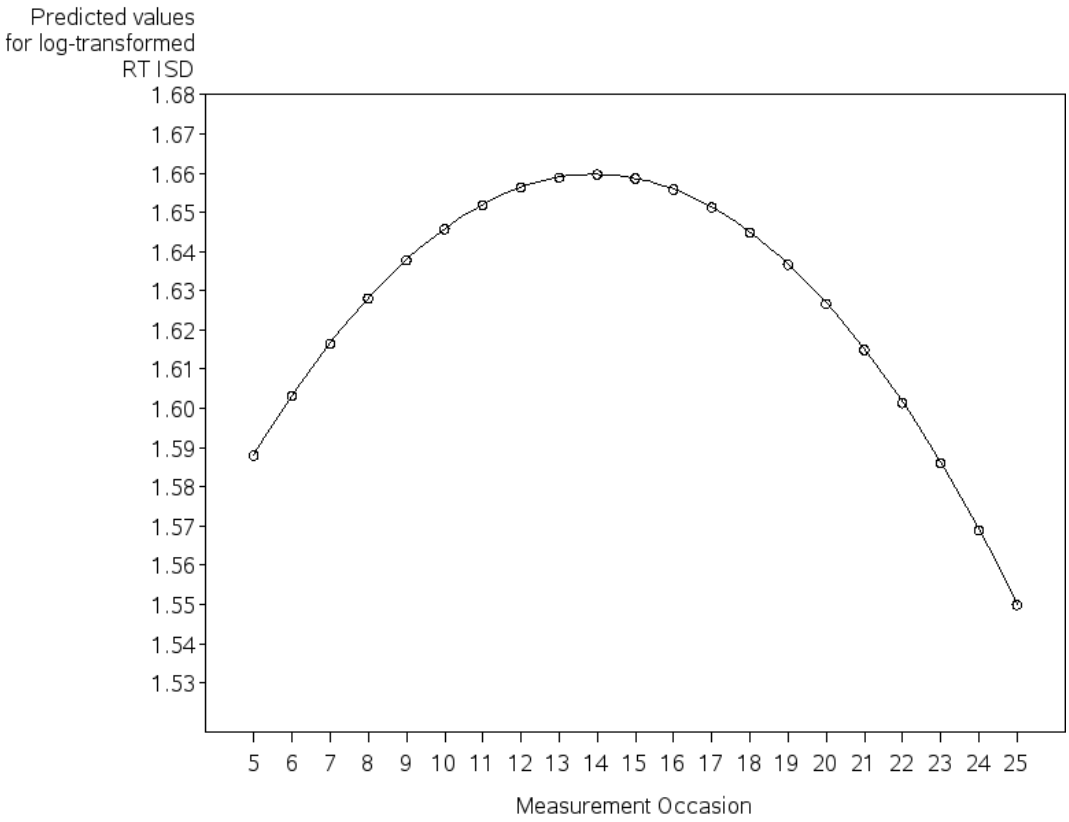

Figure 6.8. Predicted mean values for log-transformed reaction time intra-individual standard deviation (RT ISD) based on Model 2.1.

Chapter 7 -
General Discussion

This thesis has investigated determinants and correlates of intra-individual variability (IIV) in simple and choice reaction time (SRT, CRT). The two effects of primary interest were age and sex. In addition to these, various experimental task-characteristic variables have been considered, such as stimulus sequence, or preparatory interval (PI) duration. Moreover, this thesis examined the effects of a non-participant, non-task, extrinsic experimental condition, high altitude, on CRT mean and variability. In all these investigations, at least two different measures of variability were used and I compared their results.

Summary of findings

The first comprehensive systematic review with meta-analysis in this area was completed (Chapter 2) and it addressed five main issues: (1) whether there were age group differences in SRT IIV and CRT IIV; (2) whether the magnitude of these differed depending on the age gap between the groups compared; (3) whether task complexity and difficulty had an effect on the magnitude of age differences in RT IIV; (4) whether age differences were found with mean-adjusted RT IIV measures in addition to any found with raw measures; and (5) whether task and procedural factors could explain some of the between-study heterogeneity in the effects found. The results confirmed that, in cross-sectional data, RT IIV is larger in older than in younger individuals. Findings also demonstrated that the magnitude of the difference between old and young adults was greater than the magnitude of the difference between old and middle-aged adults. This finding was taken to mean that age-related decline in performance (i.e., the increase in RT variability) is not limited to old age

but, to a degree, occurs already in mid-adulthood. In terms of the importance of task complexity and difficulty, the findings were less clear: the differences between older and younger adults were greater in CRT than in SRT, but this was not observed for different numbers of choices in CRT tasks. Age differences in SRT IIV and CRT IIV were observed with both raw and mean-adjusted measures of variability; however, the effects were larger for raw measures. This finding suggests that some (but not all) of the increase in IIV with age shares common variance with increases in mean RT. Procedural factors and task characteristics did not explain much of the heterogeneity between studies, possibly because the number of studies in different subgroups was generally modest. However, effect sizes from across the vast majority of studies—regardless of the differences between them—suggested that RT IIV is greater among older than younger or middle-aged adults.

The cross-sectional study of age and sex effects across the lifespan, conducted on the Science Festivals sample (Chapter 3) confirmed the usually-found pattern of increasing speed and decreasing variability in RT across childhood and adolescence, followed by a decrease in speed and increase in variability through adulthood and into old age. The performance on the more complex CRT task showed greater deterioration in the older ages than did SRT task. The key finding of this study was in relation to sex differences in RT IIV. Adult females had greater RT variability than adult males, but this difference was not found in children. This pattern, found for the first time in a single sample, suggests that there might be an age threshold at which the sexes diverge in their RT variability. The study also demonstrated that two different methods of controlling for mean RT when investigating RT IIV (using the coefficient of variation, CV, versus including mean RT as a covariate in models of RT IIV) can lead to different conclusions in relation to age effects. On the other hand, the effects of sex seem rather robust to the two methods of controlling for mean RT.

The four waves of RT data from the three narrow age cohorts of the Twenty-07 study allowed an investigation of both the cross-sectional and the longitudinal effects of age (Chapter 4). The usual cross-sectional age differences in both SRT and CRT were confirmed in this study: youngest participants were least variable, the oldest participants were most variable, and the middle-aged participants had

variability levels between the other two groups. This investigation makes a valuable contribution to the existing body of longitudinal evidence, which is largely based on small and/or unrepresentative samples and follow-up periods which may not be sufficiently long to identify actual age trends. The key finding of this study, on a large and population-representative sample, was the longitudinal increase in CRT IIV between the ages of 55 and 75. Longitudinal SRT data were compromised due to across-wave RT apparatus-programming changes, so no conclusions could be drawn about the changes in SRT IIV over time. The main limitation of the study was the inability to separate practice from age effects; therefore the magnitude of the age effect implied from the longitudinal data may be an underestimate.

The trial-by-trial analysis of RT performance (Chapter 5) did not corroborate Reimers and Maylor's (2006) finding of slower female responses on initial trials of a RT task, followed by slightly faster responses on subsequent trials. Females had greater CRT ISD than males, but this variability appeared to persist across the whole block of 40 trials. Therefore, differential performance across trials did not explain greater female variability found in CRT in the Science Festivals sample. Moreover, although not designed to test them specifically, this study demonstrated the importance of task variables such as stimulus sequence, and preparatory interval (PI) length. A clear pattern of systematic trial-to-trial variations in RT could be observed across sexes and ages. These systematic variations were most likely to be due to the identical sequences of stimuli and PIs presented to each individual.

The final empirical study (Chapter 6) contributed to two lines of evidence: about the effects of altitude-related hypoxia on RT variables, and about whether the effects on mean and variability in RT are independent. First, it showed that mean RT increased with altitude, with a quadratic increase in RT observed above about 4000 m. RT IIV was also greater at higher altitude, but the main contributors to the increase appeared to be atmospheric temperature and the severity of mountain sickness rather than altitude per se. Second, this study showed that altitude-related effects on RT IIV were mediated by the increase in mean RT, thus suggesting that altitude does not affect RT variability directly. An unexpected finding from this study was related to practice effects: RTs continued to improve over the 25 testing sessions.

Contributions to the existing knowledge

Age and RT IIV

The general finding of this thesis regarding age and RT variability is that older individuals are more variable in their RT. This finding supports a large body of available research (Hultsch, MacDonald, & Dixon, 2002; Dixon, et al. 2007, Der & Deary, 2006). Evidence from three independent investigations suggested that age differences are not only observed between the young and the old, but that changes in RT variability occur across the lifespan. The meta-analytic review (Chapter 2) demonstrated that differences between groups of older and younger adults, in both SRT IIV and CRT IIV, were greater than between older and middle-aged adults. The finding that the groups closer together in age are more similar, suggested that some decline has already occurred in the middle-aged group. Cross-sectional comparisons between the three cohorts in the Twenty-07 study (Chapter 4) showed that the youngest cohort demonstrated less SRT and CRT variability than the middle cohort which, in turn, had lower variability than the oldest cohort. The most compelling evidence for the suggestion that RT variability increases in mid-adulthood comes from the Science Festival samples (Chapter 3), in which the changes in RT IIV could be modelled across the continuous age range. In that dataset, CRT variability increased steadily throughout the adult age range, although this effect was not found in SRT variability. A similar pattern of variability on a 2-choice RT tasks has been suggested by Williams and his colleagues (Williams, Hultsch, Strauss, Hunter, & Tannock, 2005; Williams, Strauss, Hultsch, & Hunter, 2007). The novel finding from the Science Festival sample concerns sex differences, which will be discussed below.

Another issue investigated in this thesis was whether the age effects observed in cross-sectional studies are also apparent longitudinally. This was tested in the Twenty-07 sample (Chapter 4). Despite the problem of apparatus change across waves of testing, which seemed to have severely affected the SRT data, CRT data were less affected and allowed inferences to be drawn regarding longitudinal age effects on IIV. CRT variability (intra-individual standard deviation, ISD) decreased

over 20 years in the youngest cohort (i.e., from 15 to 35 years), remained relatively stable in the middle cohort (from 35 to 55 years), and increased markedly in the oldest cohort (i.e., from 55 to 75 years). Therefore, the conclusion from this is that, whereas a longitudinal effect of age on CRT variability can be observed in the older adults, it is not clear whether age-related increases occur in mid-life. The Twenty-07 study provided valuable and rare longitudinal data from a large and representative sample. The long, 20-year follow-up period of three narrow age cohorts provided a 60-year cross-cohort span. The investigation of longitudinal effects of age in this sample is a valuable addition to the conflicting existing empirical evidence (Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994; MacDonald, Hultsch & Dixon, 2003; Maddox & Douglass, 1974), which has been largely based on studies hampered by small and unrepresentative samples and inadequate follow-up periods.

Sex and RT IIV

Two of the studies in this thesis investigated sex differences in SRT and CRT variability. Earlier findings of greater female variability in RT, already reported for two waves of the Twenty-07 study (e.g., Der & Deary, 2006) were replicated in four waves (Chapter 4). The key study to test sex effects was the Science Festival study (Chapter 3), which was the first documented study of sex differences in RT variability across the lifespan, on a single sample comprising children, adolescents, and adults up to the age of 75. The novel finding obtained from this study was that adult females were found to have greater SRT ISD and CRT ISD than adult males, but that no such difference was observed in children. This finding confirmed the suggestion of Deary and Der (2005a) that sex differences in variability in RT might occur only after puberty. Now that the pattern has been established, future studies might begin to look into potential explanations for this effect. Deary and Der (2005a) already proposed one possible explanation—the effect of sex hormones. Testing this hypothesis directly would be a clear starting point for future investigations.

Although not accounting for the discrepancy between children and adults, Reimers and Maylor (2006) suggested another explanation of sex differences in RT variability. They proposed that females improve more than males across the trials of

an RT test. This hypothesis was tested in the Science Festival sample (reported in Chapter 5). There was no evidence of differential trial-to-trial performance of males and females at the initial trials of the CRT test. Therefore, Reimers and Maylor's hypothesis was not supported in the Science Festival sample.

Controlling for mean RT

I attempted to obtain both raw and mean-adjusted measures from all studies included in the systematic review. Subsequent meta-analyses revealed that age differences in RT IIV obtained with raw measures were larger than when mean-adjusted measures were used. This result was consistently found in both RT tasks considered (SRT and CRT), as well as for the two age comparisons made (old versus young and old versus middle-aged). A similar result was obtained from the Twenty-07 study sample (Chapter 4), where cross-sectional differences in SRT and CRT variability between cohorts were found to be smaller when obtained with CV rather than ISD. It can be concluded that RT variability is greater at older ages and that this difference is only partly due to age differences in average speed. This finding was in opposition to reports that age differences in RT IIV were removed by controlling for mean RT (e.g. Shammi, Bosman, & Stuss, 1988).

However, in Chapter 6, I demonstrated that some effects (specifically that of altitude) on RT IIV are removed by controlling for mean RT. Therefore, whereas age effects on RT IIV seem to remain significant (though attenuated) after controlling for RT mean, other effects are removed altogether. From the results obtained in this thesis, it appears that age contributes independent variance to RT IIV, whereas altitude effects do not.

In Chapter 3, alongside RT ISD, I used two methods of controlling for RT mean: I used CV and included mean RT as a covariate in models of RT ISD. The age effect (increasing variability with age in adulthood) remained significant when CV was used as a measure of variability, but was removed when mean CRT was included in the model as a covariate. When the whole sample was considered, an interesting pattern appeared in which polynomial effects of age were reduced to linear when mean RT was included as a covariate. The effects of sex was not

affected by the two different approaches to the same degree. In fact, they were much more robust to the different ways of adjusting for mean RT. A conclusion from this is that the choice of method of controlling for mean RT is much more important for the study of age effects than for sex effects. It is crucial that more research resources are directed towards understanding the reasons for the differential findings obtained with different methods.

Of note is that in models of SRT ISD and CRT ISD in which mean RT was entered as a covariate, cubic terms of mean RT were significant in the whole sample and in under 18s (linear terms of mean RT were significant for the adult portion of the sample). The fact that the mean RT removes some of the age polynomials could suggest that the relationship between RT mean and RT IIV varies with age—an idea mooted by Schmiedek, Lövdén and Lindenberger (2009). The finding that a cubic function described the relationship between RT ISD and mean RT is an important one. Many methods that control for mean RT assume that the relationship between the two variables is linear and stable across age. This assumption has been recently challenged (Schmiedek et al., 2009). It is important to explore the relationship between mean and variability in RT in more detail, as a better understanding of it would lead to more optimal methods of controlling for average RT when studying RT variability. Some of the approaches used at present, for example using CV or linear regression to partial effects of mean RT, may not be adequate ways of adjusting for individual differences in mean RT.

Task complexity and task characteristics

As mentioned in Chapter 1, some researchers have suggested that group differences in RT IIV—for example, between patients with traumatic brain injury and healthy controls or between older and younger healthy adults—are greater in more complex tasks (e.g. Stuss, Murphy, Binns, & Alexander, 2003; Dixon, Garrett, Lentz, MacDonald, Strauss, & Hultsch, 2007). However, there is no clear definition of task complexity and most comparisons are made between different tasks within a study, based on the author's subjective judgement. Tasks considered to be complex usually involve greater degree of processing, and especially operations such as

retention and manipulation of items in working memory or inhibition of irrelevant information. This has led some to suggest that, at a cognitive level, failures in executive functioning may have a special role in IIV (Hultsch, Strauss, Hunter, & MacDonald, 2008).

This thesis was concerned with simple and 4-choice RT tasks, neither of which involves complex processing and no executive functions were implied in these tasks. The idea was to explore various effects on IIV with focus on relatively basic processing speed rather than more complex mental processes. Nonetheless, some evidence for greater group differences in the more complex CRT task was found. Firstly, the meta-analytic estimates of age group differences were larger for CRT than SRT (Chapter 2). Moreover, an increase in IIV with age was observed in CRT but not SRT in the Science Festival sample (Chapter 3). Similarly, a longitudinal increase in CRT IIV but not SRT IIV with age was seen between ages 55 and 75 (Chapter 4). Taken together, these findings suggest that CRT variability is more affected by age than SRT IIV. It could also be that it is affected by age earlier than SRT (see Der & Deary, 2006). Since none of the empirical samples used for the investigations described in this thesis included individuals aged above 75 years, it was not possible to investigate whether SRT IIV is affected in very old age. The meta-analysis revealed an age group difference between older groups (including people in their 80s and 90s) and younger groups in SRT as well as in CRT variability, suggesting that SRT variability increases with age also, but perhaps the onset of the performance decline is later than for CRT IIV.

This thesis also demonstrated the importance of various task characteristics for the study of IIV in RT. In Chapter 5 variability was investigated on a trial-by-trial basis, and a clear pattern of systematic variations across trials was revealed. The differences between average RTs on different trials were substantial and magnitude and larger than sex differences in RT on any given trial. The consistency with which both sexes and all age groups followed the same trajectory of RTs across trials was remarkable. The potential to introduce trial-by-trial variability by varying characteristics such as stimulus sequence and (PI) variability and length should be given more attention by researchers studying RT IIV, especially in the light of findings suggesting that the effects of these factors differ by age (e.g., Lahtela,

Niemi, & Kuusela, 1985). Admittedly, the systematic review did not confirm a clear pattern of differences between older and younger adults based on task characteristics. However, small numbers of studies in various categories are the most likely explanation for the lack of clear, significant effects.

Another, although weaker, suggestion of the potential importance of the task and procedural factors is the discrepancy in findings of my trial-by-trial investigation of sex differences and of Reimers and Maylor (2006). My study revealed that females remained more variable than males across 40 trials of a CRT task, whereas Reimers and Maylor found that women were initially slower but became marginally faster than their male counterparts in a gender classification task. Among the potential reasons for the contrasting findings may be that a number of task and procedural factors differed between the two studies. The type of stimulus used (face vs. digit), the way they were presented (short, pre-determined period vs. remaining visible unless a response is made), the amount of practice (none vs. some), and the mode of administration (Internet vs. face-to-face) were among the few possible influences. It would be informative to experimentally investigate the importance of these differences and their potential to affect trial-by-trial performance of males and females.

Practice effects

Three of the studies included in this thesis allowed inferences to be drawn about practice effects on RT. On a small scale—within a task—CRTs slightly increased across the 40 test trials (Chapter 5). This was taken to mean that, rather than improving on the task over time, participants experience a little fatigue. Practice effects across occasions were found in the study of altitude effects (Chapter 6). Mean RT decreased steadily across the 25 testing sessions. Such a continuing pattern of improvement, with no evidence of the practice levelling off, was unexpected. They also contradicted the frequent reports of practice effects in RT diminishing (Abraini, Bouquet, Joulia, Nicolas, & Kriem 1998; Denison, Ledwith, Poulton, 1966) or disappearing with time (Parkin, Kerr, & Hindmarch, 1997; Rabbitt & Goward,

1994). One explanation for that was in terms of implicit learning of the stimulus sequence, which was the same across occasions. This suggestion requires further inquiry. Practice effects for RT variability were less marked and, when mean RT was controlled, RT IIV actually (though very slightly) increased across testing sessions.

The final study with potential to shed light on practice effects in RT, including RT variability, is the Twenty-07 study (Chapter 4). Here, practice effects were investigated longitudinally, across relatively long time periods. However, the previously-noted problems with apparatus changes across waves made drawing inferences problematic. It appears that some practice occurred across the four waves, as evidenced by a decreasing trend for CRT ISD that was observed in the youngest cohort. However, as discussed in Chapter 4, practice effects could not be modelled in this dataset due to collinearity of age and wave number. Because the age cohorts were very narrow and testing sessions were separated by similar intervals for all participants, in each cohort, age correlated very highly with the wave number. Therefore, the resulting longitudinal effects are necessarily a mixture of ageing effects and various wave effects (including practice, but not limited to it). Practice and age effects are likely to act in opposite directions: age is expected to be associated with a decline, and practice with an improvement in performance (e.g., Ferrer, Salthouse, Stewart, & Schwartz, 2004). Therefore, although the investigation demonstrated a clear increase in RT IIV in the oldest cohort in the Twenty-07 study (55-75 years), better estimates of longitudinal effects of age remain to be determined. It would be very useful for future investigations to attempt to separate the effects of age and practice more effectively, in order to obtain a better estimate of the increase in RT IIV in very old age. Modelling practice and age separately would also clarify whether the apparent lack of change in CRT ISD in mid-adulthood is due to age-related decline occurring is masked by practice (i.e., both effects counteract each other), or whether there is no effect of either age or practice on RT IIV in that age group.

Future directions and recommendations

This thesis investigated some correlates and determinants of RT variability, with novel empirical studies and formal meta-analysis of existing studies. It considered whether and under what circumstances factors such as age, sex, oxygen deprivation at high altitude, and various task characteristics are associated with RT IIV. However, the mechanisms of these associations are still to be explored and understood. The investigations described in this thesis contributed to knowledge about the development and ageing of SRT IIV and CRT IIV across the lifespan, about the sex differences, about the effects of altitude-related hypoxia, and about task characteristics which contribute to RT variability from trial to trial. Future investigations should focus on exploring some of the potential mechanisms behind these associations. For example, the role of sex hormones in RT IIV could be investigated. In particular, it remains to be explored whether the emergence of sex differences in RT inconsistency could be explained by the action of sex hormones on the brain.

There is growing popularity of RT variability as a performance measure; the number of studies investigating various influences on it is increasing. However, despite the knowledge of the positive correlation between mean RT and RT IIV, little is known about the precise shape of the relationship between the two variables. Some suggest that the relationship is largely linear, with deviations from linearity only at the extreme values of RT (Der & Deary, 2006). Others, for example Schmiedek et al. (2009), argue that the relationship is not linear and that it may be different in younger and older adults. If the suggestion of Schmiedek et al. is true, then many of the commonly-used methods of controlling for effects of mean RT may be inadequate. It is important that more research effort is directed to the exploration of the mean and variability relationship. Until more is known and understood about it, my main recommendation regarding the use of different measures of IIV, is to report findings obtained using both mean-adjusted and raw measures of variability. It is also advisable to use multiple methods of adjusting for mean RT and testing the robustness of any effects found across the different mean-adjusted IIV measures.

Longitudinal designs are very useful for studying development or ageing and have many benefits over cross-sectional studies. However, they also have their limitations, one of which is practice occurring from one to another measurement occasion. Unless practice effects are considered and controlled for, age effect estimates obtained from longitudinal analyses may be underestimated (Ferrer, et al. 2004). However, the way in which a study is designed might preclude investigating independent effects of age and practice; most notably this can occur when age is collinear with the measurement occasion number. Therefore, in designing future longitudinal studies, attempts should be made to allow separation of age and wave (or practice) effects. For example, to decrease the correlation between chronological age and wave number, measurements could be taken at irregular intervals, varying between subjects. Another way to reduce practice effects would be to use different versions of tasks at each measurement occasion. For example, different sequences of stimuli and possibly also PIs should be used in longitudinal investigations of RT in general and RT IIV in particular.

Following from the findings of the systematic trial-to-trial variations in RT when identical sequences of stimuli and PIs were used for all participants, the importance of procedure and task characteristics for RT variability is brought to the fore. It is important to pay closer attention to task characteristics, such as stimulus sequence or length of the PIs. More information about these should be given in study reports. These variables should also be considered in relation to different groups about which inferences are to be made. For example, interactions with age should be tested before conclusions about age effects on RT IIV are drawn.

Final summary

This thesis investigated correlates and determinants of RT IIV. The collective contribution of the studies included in it were more in terms of *what* and *when* than *how* and *why*. Therefore, a useful future direction would be to investigate some of the mechanisms bringing about inconsistency in RT. There is still much to be learned about the causes of RT IIV. Certainly much of the evidence supports a neural basis for RT variability, but the precise mechanisms bringing about inconsistency in RT

performance remain to be determined. It is also still unclear to what extent variability is a primary phenomenon and to what extent it is subservient to average speed. On the other hand, there is much more to be learned about the consequences of RT variability for cognition, for health, and for longevity. If RT variability is a phenomenon primary to, and causing to some extent, RT mean, it could become a valuable screening tool for early signs of neurological dysfunction with potential practical applications.

References

References marked with an asterisk indicate studies included in the meta-analysis

Abraini, J. H., Bouquet, C., Joulia, F., Nicolas, M., & Kriem, B. (1998). Cognitive performance during a simulated climb of Mount Everest: implications for brain function and central adaptive processes under chronic hypoxic stress. *Pflügers Archiv European Journal of Physiology*, *436*, 553-559.

*Adam, M., Retey, J. V., Khatami, R., & Landolt, H.-P. (2006). Age-related changes in the time course of vigilant attention during 40 hours without sleep in men. *Sleep*, *29*, 55-57.

Anstey, K. J. (1999). Sensorimotor variables and forced expiratory volume as correlates of speed, accuracy and variability in reaction time performance in late adulthood. *Aging, Neuropsychology, and Cognition*, *6*, 84-95.

*Anstey, K. J., Dear, K., Christensen, H., & Jorm, A. F. (2005). Biomarkers, health, lifestyle, and demographic variables as correlates of reaction time performance in early, middle, and late adulthood. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *58*, 5-21.

Beck, L. F. (1933). The role of speed in intelligence. *Psychological Bulletin*, *30*, 169–178.

Bellis, C. J. (1933). Reaction time and chronological age. *Proceedings of the Society for Experimental Biology & Medicine*, *30*, 801-803.

*Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Becic, E. (2006). Testing the limits of cognitive plasticity in older adults: Application to attentional control. *Acta Psychologica*, *123*, 261-278.

Blakemore, S-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, *47*, 296-312.

Bonnardel, R., & Liberson, W. (1933). Recherches sur la physiologie de l'homme aux hautes altitudes. *Le Travail Humain*, *1*, 432-444.

Bolmont, D., Thullier, F., & Abraini, J.H. (2000). Relationships between mood states and performances in reaction time, psychomotor ability, and mental efficiency during a 31-day gradual decompression in a hypobaric chamber from sea level to 8848 m equivalent altitude. *Physiology & Behavior*, *71*, 469-476

Bonney, K.R., Osvaldo, P.A., Flicker, L., Davies, S., Clarnette, R., Anderson, M., Lautenschlager, N.T. et al., (2006). Inspection time in non-demented older adults with mild cognitive impairment. *Neuropsychologia*, *44*, 1452-1456.

Borenstein, M., Hedges, L. V., Higgins, J. P. T, Rothstein, H.R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons Ltd.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B, 26,* 211–252.

Brebner, J. T. & Welford, A. T. (1980). Introduction: an historical background sketch. In A. T. Welford (Ed.), *Reaction Times.* Academic Press, New York, pp. 1-23.

Bruhn, P. & Parsons, O. A. (1971). Continuous reaction time in brain damage. *Cortex*, *7*, 278-291.

Buchanan, T. (2002). Online assessment: Desirable or dangerous? *Professional Psychology: Research and Practice*, *33*, 148-154.

Buchanan, T., Ali, T., Heffernan, T. M., Ling, J., Parrott, A. C., Rodgers, J., & Scholey, A. B.(2005). Non-equivalence of online and paper-and-pencil psychological tests: The case of the Prospective Memory Questionnaire. *Behavior Research Methods*, *37*, 148-154.

Buchanan, T. & Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the World-Wide Web. *British Journal of Psychology*, *90*, 125-144.

Bunce, D., Anstey, K. J., Cherbuin, N., Burns, R., Christensen, H., Wen, W., & Sachdev, P. S. (2010). Cognitive deficits are associated with frontal and temporal lobe white matter lesions in middle-aged adults living in the community. *PLoS ONE: e13567*. doi:10.1371/journal.pone.0013567.

Bunce, D., Anstey, K.J., Christensen, H., Dear, K., Wen, W., & Sachdev, P. (2007). White matter hyperintensities and within-person variability in community-dwelling adults aged 60-64 years. *Neuropsychologia*, *45*, 2009-2015.

*Bunce, D., Handley, R., & Gaines, S. O., Jr. (2008). Depression, anxiety, and within-person variability in adults aged 18 to 85 years. *Psychology and Aging*, *23*, 848-58.

*Bunce, D., MacDonald, S. W. S., & Hultsch, D. F. (2004). Inconsistency in serial choice decision and motor reaction times dissociate in younger and older adults. *Brain and Cognition*, *56*, 320-327.

*Bunce, D., Tzur, M., Rarachurn, A., Gain, F., & Bond, F. W. (2008). Mental health and cognitive function in adults aged 18 to 92 years. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences*, *63*, P67-P74.

Burns, N. R., & Nettelbeck, T. (2005). Inspection time and speed of processing: Sex differences on perceptual speed but not IT *Personality and Individual Differences 39*, 439–446.

Burton, C. L., Strauss, E., Hultsch, D. F., Moll, A., & Hunter, M. A. (2006). Intraindividual variability as a marker of neurological dysfunction: A comparison of Alzheimer's Disease and Parkinson's Disease. *Journal of Clinical and Experimental Neuropsychology*, *28*, 67-83.

Burton, C. L., Strauss, E., Hultsch, D. F., Hunter, M. A. (2009). The relationship between everyday problem solving and inconsistency in reaction time in older adults. *Aging, Neuropsychology and Cognition, 16*, 607-32.

Cattell, R.B. (1966). The data box: Its ordering of total resources in terms of possible relational systems. In R.B. Cattell (Ed). *Handbook of multivariate experimental psychology* (pp. 67-128). Chicago, USA: Rand McNally.

Cherbuin, N., Sachdev, P., & Anstey, K. J. (2010). Neuropsychological predictors of transition from healthy cognitive aging to mild cognitive impairment: the PATH Through Life Study. *American Journal of Geriatric Psychiatry*, *18*, 723-733.

Christensen, H., Dear, K.B.G., Anstey, K.J., Parslow, R.A., Sachdev, P., & Jorm, A.F. (2005). Within-occasion intraindividual variability and preclinical diagnostic status: Is intraindividual variability an indicator of mild cognitive impairment? *Neuropsychology*, *19*, 309-317.

Christensen, H., Mackinnon, A. J., Korten, A., & Jorn, A. F. (2001). The "common cause hypothesis" of cognitive aging: Evidence for not only a common factor but also specific associations of age with vision and grip strength in cross-sectional analysis. *Psychology and Aging, 16*, 588–599.

Christensen, H., Mackinnon, A.J., Korten, A.E., Jorm, A.F., Henderson, A.S., & Jacomb, P. (1999). Dispersion in cognitive ability as a function of age: A longitudinal study of an elderly community sample. *Aging, Neuropsychology, and Cognition*, *6*, 214-228.

Codorniu-Raga, M.J.,& Vigil-Colet, A.(2003). Sex differences in psychometric and chronometric measures of intelligence among young adolescents. *Personality and Individual Differences*, *35*, 681–689.

Cohen, J (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Collins, L.F., & Long, C.J. (1996). Visual reaction time and its relationship to neuropsychological test performance. *Archives of Clinical Neuropsychology, 11*, 613-623.

Cox, B. D., Huppert, F. A., & Whichelow, M. J. (1993). *The Health and Lifestyle Survey: seven years on.* Aldershot, U.K.: Dartmouth Publishing.

Crawford, T., Goodrich, S., Henderson, L., & Kennard, C. (1989). Predictive responses in Parkinson's disease: manual keypresses and saccadic eye movements to regular stimulus events. *Journal of Neurology, Neurosurgery and Psychiatry*, *52*, 1033-1042.

De Bellis, M. D., Keshavan, M. S., Beers, S. R., Hall, J., Frustaci, K., Masalehdan, A., … Boring, A.M. (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex*, *11*, 552-7.

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*, 2499-2507.

*Deary, I. J., & Der, G. (2005a). Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology, and Cognition*, *12*, 187-215.

Deary, I. J., & Der, G. (2005b). Reaction time explains IQ's association with death. *Psychological Science, 16*, 64-69.

Deary, I. J., Der, G., & Ford, G. (2001). Reaction times and intelligence differences: A population-based cohort study. *Intelligence, 29,* 389–399.

Denison, D. M., Ledwith, F., & Poulton, E. C. (1966). Complex reaction times at simulated cabin altitudes of 5,000 feet and 8,000 feet. *Aerospace Medicine, 37*, 1010-1013.

Department of Communities and Local Government (2007). Indices of Deprivation 2007 [data file]. Retrieved from http://www.communities.gov.uk/documents/communities/zip/indices2007.zip

Der, G. (1998). *A comparison of the West of Scotland Twenty-07 study sample and the 1991 Census SARs*. Working Paper No. 60. Glasgow: MRC Medical Sociology Unit.

*Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, *21*, 62-73.

Der, G., & Everitt, B. S. (2009). *A handbook of statistical analyses using SAS* (3rd ed.). Boca Raton, FL: Chapmann & Hall/CRC.

Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W. S., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. *Neuropsychology*, *21*, 381-399.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431.

*Duchek, J. M., Balota, D. A., Tse, C.S. Holtzman, D. M., Fagan, A. M., Goate, A. M. (2009). The utility of intraindividual variability in selective attention tasks as an early marker for Alzheimer's disease. *Neuropsychology*, *23*, 746-58.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*, 89- 98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455-463.

Dykiert, D., Der, G., Starr, M.A., & Deary, I.J. (in revision). Sex differences in reaction time mean and intra-individual variability across the life span. *Developmental Psychology.*

Dykiert, D., Gale, C.R., & Deary, I.J. (2009). Are apparent sex differences in mean IQ scores created in part by sample restriction and increased male variance? *Intelligence*, *37*, 42–47.

Ecob, R. (1987). *West of Scotland: the sampling frame and procedures for the cohort studies*. Working Paper No. 6. Glasgow: MRC Medical Sociology Unit.

Ecob, R, & Der G. (2003). An iterative method for the detection of outliers in longitudinal growth data using multilevel models. In: S. Reise, & N. Duan (eds). *Multilevel Modeling: Methodological Advances, Issues and Applications* (pp. 229-254). New Jersey: Lawrence Erlbaum.

Eckert, H. M., & Eichorn, D. H. (1977). Developmental variability in reaction time. *Child Development*, *48*, 452-358.

Fairweather, H., & Hutt, S. J. (1972). Sex differences in a perceptual-motor skill in children. In C. Ounsted, & D.C. Taylor (Eds.) *Gender Differences: Their Ontogeny and Significance.* London: Churchill Livingstone.

Ferrer, E., Salthouse, T.A., Stewart W.F., & Schwartz, B.S, (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging*, *19*, 243-259.

*Finkel, D., & McGue, M. (2007). Genetic and environmental influences on intraindividual variability in reaction time. *Experimental Aging Research*, *33*, 13-35.

Fiske, D.W. & Rice, L. (1955). Intra-individual response variability. *Psychological Bulletin*, *52*, 217-250.

Flehmig, H.C., Steinborn, M., Lagner, R., Scholz, A. & Westhoff, K (2007). Assessing intraindividual variabilty in sustained attention: reliability, relation to speed and accuracy, and practice effects. *Psychology Science*, *49*, 132-149.

*Fontani, G., Lodi, L., Felici, A., Corradeschi, F., & Lupo, C. (2004). Attentional, emotional and hormonal data in subjects of different ages. *European Journal of Applied Physiology*, *92*, 452-461.

Fozard, J. L., Thomas, J. C., & Waugh, N. C. (1976). Effects of age and frequency of stimulus repetitions on two-choice reaction time. *Journal of Gerontology*, *31*, 556-563.

*Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R.E. (1994). Age differences and changes in reaction time: The Baltimore Longitudinal Study of Aging. *Journals of Gerontology*, *49*, P179-P189.

Frey, D.J., Badia, P., & Wright Jr., K.P. (2004). Inter- and intra-individual variability in performance near the circadian nadir during sleep deprivation. *Journal of Sleep Research*, *13*, 305-315.

Gardepe, J. P., & Runcie, D. (1983) Heart rate and reaction times for mentally retarded and nonretarded adults in subject- and experimenter-initiated tasks. *American Journal of Mental Deficiency, 88*, 314 – 320.

Gerard, A. B., McElroy, M. K., Taylor, M. J., Grant, I., Powell, F. L., Holverda, S., et al. (2000). Six percent oxygen enrichment of room air at simulated 5000 m altitude improves neuropsychological function. *High Altitude Medicine & Biology*, *1*, 51-61.

Gilbert, J.A. (1894), Researches on the mental and physical development of school-children. *Studies from the Yale Psychological Laboratory*, *2*, 40-100. Retrieved from http://www.archive.org/stream/studiesfromyale00labogoog#page/n161/mode/1up

Godard, O., & Fiori, N. (2010). Sex differences in face processing: Are women less lateralized and faster than men? *Brain and Cognition, 73*, 167-175.

Gogaty, N., Giedd, J.N., Luysk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C. … & Thompson, P. M. (2004) Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Science of the USA, 101*, 8174-8179.

*Gooch, C. M., Stern, Y., & Rakitin, B. C. (2009). Evidence for age-related changes to temporal attention and memory from the choice time production task. *Aging Neuropsychology & Cognition*, *16*, 285-310.

Goodenough, F. L. (1935). The development of the reactive process from early childhood to maturity. *Journal of Experimental Psychology*, *18*, 431-450.

*Gorus, E., De Raedt, R., & Mets, T. (2006). Diversity, dispersion and inconsistency of reaction time measures: effects of age and task complexity. *Aging-Clinical & Experimental Research*, *18*, 407-417.

Hale, S., Myerson, J., Smith, G.A., & Poon, L. W. (1988). Age, variability, and speed: Between-subject diversity. *Psychology and Aging, 3*, 407-410.

Head, H. (1926). *Aphasia and kindred disorders of speech*. Cambridge: Cambridge University Press.

Hetherington, C. R., Stuss, D. T., & Finlayson, M. A. J. (1996). Reaction time and variability 5 and 10 years after traumatic brain injury *Brain Injury*, *10*, 473-486.

*Hogan, M. J. (2003). Average speed or variable speed--what do older adults really need? *Irish Journal of Psychology, 24*, 161-183.

Hendrickson, A. E. (1982). The biological basis of intelligence. Part I: Theory. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 151–196). Berlin: Springer-Verlag.

Hultsch, D. F., & MacDonald, S. W. S. (2004). Intraindividual variability in performance as a theoretical window onto cognitive aging. In R. A. Dixon, L. Bäckman, and L.-G/ Nilsson (Eds.), *New frontiers in cognitive aging* (pp. 65-88). New York, USA: Oxford University Press.

*Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *57*, P101-P115.

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Levy-Bencheton, J. & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology*, *14*, 588-598.

Hultsch, D. F., Strauss, E., Hunter, M. A., & MacDonald, S.W.S. (2008). Intraindividual variability, cognition, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.). *The handbook of aging and cognition* (3rd ed., pp. 491-556). New York: Psychology Press.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), A *Model for Intelligence* (pp. 93-132). Berlin: Springer-Verlag.

Jensen, A. R. (1987a). Individual differences in the Hick paradigm. In Vernon, P. A. (Ed.), *Speed of information processing and intelligence.* Norwood, NJ: Ablex.

Jensen, A. R. (1987b). Process differences and individual differences in some cognitive tasks. *Intelligence, 2,* 107-136.

Jensen, A. R. (1991). Speed of elementary cognitive processes: A chronometric anchor for psychometric tests of g. *Psychological Test Bulletin, 4, 59-70.*

Jensen, A. R. (1992). The importance of intraindividual variation in reaction time. *Personality and Individual Differences*, *13*, 869-881.

Jensen, A.R., (2006). Clocking the Mind: Mental Chronometry and Individual Differences. Amsterdam, The Netherlands: Elsevier.

Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahern, F. M. & Cole, R. E. (1985). Galtons data a century later. *American Psychologist*, *40*, 875-892.

Kalb, R., Jansen, S., Reulbach, U., & Kalb, S. (2004). Sex differences in simple reaction tasks. *Perceptual and Motor Skills*, *98*, 793-802.

Kellas, G. (1969). Reaction-time and response variability of normal and retarded individuals. *American Journal of Mental Deficiency*, *74*, 409-414.

Kirby, N.H., Nettelbeck, T., & Tiggemann, M. (1977). Reaction time in retarded and nonretarded young adults: sequential effects and response organization. *American Journal of Mental Deficiency*, *81*, 492-498.

Kobrick, J. L., & Dusek, E. R. (1970). Effects of hypoxia on voluntary response time to peripherally located visual stimuli. *Journal of Applied Physiology*, *29*, 444-448.

Koga, Y., & Morant, G. M. (1923). On the degree of association between reaction times in the case of different senses. *Biometrika*, *15*(3/4), 346–372.

Kirby, N. (1980). Sequential effects in choice reaction time. In A.T. Welford (ed.). *Reaction Times* (pp. 129-172). London, UK: Academic Press.

Kramer, A. F., Coyne, J. T., & Strayer, D. L. (1993). Cognitive function at altitude. *Human Factors, 35*, 329-344.

Lahtela, K., Niemi, P., Kuusela, V. (1985). Adult visual choice-reaction time, age, sex, and preparedness: A test of Welford's problem in a large population sample. *Scandinavian Journal of Psychology*, *26*, 357-362.

Landauer, A. A., Armstrong, S., & Digwood, J. (1980). Sex differences in choice reaction time. *British Journal of Psychology*, *71*, 551-555.

Lang, I. A., Llewellyn, D.J., Langa, K.M., Wallace, R. B., Huppert, F.A., Melzer, D. (2008). Neighborhood deprivation, individual socioeconomic status, and

cognitive function in older people: analyses from the English Longitudinal Study of Ageing. *Journal of the American Geriatric Society*, *56*, 191–8.
Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*, *14*, 309-325.
Ledwith, F. (1970). The effects of hypoxia on choice reaction time and movement time. *Ergonomics*, *13*, 465-482.
Lenroot, R. K., Gogtay, N., Greenstein, D. K., Molloy Wells, E., Wallace, G. L., Clasen, L. S., ... Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, *36*, 1065–1073.
Lewin, C. & Herlitz, A.(2002). Sex differences in face recognition—Women's faces make the difference. *Brain and Cognition*, *50*, 121-128.
Li, S.-C., Aggen, S .H., Nesselroade, J. R., & Baltes, P.B. (2001). Short-term fluctuations in elderly people's sensorimotor functioning in predict text and spatial memory performance: The MacArthur succesful aging studies. *Gerontology, 47*, 100-116.
Li, S. C., & Lindenberger, U. (1999). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilities in old age. In L.-G. Nilsson & H. Markowitsch (Eds.), *Cognitive neuroscience and memory* (pp. 103–146). Toronto, Canada: Hogrefe & Huber.
*Li, S.-C., Hammerer, D., Muller, V., Hommel, B., & Lindenberger, U. (2009). Lifespan development of stimulus-response conflict cost: Similarities and differences between maturation and senescence. *Psychological Research/Psychologische Forschung*, *73*, 777–785.
Lövdén, M., Li, S-C, Shing, Y.L. & Lindenberger, U. (2007). Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: Longitudinal data from the Berlin Aging Study. *Neuropsychologia*, *45*, 2827-2838.
Low, L-F., Anstey, K. J., Jorm, A. F., Christensen, H., & Rodgers, B. (2006). Hormone replacement therapy and cognition in an Australian representative sample aged 60-64 years. *Maturitas*, *54*, 86-94.
Lowery, D.P., Wesnes, K., Brewster, N., & Ballard, C. (2008). Subtle attentional deficits in the absence of dementia are associated with an increased risk of post-operative delirium. *International Journal of Geriatric Psychiatry*, *23*, 1253-1260.
Luciano, M., Smith, G.A., Wright, M.J., Geffen, G.M, Geffen, L.B., & Martin, N.G. (2001). On the heritability of inspection time and its covariance with IQ: a twin study. *Intelligence*, *29*, 443-457.
Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.
MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2003). Performance variability is related to change in cognition: Evidence from the Victoria Study. *Psychology and Aging*, *18*, 510–523.
MacDonald, S.W.S., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, *29*, 474-480.

Macintyre, S., Annandale, E., Ecob, R., Ford, G., Hunt, K., Jamieson, B., MacIver, S., West, P., & Wyke, S. (1989). The West of Scotland Twenty-07 Study: Health in the community. In C. J. Martin, & D. V. McQueen (Eds.), *Readings for a new public health* (pp. 56–74). Edinburgh, U.K.: Edinburgh University Press.

Mackintosh, J. H., Thomas, D. J., Olive, J. E., Chesner, I. M., & Knight, R. J. (1988). The effect of altitude on tests of reaction time and alertness. *Aviation Space & Environmental Medicine*, *59*, 246-248.

Maddox, G.L. & Douglass, E., (1974). Aging and individual differences: A longitudinal analysis of social, psychological and physiological indicators. Journal of *Gerontology*, *29*, 555–563.

*Martin, D. M., Burns, N. R., & Wittert, G. (2009). Free Testosterone Levels, Attentional Control, and Processing Speed Performance in Aging Men. *Neuropsychology*, *23*, 158-167.

Maruff, P., Falleti, M. G., Collie, A., Darby, D., & McStephen, M. (2005). Fatigue-related impairment in the speed, accuracy and variability of psychomotor performance: Comparison with blood alcohol levels. *Journal of Sleep Research*, *14*, 21-27.

McBain, R., Norton, D., & Chen, Y.(2009). Females excel at basic face perception. *Acta Psychological*, *130*, 168-173.

McEwen, B. S., (2001). Genome and Hormones: Gender Differences in Physiology: Invited Review: Estrogens effects on the brain: multiple sites and molecular mechanisms. *Journal of Applied Physiology*, *91*, 2785-2801.

McFarland, R. A. (1937a). Psycho-physiological studies at high altitude in the Andes I: The effects of rapid ascents by aeroplane and train. *Journal of Comparative Psychology*, *23*, 191-225.

McFarland, R. A. (1937b). Psycho-physiological studies at high altitude in the Andes II: Sensory and motor responses during acclimatization. *Journal of Comparative Psychology*, *23*, 227-258.

*McAuley, T., Yap, M., Christ, S. E., & White, D. E. A. (2006). Revisiting inhibitory control across the life span: Insights from the ex-gaussian distribution. *Developmental Neuropsychology*, *29*, 447-458.

Murtha, S., Cismaru, R., Waechter, R., & Chertkow, H. (2002). Increased variability accompanies frontal lobe damage in dementia. *Journal of the International Neuropsychological Society*, *8*, 360-372.

Nesselroade, J. R. (1991). The warp and and woof of the developmental fabric. In R. Downs, L. Liben, & D.S. Palermo (Eds.) *Can personality change?* (pp. 121-147). Washington, USA: American Psychological Association.

Nesselroade, J. R., & Salthouse,, T. A. (2004). Methodological and theoretical implications of intraindividual variability in perceptual-motor performance. *Journal of Gerontology: Psychological Science, 59B*, P49-P55.

Noble, C. E., Baker, B. L., & Jones, T. A. (1964). Age and sex parameters in psychomotor learning. *Perceptual and Motor Skills*, *19*, 935-945.

*Obrist, W. D. (1953). Simple auditory reaction time in aged adults. *Journal of Psychology: Interdisciplinary and Applied*, *35*, 259-266.

O'Carroll, R. E., & MacLeod, D. (1997). Moderate altitude has no effect on choice reaction time in international rugby players. *British Journal of Sports Medicine, 31*, 151-152.

Office For National Statistics (2006). National Statistics Postcode Directory [data file]. Census Geography Data Unit (UKBORDERS), EDINA (University of Edinburgh).

Office For National Statistics (n.d.). *Similarities and differences between the Indices of Deprivation across the UK. Retrieved from* http://www.neighbourhood.statistics.gov.uk/dissemination/Info.do?page=analysisandguidance/analysisarticles/indices-of-deprivation.htm.

Parkin, C., Kerr, J. S., & Hindmarch, I. (1997). The effects of practice on choice reaction time and critical flicker fusion threshold. *Human Psycholpharmacology, 12*, 65-70.

Paul, M. A., & Fraser, W. D. (1994). Performance during mild acute hypoxia. *Aviation Space & Environmental Medicine*, *65*, 891-899.

Peak, H., & Boring, E. G. (1926). The factor of speed in intelligence. *Journal of Experimental Psychology*, *9*, 71–94.

Pesta, B.J., Bertsch, S., Poznanski, P.J., & Bommer, W.H. (2008). Sex differences on elementary cognitive tasks despite no differenceson the Wonderlic Personnel Test. *Personality and Inividual Differences*, *45*, 429-431.

Pierson, W. R., & Montoye, H. J. (1958). Movement time, reaction time, and age. *Journal of Gerontology*, *13*, 418-421.

Petersen, R .C., & Kopmann, D. S. (2004). MCI is a clinically useful concept. *International Psychogeriatrics, 18*, 394-402.

Rabbitt, P., & Goward, L. (1994). Age, information processing speed, and intelligence. *The Quarterly Journal of Experimental Psychology*, *47 A*, 741-760.

Rabbitt, P., Diggle, P., Smith, D., Holland, F., & McInnes, L. (2001). Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia, 39*, 532 - 543.

Rabbitt, P., Osman, P., Moore, B., & Stollery, B. (2001). There are stable individual differences in performance variability, both from moment to moment and from day to day. *The Quarterly Journal of Experimental Psychology*, *54A*, 981-1003.

*Rakitin, B. C., Scarmeas, N., Li, T., Malapani, C., & Stern, Y. (2006). Single-dose levodopa administration and aging independently disrupt time production. *Journal of Cognitive Neuroscience*, *18*, 376-387.

Reed, C.L. & Franks, I.M. (1998) Evidence for movement preprogramming and on-line control in differentially impaired patients with Parkinson's Disease. *Cognitive Neuropsychology*, *15*, 723-745.

Reed, T. E., Vernon, P. A., & Johnson, A. M. (2004). Sex differences in brain nerve conduction velocity in normal humans. *Neuropsychologia*, *42*, 1709–1714.

Reimers, S., & Maylor, E.A. (2006). Gender effects on reaction time variability and trial-to-trial performance: Reply to Deary and Der (2005). *Aging, Neuropsychology, and Cognition*, *13*, 479–489.

Regard, M., & Landis, T. (1991). Cognitive changes at high altitude on healthy climbers developing acute mountain sickness. *Aviation Space and Environmental Medicine*, *62*, 291-295.

Roach R.C., Bartsch P., Hackett P.H., Oelz O. (1993). The Lake Louise acute mountain sickness scoring system. In: J.R. Sutton, G. Coates, C.S. Huston

(Eds.), *Hypoxia and Molecular Medicine: Proceedings of the 8th International Hypoxia Symposium* (pp. 272-274). Burlington, VT: Queen City Printers.
Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.
Salthouse, T. A. (2007). Implications of within-person variability in cognitive and neuropsychological functioning for the interpretation of change. *Neuropsychology, 21*, 401-411.
Salthouse, T. A. (1996). Relation of successive percentiles of reaction time distributions to cognitive variables and adult age. *Intelligence, 26*, 153-166.
Salthouse, T. A., & Berish, D. E. (2005). Correlates of within-person (across-occasion) variabilityin reaction time. *Neuropsychology, 19*, 77-87.
Schaie, K.W. (2005). *Developmental influences in adult intelligence: The Seattle Longitudinal Study*. New York, USA: Oxford University Press.
Schmiedek, F., Lövdén, M., & Lindenberger, U. (2009). On the relation of mean reaction time and intraindividual reaction time variability. *Psychology and Aging*, *24*, 841-857.
Scottish Government (2006). *Scottish Index of Multiple Deprivation 2006 - background data* [data file]. Retrieved from http://www.scotland.gov.uk/Resource/Doc/933/0041675.xls
*Shammi, P., Bosman, E., & Stuss, D. T. (1998). Aging and variability in performance. *Aging, Neuropsychology, and Cognition*, *5*, 1-13.
Shipley, B. A., Der, G.., Taylor, M. D., & Deary, I. J. (2007). Association between mortality and cognitive change over 7 years in a large representative sample of UK residents. *Psychosomatic Medicine*, *69*, 640-650.
Shipley, B. A., Der, G., Taylor, M. D., & Deary, I. J. (2006). Cognition and all-cause mortality across the entire adult age range: Health and lifestyle survey. *Psychosomatic Medicine*, *68*, 17–24.
Silverman, I. W. (2006). Sex differences in simple visual reaction time: A historical meta-analysis. *Sex Roles*, *54*, 57-68.
Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, *23*, 323-355.
Sliwinski, M. J., Smyth, J. Hofer, S. M., & Stawski, R. S. (2006). Inraindividual coupling of daily stress and cognition. *Psychology and Aging, 21*, 545-557.
*Smulders, F. T., Kenemans, J. L., Jonkman, L. M., & Kok, A. (1997). The effects of sleep loss on task performance and the electroencephalogram in young and elderly subjects. *Biological Psychology*, *45*, 217-239.
Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., Toga, A. W. (2004). Mapping cortical change across the human life span. *Nature Neuroscience, 6*, 309-315.
*Sparrow, W. A., Begg, R. K., & Parker, S. (2006). Aging effects on visual reaction time in a single task condition and when treadmill walking. *Motor Control*, *10*, 201-211.
Spirduso, W. W., & Clifford, P. (1978). Replication of age and physical activity effects on reaction and movement time. *Journal of Gerontology*, *33*, 26-30.
Strauss, E., Bielak, A. A. M., Bunce, D., Hunter, M. A., & Hultsch, D. F. (2007). Within-person variability in response speed as an indicator of cognitive

impairment in older adults. *Aging, Neuropsychology, and Cognition*, *14*, 608-630.
Stuss, D. T., Murphy, K. J., Binns, M. A., & Alexander, M. P. (2003). Staying on the job: The frontal lobes control individual performance variability. *Brain*, *126*, 2363-2380.
Stuss, D. T., Pogue, J., Buckle, L., & Bondar, J. (1994). Characterization of stability of performance in patients with traumatic brain injury: Variability and consistency on reaction time tests. *Neuropsychology*, *8*, 316-324.
Stuss, D. T., Stethem, L. L., Hugenholtz, H., Picton, T., Pivik, J., & Richard, M. T. (1989). Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance. *Journal of Neurology, Neurosurgery, and Psychiatry*, *52*, 742–748.
Surwillo, W. W. (1963). The relation of response-time variability to age and the influence of brain wave frequency. *Electroencephalography & Clinical Neurophysiology*, *15*, 1029-1032.
Vernon, P. A. (1981). Reaction time and intelligence in the mentally retarded. *Intelligence*, *5*, 345-355.
Virués-Ortega, J., Buela-Casal, G., Garrido, E., & Alcázar, B. (2004). Neuropsychological functioning associated with high-altitude exposure. *Neuropsychology Review*, *14*, 197-224.
Walker, M. P., Ayre, G. A., Ashton, C. H., Marsh, V. R., Wesnes, K., Perry, E. K., O'Brien, J. T., McKeith, I. G. & Ballard, C. G. (1999). A psychophysiological investigation of fluctuating consciousness in neurodegenerative dementias. *Human Psychopharmacology*, *14*, 483-489.
Walker M. P., Ayre G. A., Cummings J. L., Wesnes K., McKeith I. G., O'Brien J. T. & Ballard C. G. (2000). Quantifying fluctuation in dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. *Neurology*, *54*, 1616-1625.
Walker, M. P., Ayre, G. A., Perry, E. K., Wesnes, K., McKeith, I. G., Tovee, M., Edwardson, J. A. & Ballard, C. G. (2000). Quantification and characterisation of fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *11*, 327-335.
Waugh, N. C., Fozard, J. L., Talland, G. A., & Erwin, D. E. (1973). Effects of age and stimulus repetition on two-choice reaction time. *Journal of Gerontology*, *28*, 466-470.
Wegesin, D. J. & Stern, Y (2004). Inter- and intraindividual variability in recognition memory: Effects of age and estrogen use. *Neuropsychology*, *18*, 646-657.
Welford, A.T. (1980). Choice reaction time: basic concepts. In A.T. Welford (ed.). *Reaction Times* (pp.73-128). London, UK: Academic Press.
West, J. B. (1984). Human physiology at extreme altitudes on Mount Everest. *Science*, *223*, 784-788.
*West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. *Brain and Cognition*, *49*, 402-419.
West, J. B., Schoene, R. B., & Milledge, J. S. (2007). *High Altitude Medicine and Physiology* (4th ed.): Hodder Arnold.
Wesensten, N. J., Crowley, J. B., Balkin, T., Kamimori, G., Iwanyk, E., Pearson, N., et al. (1993). Effects of simulated high altitude exposure on long latency event-

related brain potentials and performance. *Aviation Space and Environmental Medicine*, *64*, 30-36.

*Wilkinson, R. T., & Allison, S. (1989). Age and simple reaction time: Decade differences for 5,325 subjects. *Journals of Gerontology*, *44*, 29-35.

*Williams, B. R., Hultsch, D. F., Strauss, E. H., Hunter, M. A., & Tannock, R. (2005). Inconsistency in Reaction Time Across the Life Span. *Neuropsychology*, *19*, 88-96.

*Williams, B. R., Strauss, E. H., Hultsch, D. F., & Hunter, M. A. (2007). Reaction time inconsistency in a spatial Stroop task: Age-related differences through childhood and adulthood. *Aging, Neuropsychology, and Cognition*, *14*, 417-439.

Wolf, F.M. (1986). *Meta-analysis: Quantitative Methods for Research Synthesis*. Newbury Park, CA: Sage Publications, Inc.

*Yan, J. H., Thomas, J. R., & Stelmach, G. E. (1998). Aging and rapid aiming arm movement control. *Experimental Aging Research*, *24*, 155-168.

Appendices

Appendix A. Search terms used for each database

PsycINFO

The first, general search of PsycINFO database was done through EBSCO Host gateway, using advanced search facility of all available fields. The search terms used were:

((reaction time) OR (cognition) OR (processing speed) OR (response speed))
AND
(variability OR inconsistency OR lability OR instability).

The update search was performed via OvidSP using the following search terms:
1. reaction time.mp.
2. exp Reaction Time/
3. cognition.mp.
4. exp Cognition/
5. processing speed.mp.
6. exp Cognitive Processing Speed/
7. response time.mp.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. variability.mp.
10. inconsistency.mp.
11. lability.mp.
12. instability.mp.
13. standard deviation/
14. 9 or 10 or 11 or 12 or 13
15. 8 and 14
16. limit 15 to yr="2008 - 2009"

MEDLINE and EMBASE

Searched via Ovid SP gateway with the same search criteria used for both databases:

1. reaction time.mp.
2. processing speed.mp.
3. response speed.mp.
4. cognition.mp.
5. variability.mp.
6. inconsistency.mp.
7. lability.mp.
8. instability.mp.
9. 1 or 2 or 3 or 4
10. 5 or 6 or 7 or 8
11. 9 and 10

Web of Science

Accessed via ISI Web of Knowledge Platform and searched using the following search strategy:

1. Topic=(reaction time) OR Title=(reaction time)
2. Topic=(processing speed) OR Title=(processing speed)
3. Topic=(response speed) OR Title=(response speed)
4. Topic=(cognition) OR Title=(cognition)
5. Topic=(variability) OR Title=(variability)
6. Topic=(inconsistency) OR Title=(inconsistency)
7. Topic=(lability) OR Title=(lability)
8. Topic=(instability) OR Title=(instability)
9. 1 OR 2 OR 3 OR 4
10. 5 OR 6 OR 7 OR 8
11. 9 AND 10

Appendix B. Data extraction sheet

Study ID: ____

Data Extraction Sheet

Author(s)	**Year**
Title	**Publication type**
Citation	**Source** ☐ Database search ☐ Hand search

Verification of study eligibility	**Corresponding author's details**
☐ Correct population investigated ☐ Empirical data collected ☐ Appropriate RT task administered ☐ Intra-individual variability measured ☐ Matter relevant	**Author to be contacted? YES/NO** **Date contacted:** _____

Design, sample and data collection

Study design Cross-sectional / longitudinal	Details:
If applicable, name of the study/project/dataset:	**Country where data collected:**

Sample size:	**N of males/females: /**	**Overall age range:**
Description of the sample (e.g. community dwelling adults)		
Sampling method (initial sample)	☐ Representative (Random sampling or whole population) ☐ Likely to be representative (systematic, stratified sampling) ☐ Likely to be non-representative (Non-probability sampling, e.g. purposive, convenience, snowball sampling) ☐ No sufficient information Details: ……	

If applicable, details of comparisons between groups on demographic/ cognitive measures etc	Group	Young	Old	Comments/Notes:
	Education(yrs)			
	IQ			
	Other _____ _____ _____ _____			

Task details

General and data treatment

Apparatus used	1. Purpose-built RT apparatus or response console 2. Standard PC and keyboard/mouse 3. Other _____ 4. Not stated	
Precision of measurement	±0.001s ±0.01 ±0.1 other: ____ Not stated	
Reaction time measure available	1. Decision time 2. Movement time 3. Composite only (DT+MT) 4. Other _____	1. Mean 2. Median 3. Other _____
Variability measure used	1. ISD 2. ISD/mean 3. CV (SD/mean * 100) 4. Other _____	
Trial-level data cleaning	☐ Performed ☐ Not performed / Not stated	
RT data trimming etc.	SRT CRT Lower cut-off: _____ _____ Upper cut-off: _____ _____ Excluded (n or %): _____ _____	Other details:
Missing data treatment (describe)		
Other data preparation details		

Simple RT

Stimulus modality/type 1. Visual a) Light d) timer b) Digit e) word c) Letter f) other 2. Auditory 3. Other 4. Not stated Details:	**Response modality** 1. Manual a) button press b) button release c) other 2. Oral 3. Other 4. Not stated Details:
Was location of the stimulus varied (were different stimuli presented in different locations)? 1. Yes 2. No 3. N/A	
Instructions emphasized: 1. Speed 2. Accuracy (avoid pre-presses) 3. Neither 1 or 2 4. Not stated	**If hand used to respond, specify:** 1. Left 3. Preferred 5. Not stated 2. Right 4. Non-preferred **Was the choice of hand counter-balanced?** 1. Yes 2. No 3. Not stated 4. N/A
ISI 1. Fixed a) Same value b) Different 2. Variable 3. Not stated Values:	**Number of practice trials**: _____ **Number of test trials:** _____

Choice RT

Stimulus modality/type 1. Visual a) Light d) timer b) Digit e) word c) Letter f) other 2. Auditory 3. Other 4. Not stated Details:	**Response modality** 1. Manual a) Button press b) Button release c) Other 2. Oral 3. Other 4. Not stated Details:
Was location of the stimulus varied? 1. Yes 2. No 3. N/A	**Number of possible choices (circle):** 2 3 4 5 6 7 8 other:
Instructions emphasized: 1. Speed 2. Accuracy (avoid errors) 3. Neither 1 or 2 4. Not stated	**If hand used to respond, specify:** 1. Left 3. Preferred 5. Not stated 2. Right 4. Non-preferred **Was the choice of hand counter-balanced?** 1. Yes 2. No 3. N/A
ISI 1. Fixed a) Same value b) Different values 2. Variable 3. Not stated Values:	**Number of practice trials**: _____ **Number of test trials:** _____ **Number of blocks**: _____

Results

For group comparisons, specify for each group:

Age range						
Mean (SD)						
n for group						
Gender distr. (M/F)	/	/	/	/	/	/
SRT Mean (SD)						
SRT IIV (SD)						
CRT Mean (SD)						
CRT IIV (SD)						

Statistical tests results

(Include test statistic and p value in full if reported; if applicable, specify covariates)

Appendix C. Details of data extraction from studies for which sufficient information were not reported in the paper nor obtainable from the authors.

Study	Data available	Method used to extract the data	Obtained values
Hogan, M. J. (2003). Average speed or variable speed--what do older adults really need? Irish Journal of Psychology, 24(3), 161-183.	From Figure 1 - age-group variance (R^2) in SRT ISD and CRT ISD from regression models controlling for the relevant RT median.	1. Digitised graph to obtain R^2 2. Converted of R^2 into Cohen's d using the following formula: $d = \frac{2R}{\sqrt{1-R^2}}$ [a]	Values for adjusted SRT and CRT: SRT ISD R^2 for age = CRT ISD R^2 for age = SRT IIV d from R^2 = **0.344** SRT IIV d from R^2 = **0.050**
Pierson, W. R., & Montoye, H. J. (1958). Movement time, reaction time, and age. Journal of Gerontology, 13, 418-421.	From Figure 3 – mean frequency of intra-individual mode for different age groups	1. Digitised graph to obtain estimates of the mean frequency of mode 2. Averaged the values across groups to obtain the required age ranges	Frequency of mode for age ranges: 19-20: 4.427 21-25: 4.118 26-30: 5.01 } **19-30: 4.518** 41-45: 3.275 46-55: 2.964 } **41-55: 3.120** 66-75: 2.014 76-85: 1.93 } **66-85: 1.972**
Shammi, P., Bosman, E., & Stuss, D. T. (1998). Aging and variability in performance. Aging, Neuropsy-chology, and Cognition, 5, 1-13.	Results from a significance tests; main effect of age group (young, old) on raw CRT ISD: F(1.32)=4.99, p<.03	1. Converted F to Cohen's d using the following formula: $d = \frac{2\sqrt{F}}{\sqrt{df\,(error)}}$ [a]	Effect size for old-young difference: **d = 0.790**

Study	Data available	Method used to extract the data	Obtained values
Smulders, F. T., Kenemans, J. L., Jonkman, L. M., & Kok, A. (1997). The effects of sleep loss on task performance and the electroencephalogram in young and elderly subjects. Biological Psychology, 45(1-3), 217-239.	Results from a significance tests; main effect of age group (young, old) on raw CRT IQR: F(1.22)=5.4, p<.05	1. Converted F to Cohen's d using the following formula: $d = \frac{2\sqrt{F}}{\sqrt{df\,(error)}}$ [a]	Effect size for old-young difference: **d = 0.991**
Sparrow, W. A., Begg, R. K., & Parker, S. (2006). Aging Effects on Visual Reaction Time in a Single Task Condition and When Treadmill Walking. Motor Control, 10(3), 201-211.	From Figure 2 – means and SD of RTs for each participant	1. Digitise graph to obtain the values for individual RT mean and SD 2. Calculate mean RT ISD for young and old groups 3. Calculate CV for each individual and obtain mean CV for young and old groups	Mean ISD (SD) and Mean CV (SD) for young and old groups: Young ISD = **0.046 (0.010)** Old ISD = **0.061 (0.032)** Young CV = **0.172 (0.028)** Old CV = **0.200 (0.089)**
Surwillo, W. W. (1963). The relation of response-time variability to age and the influence of brain wave frequency. Electroencephalography & Clinical Neurophysiology, 15, 1029-1032.	Correlation coefficient for age and ISD: r=0.26, p<.01 But age entered as a continuous variable and not discrete groups, so converting r into d would not yield d comparable to other studies. Regression equation: ISD=0.18 Age + 19.29	1. Calculate predicted values from the regression equation for midpoints of age ranges of interest	Predicted values for ISD: At 45 predicted ISD = **27.39** At 65 predicted ISD = **30.99**

Study	Data available	Method used to extract the data	Obtained values
Wilkinson, R. T., & Allison, S. (1989). Age and simple reaction time: Decade differences for 5,325 subjects. Journals of Gerontology, 44(2), 29-35.	Mean difference for old vs. young and old vs. middle-aged available, but no exact p values reported. From Figure 2 – mean CV with 95% CI for different age groups	1. Digitised graph to obtain mean CV and 95%CI 2. Converted 95% CI to SD	Mean CV (SD) 20-29 CV = **63.298 (18.763)** 40-49 CV = **66.062 (31.389)** 60-69 CV = **73.839 (41.333)**

[a] Formula from Wolf (1986)

Appendix D. Excluded studies - general search

No.	Study	Reason for exclusion[a]
1	Aartsen, M. J., Martin, M., Zimprich, D., & Longitudinal Aging Study, A. (2004). Gender differences in level and change in cognitive functioning. Results from the Longitudinal Aging Study Amsterdam. Gerontology, 50(1), 35-38.	No RT
2	Abrahams, S., Leigh, P. N., & Goldstein, L. H. (2005). Cognitive change in ALS: a prospective study. Neurology, 64(7), 1222-1226.	Wrong task
3	Abrams, R. A., Pratt, J., & Chasteen, A. L. (1998). Aging and movement: variability of force pulses for saccadic eye movements. Psychology & Aging, 13(3), 387-395.	Wrong task
4	Advokat, C., Martino, L., Hill, B. D., & Gouvier, W. (2007). Continuous Performance Test (CPT) of college students with ADHD, psychiatric disorders, cognitive deficits, or no diagnosis. Journal of Attention Disorders, 10(3), 253-256.	Wrong task
5	Ahsberg, E., Gamberale, F., & Gustafsson, K. (2000). Perceived fatigue after mental work: an experimental evaluation of a fatigue inventory. Ergonomics, 43(2), 252-268.	No IIV
6	Alarcon, M., Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1998). Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado adoption project. Behavior Genetics, 28(4), 255-264.	No RT
7	Aldenkamp, A. P., van Meel, H. F., Baker, G. A., Brooks, J., & Hendriks, M. P. H. (2002). The A-B neuropsychological assessment schedule (ABNAS): the relationship between patient-perceived drug related cognitive impairment and results of neuropsychological tests. Seizure, 11(4), 231-237.	No IIV
8	Allain, H., & Bentue-Ferrer, D. (1998). Clinical efficacy of almitrine-raubasine. An overview. European Neurology, 39 Suppl 1, 39-44.	No empirical data
9	Allaire, J. C., & Marsiske, M. (2005). Intraindividual Variability May Not Always Indicate Vulnerability in Elders' Cognitive Performance. Psychology and Aging, 20(3), 390-401.	No RT
10	Allen, P. A. (1990). Influence of processing variability on adult age differences in memory distribution of order information. Cognitive Development, 5(2), 177-192.	Wrong task
11	Alvarez, J. A., & Freides, D. (2004). Research on attention deficit hyperactivity disorder using the covert orienting paradigm of Posner. Developmental Neuropsychology, 26(2), 627-645.	No empirical data
12	Alvarez, L., Macias, R., Lopez, G., Alvarez, E., Pavon, N., Rodriguez-Oroz, M. C., et al. (2005). Bilateral subthalamotomy in Parkinson's disease: initial and long-term response. Brain, 128(3), 570-583.	No RT
13	American Academy of Pediatrics. Behavioral and cognitive effects of anticonvulsant therapy. Committee on Drugs. (1985). Pediatrics, 76(4), 644-647.	No empirical data
14	Ando, S., Kida, N., & Oda, S. (2001). Central and peripheral visual reaction time of soccer players and nonathletes. Perceptual and Motor Skills, 92(3), 786-794.	No IIV
15	Andre-Petersson, L., Elmstahl, S., Hagberg, B., Janzon, L., Reinprecht, F., & Steen, G. (2003). Is blood pressure at 68 an independent predictor of cognitive decline at 81? Results from follow-up study "Men born in 1914", Malmo, Sweden. Aging & Mental Health, 7(1), 61-72.	No RT
16	Anson, O., & Paran, E. (2005). Hypertension and cognitive functioning among the elderly: an overview. American Journal of Therapeutics, 12(4), 359-365.	No empirical data
17	Anstey, K. J. (2004). Within-Person Variability as a Dynamic Measure of Late-Life Development: New Methodologies and Future Directions. Gerontology, 50(4), 255-258.	No empirical data
18	Archer, T. (2005). Comorbidity influences in neuropsychiatry: Symptom plasticity constraints on etiology and treatment. European Journal of Psychiatry, 19(4), 197-207.	No empirical data
19	Armstrong, J. G., & Loewenstein, R. J. (1990). Characteristics of patients with multiple personality and dissociative disorders on psychological testing. Journal of Nervous & Mental Disease, 178(7), 448-454.	No RT
20	Arnedt, J. T., Owens, J., Crouch, M., Stahl, J., & Carskadon, M. A. (2005). Neurobehavioral Performance of Residents After Heavy Night Call vs After Alcohol Ingestion. JAMA: Journal of the American Medical Association, 294(9), 1025-1033.	No IIV
21	Artes, P. H., McLeod, D., & Henson, D. B. (2002). Response time as a discriminator between true- and false-positive responses in suprathreshold perimetry. Investigative Ophthalmology & Visual Science, 43(1), 129-132.	Wrong task
22	Auriel, E., Hausdorff, J. M., Herman, T., Simon, E. S., & Giladi, N. (2006). Effects of methylphenidate on cognitive function and gait in patients with Parkinson's disease: a pilot study. Clinical Neuropharmacology, 29(1), 15-17.	Wrong task
23	Avni, N., Avni, I., Barenboim, E., Azaria, B., Zadok, D., Kohen-Raz, R., et al. (2006). Brief posturographic test as an indicator of fatigue. Psychiatry and Clinical Neurosciences, 60(3), 340-346.	No IIV
24	Bachoud-Levi, A. C., Maison, P., Bartolomeo, P., Boisse, M. F., Dalla Barba, G., Ergis, A. M., et al. (2001). Retest effects and cognitive decline in longitudinal follow-up of patients with early HD. Neurology, 56(8), 1052-1058.	No IIV
25	Backman, L., Wahlin, Ã. k., Small, B. J., Herlitz, A., Winblad, B., & Fratiglioni, L. (2004). Cognitive functioning in aging and Dementia: The Kungsholmen project. Aging, Neuropsychology, and Cognition, 11(2), 212-244.	No RT
26	Balfour, C. A., Clark, C. R., & Geffen, G. M. (1991). Interhemispheric differences in the planning and execution of sequences of skilled finger movements. Neuropsychologia, 29(9), 889-905.	No IIV
27	Balfour, D., & Le Houezec, J. (2002). Advances in neuroscience and pharmacology of Nicotine. Nicotine and Tobacco Research, 4(2), 229-244.	No empirical data

28	Ballard, C., Walker, M., O'Brien, J., Rowan, E., & McKeith, I. (2001). The characterisation and impact of 'fluctuating' cognition in dementia with Lewy bodies and Alzheimer's disease. International Journal of Geriatric Psychiatry, 16(5), 494-498.	Wrong task
29	Banks, S., & Dinges, D. F. (2007). Behavioral and physiological consequences of sleep restriction. Journal of Clinical Sleep Medicine, 3(5), 519-528.	No empirical data
30	Barrett, N. C., & Glencross, D. J. (1991). Visually based amendments of aiming movements: a reappraisal of Christina (1970). Perceptual & Motor Skills, 72(2), 531-535.	No empirical data
31	Barrett, P. T., & Eysenck, H. J. (1992). Brain Evoked-Potentials and Intelligence - the Hendrickson Paradigm. Intelligence, 16(3-4), 361-381.	No empirical data
32	Bartolomeo, P., SiÃ©roff, E., Chokron, S., & Decaix, C. (2001). Variability of response times as a marker of diverted attention. Neuropsychologia, 39(4), 358-363.	Wrong task
33	Bartres-Faz, D., Junque, C., Clemente, I. C., Serra-Grabulosa, J. M., Guardia, J., Lopez-Alomar, A., et al. (2001). MRI and genetic correlates of cognitive function in elders with memory impairment. Neurobiology of Aging, 22(3), 449-459.	No RT
34	Bartres-Faz, D., Junque, C., Serra-Grabulosa, J. M., Lopez-Alomar, A., Moya, A., Bargallo, N., et al. (2002). Dopamine DRD2 taq I polymorphism associates with caudate nucleus volume and cognitive performance in memory impaired subjects. NeuroReport, 13(9), 1121-1125.	No RT
35	Bashore, T. R. (1981). Vocal and manual reaction time estimates of interhemispheric transmission time. Psychological Bulletin, 89(2), 352-368.	No empirical data
36	Bates, M. E., & Pandina, R. J. (1992). Familial alcoholism and premorbid cognitive deficit: a failure to replicate subtype differences. Journal of Studies on Alcohol, 53(4), 320-327.	No RT
37	Baumeister, A. A. (1968). Behavioral inadequacy and variability of performance. American Journal of Mental Deficiency, 73(3), 477-483.	No empirical data
38	Baumeister, A. A. (1998). Intelligence and the 'personal equation '. Intelligence, 26(3), 255-265.	No empirical data
39	Baune, B. T., Ponath, G., Rothermundt, M., Riess, O., Funke, H., & Berger, K. (2008). Association between genetic variants of IL-1beta, IL-6 and TNF-alpha cytokines and cognitive performance in the elderly general population of the MEMO-study. Psychoneuroendocrinology, 33(1), 68-76.	No RT
40	Bayles, K. A., Tomoeda, C. K., McKnight, P. E., Helm-Estabrooks, N., & Hawley, J. N. (2004). Verbal perseveration in individuals with Alzheimer's disease. Seminars in Speech & Language, 25(4), 335-347.	No RT
41	Beblo, T., Saavedra, A. S., Mensebach, C., Lange, W., Markowitsch, H.-J., Rau, H., et al. (2006). Deficits in visual functions and neuropsychological inconsistency in Borderline Personality Disorder. Psychiatry Research, 145(2-3), 127-135.	No RT
42	Beglinger, L. J., Crawford-Miller, J., Kareken, D. A., Ahmed, S., Derby, M. A., Siemers, E., et al. (2003). Neuropsychological practice effects and change detection in people with schizophrenia. Schizophrenia Research, 62(1), 191-194.	Wrong IIIV
43	Bell, E. C., Willson, M. C., Wilman, A. H., Dave, S., & Silverstone, P. H. (2006). Males and females differ in brain activation during cognitive tasks. Neuroimage, 30(2), 529-538.	No IIV
44	Bellelli, G., Pezzini, A., Bianchetti, A., Trabucchi, M., Birkenhager, W. H., Forette, F., et al. (2002). Increased blood pressure variability may be associated with cognitive decline in hypertensive elderly subjects with no dementia [1] (multiple letters). Archives of Internal Medicine, 162(4), 483-484.	No RT
45	Bellgrove, M. A., Hester, R., & Garavan, H. (2004). The functional neuroanatomical correlates of response variability: evidence from a response inhibition task. Neuropsychologia, 42(14), 1910-1916.	Wrong task
46	Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St George, M. (2000). The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. Journal of Cognitive Neuroscience, 12, 7-29.	No RT
47	Berntson, G. G., Cacioppo, J. T., Binkley, P. F., Uchino, B. N., Quigley, K. S., & Fieldstone, A. (1994). Autonomic Cardiac Control .3. Psychological Stress and Cardiac Response in Autonomic Space as Revealed by Pharmacological Blockades. Psychophysiology, 31(6), 599-608.	No IIV
48	Berr, C., Richard, F., Dufouil, C., Amant, C., Alperovitch, A., & Amouyel, P. (1998). Polymorphism of the prion protein is associated with cognitive impairment in the elderly: the EVA study. Neurology, 51(3), 734-737.	No IIV
49	Berryman, M. L., & Kennelly, K. J. (1992). Letter memory loads change more than visual-field advantage: Interhemispheric coupling effects. Brain and Cognition, 18(2), 152-168.	Matter not relevant
50	Bever, C. T., Jr., Grattan, L., Panitch, H. S., & Johnson, K. P. (1995). The Brief Repeatable Battery of Neuropsychological Tests for Multiple Sclerosis: a preliminary serial study. Multiple Sclerosis, 1(3), 165-169.	No RT
51	Bielak, A. A. M., Hultsch, D. F., Levy-Ajzenkopf, J., Macdonald, S. W. S., Hunter, M. A., & Strauss, E. (2007). Short-term changes in general and memory-specific control beliefs and their relationship to cognition in younger and older adults. International Journal of Aging & Human Development, 65(1), 53-71.	Wrong task
52	Birkett, P., Sigmundsson, T., Sharma, T., Toulopouiou, T., Grifflths, T. D., Reveley, A., et al. (2007). Reaction time and sustained attention in schizophrenia and its genetic predisposition. Schizophrenia Research, 95(1-3), 76-85.	Wrong task
53	Black, D. N., Stip, E., Bedard, M., Kabay, M., Paquette, I., & Bigras, M. J. (2000). Leukotomy revisited: late cognitive and behavioral effects in chronic institutionalized schizophrenics. Schizophrenia Research, 43(1), 57-64.	Wrong task
54	Blackman, J. D., Towle, V. L., Sturis, J., Lewis, G. F., Spire, J. P., & Polonsky, K. S. (1992). Hypoglycemic Thresholds for Cognitive Dysfunction in Iddm. Diabetes, 41(3), 392-399.	No IIV
55	Blagrove, M., Colemorgan, D., & Lambe, H. (1994). Interrogative Suggestibility - the Effects of Sleep-Deprivation and Relationship with Field-Dependence. Applied Cognitive Psychology, 8(2), 169-179.	No IIV

56	Blasi, G., Mattay, V. S., Bertolino, A., Elvevag, B., Callicott, J. H., Das, S., et al. (2005). Effect of catechol-O-methyltransferase val158met genotype on attentional control. Journal of Neuroscience, 25(20), 5038-5045.	No IIV
57	Bleecker, M. L., Lindgren, K. N., Tiburzi, M. J., & Ford, D. P. (1997). Curvilinear relationship between blood lead level and reaction time. Differential association with blood lead fractions derived from exogenous and endogenous sources. Journal of Occupational & Environmental Medicine, 39(5), 426-431.	No IIV
58	Bleom, B. R., Grimbergen, Y. A. M., van Dijk, J. G., & Munneke, M. (2006). The "posture second" strategy: A review of wrong priorities in Parkinson's disease. Journal of the Neurological Sciences, 248(1-2), 196-204.	No empirical data
59	Bliese, P. D., Wesensten, N. J., & Balkin, T. J. (2006). Age and individual variability in performance during sleep restriction. Journal of Sleep Research, 15(4), 376-385.	Wrong IIIV
60	Blinkov, S. M., & Nikandrov, M. G. (2002). On the Problem of Cooperation of the Reaction of Hands: I Cooperation of Hand Reaction to Sound and Light Stimuli in Healthy Subjects. Human Physiology, 28(6), 698-701.	No IIV
61	Bock, O., & Girgenrath, M. (2006). Relationship between sensorimotor adaptation and cognitive functions in younger and older subjects. Experimental Brain Research, 169(3), 400-406.	Wrong IIIV
62	Boessenkool, J. J., Nijhof, E. J., & Erkelens, C. J. (1999). Variability and correlations in bi-manual pointing movements. Human Movement Science, 18(4), 525-552.	Matter not relevant
63	Bora, E., Vahip, S., & Akdeniz, F. (2006). Sustained attention deficits in manic and euthymic patients with bipolar disorder. Progress in Neuro-Psychopharmacology & Biological Psychiatry, 30(6), 1097-1102.	Wrong task
64	Borah, D., Wadhwa, S., Singh, U., Yadav, S. L., Bhattacharjee, M., & V, S. (2007). Age related changes in postural stability. Indian Journal of Physiology and Pharmacology, 51(4), 395-404.	No IIV
65	Borello-France, D. F., Gallagher, J. D., Furman, J. M., Redfern, M. S., & Carvell, G. E. (2002). Voluntary upper-extremity movements in patients with unilateral peripheral vestibular hypofunction. Physical Therapy, 82(3), 216-227.	No RT
66	Borg, J., Andree, B., Lundberg, J., Halldin, C., & Farde, L. (2006). Search for correlations between serotonin 5-HT1A receptor expression and cognitive functions--a strategy in translational psychopharmacology. Psychopharmacology, 185(3), 389-394.	No IIV
67	Bosley, T. M., Salih, M. A., Alorainy, I. A., Oystreck, D. T., Nester, M., Abu-Amero, K. K., et al. (2007). Clinical characterization of the HOXA1 syndrome BSAS variant. Neurology, 69(12), 1245-1253.	Sample consisted of children
68	Boston, P. F., Dennis, M. S., Jagger, C., Jarman, M., & Lamers, C. (2001). Unequal distribution of cognitive deficits in vascular dementia--is this a valid criterion in the ICD-10? International Journal of Geriatric Psychiatry, 16(4), 422-426.	No RT
69	Bosworth, H. B., Warner Schaie, K., Willis, S. L., & Siegler, I. C. (1999). Age and distance to death in the Seattle Longitudinal Study. Research on Aging, 21(6), 723-738.	No RT
70	Bowler, R. M., Nakagawa, S., Drezgic, M., Roels, H. A., Park, R. M., Diamond, E., et al. (2007). Sequelae of fume exposure in confined space welding: A neurological and neuropsychological case series. NeuroToxicology, 28(2), 298-311.	No RT
71	Braun, D. L., Dulit, R. A., Adler, D. A., Berlant, J., Dixon, L., Fornari, V., et al. (2004). Attention-deficit/hyperactivity disorder in adults: Clinical information for primary care physicians. Primary Psychiatry, 11(9), 56-65.	No empirical data
72	Braun, D., Weber, H., Mergner, T., & Schultemonting, J. (1992). Saccadic Reaction-Times in Patients with Frontal and Parietal Lesions. Brain, 115, 1359-1386.	Wrong task
73	Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. Nature Reviews Neuroscience, 3(3), 243-249.	No empirical data
74	Brown, L. A., & Woollacott, M. H. (1998). The effects of aging on the control of posture and locomotion in healthy older adults: An emphasis on cognition. Psychologische Beitrage, 40(1), 27-43.	No empirical data
75	Brumaghim, J. T., Klorman, R., Strauss, J., Lewine, J. D., & et al. (1987). Does methylphenidate affect information processing? Findings from two studies on performance and P3b latency. Psychophysiology, 24(3), 361-373.	Wrong task
76	Brumback, C. R., Low, K. A., Gratton, G., & Fabiani, M. (2005). Putting things into perspective - Individual differences in working-memory span and the integration of information. Experimental Psychology, 52(1), 21-30.	No IIV
77	Buckwalter, J. G., Buckwalter, D. K., Bluestein, B. W., & Stanczyk, F. Z. (2001). Pregnancy and post partum: Changes in cognition and mood. Progress in Brain Research, 133, 303-319.	No empirical data
78	Burdick, K. E., Funke, B., Goldberg, J. F., Bates, J. A., Jaeger, J., Kucherlapati, R., et al. (2007). COMT genotype increases risk for bipolar I disorder and influences neurocognitive performance. Bipolar Disorders, 9(4), 370-376.	No IIV
79	Burdick, K. E., Goldberg, J. F., Harrow, M., Faull, R. N., & Malhotra, A. K. (2006). Neurocognition as a stable endophenotype in bipolar disorder and schizophrenia. Journal of Nervous & Mental Disease, 194(4), 255-260.	No RT
80	Burton, C. L., Hultsch, D. F., Strauss, E., & Hunter, M. A. (2002). Intraindividual variability in physical and emotional functioning: comparison of adults with traumatic brain injuries and healthy adults. Clinical Neuropsychologist, 16(3), 264-279.	No RT
81	Cacabelos, R. (2006). Homocysteine and cognition: From Galenic dogmatism to genetic relativism. Aging Health, 2(5), 783-786.	No RT

82	Cacabelos, R. (2008). Influence of pharmacogenetic factors on Alzheimer's disease therapeutics. Neurodegenerative Diseases, 5(3-4), 176-178.	No empirical data
83	Calil, H. M., Zwicker, A. P., & Klepacz, S. (1990). The effects of lithium carbonate on healthy volunteers: mood stabilization? Biological Psychiatry, 27(7), 711-722.	No IIV
84	Callaway, E. I. (1961). Day-to-day variability in relationship between electroencephalographic alpha phase and reaction time to visual stimuli. Annals of the New York Academy of Sciences, 92, 1183-1186.	No IIV
85	Cameron, O. G., Modell, J. G., Hichwa, R. D., Agranoff, B. W., & Koeppe, R. A. (1990). Changes in sensory-cognitive input: effects on cerebral blood flow. Journal of Cerebral Blood Flow & Metabolism, 10(1), 38-42.	No RT
86	Cannon, T. D., Huttunen, M. O., Lonnqvist, J., Tuulio-Henriksson, A., Pirkola, T., Glahn, D., et al. (2000). The inheritance of neuropsychological dysfunction in twins discordant for schizophrenia. American Journal of Human Genetics, 67(2), 369-382.	No IIV
87	Cantagallo, A., Maietti, A., Hall, K. M., & Bushnik, T. (2000). Comparison among functional outcome measures for traumatic brain injury: Assessment in the Italian community. Europa Medicophysica, 36(4), 171-182.	No RT
88	Cantin, J. F., McFadyen, B. J., Doyon, J., Swaine, B., Dumas, D., & Vallee, M. (2007). Can measures of cognitive function predict locomotor behaviour in complex environments following a traumatic brain injury? Brain Injury, 21(3), 327-334.	No RT
89	Capizzano, A. A., AciÃ³n, L., Bekinschtein, T., Furman, M., Gomila, H., MartÃnez, A., et al. (2004). White matter hyperintensities are significantly associated with cortical atrophy in Alzheimer's disease. Journal of Neurology, Neurosurgery & Psychiatry, 75(6), 822-827.	No RT
90	Carlsson, G. E., Moller, A., & Blomstrand, C. (2003). Consequences of mild stroke in persons <75 years - A 1-year follow-up. Cerebrovascular Diseases, 16(4), 383-388.	No RT
91	Carlton, M. J., Robertson, R. N., Carlton, L. G., & Newell, K. M. (1985). Response timing variability: Coherence of kinematic and EMG parameters. Journal of Motor Behavior, 17(3), 301-319.	Matter not relevant
92	Carnahan, H., Elliott, D., & Velamoor, V. R. (1995). Manual performance in leukotomized and unleukotomized individuals with schizophrenia. Schizophrenia Research, 17(3), 267-278.	No RT
93	Carnahan, H., Vandervoort, A. A., & Swanson, L. R. (1998). The influence of aging and target motion on the control of prehension. Experimental Aging Research, 24(3), 289-306.	No RT
94	Carpenter, R. H. S. (1999). A neural mechanism that randomises behaviour. Journal of Consciousness Studies, 6(1), 13-22.	No empirical data
95	Carr, V., & Wale, J. (1986). Schizophrenia: an information processing model. Australian & New Zealand Journal of Psychiatry, 20(2), 136-155.	No empirical data
96	Carron, A. V. (1971). Motor performance and response consistency as a function of age. Journal of Motor Behavior, 3(2), 105-109.	Sample consisted mainly of children
97	Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. Trends in Cognitive Sciences, 10(3), 117-123.	No empirical data
98	Cauraugh, J. H., & Kim, S. B. (2003). Stroke motor recovery: active neuromuscular stimulation and repetitive practice schedules. Journal of Neurology Neurosurgery and Psychiatry, 74(11), 1562-1566.	No IIV
99	Cervilla, J. A., Prince, M., Joels, S., Russ, C., & Lovestone, S. (2004). Genes Related to Vascular Disease (APOE, VLDL-R, DCP-1) and Other Vascular Factors in Late-Life Depression. American Journal of Geriatric Psychiatry, 12(2), 202-210.	No RT
100	Chan, R. C. K., Chen, R. Y. L., Chen, E. Y. H., Hui, T. C. K., Cheung, E. F. C., Cheung, H. K., et al. (2005). The differential clinical and neurocognitive profiles of COMT SNP rs165599 genotypes in schizophrenia. Journal of the International Neuropsychological Society, 11(2), 202-204.	No IIV
101	Chatterjea, R. G., & et al. (1979). Reaction time in schizophrenics and normals. Psychological Studies, 24(2), 79-81.	No IIV
102	Chei, T. T., & Demetrio, F. N. (2006). Memantine may acutely improve cognition and have a mood stabilizing effect in treatment-resistant bipolar disorder [2]. Revista Brasileira de Psiquiatria, 28(3), 252-254.	No RT
103	Cheung, S. S., Westwood, D. A., & Knox, M. K. (2007). Mild body cooling impairs attention via distraction from skin cooling. Ergonomics, 50(2), 275-288.	No RT
104	Chow, S.-M., Hamagani, F., & Nesselroade, J. R. (2007). Age differences in dynamical emotion-cognition linkages. Psychology and Aging, 22(4), 765-780.	Wrong task
105	Christensen, H. (2001). What cognitive changes can be expected with normal ageing? Australian & New Zealand Journal of Psychiatry, 35(6), 768-775.	No empirical data
106	Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., & Jacomb, P. (1999). Dispersion in cognitive ability as a function of age: A longitudinal study of an elderly community sample. Aging, Neuropsychology, and Cognition, 6(3), 214-228.	Wrong IIIV
107	Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., Jacomb, P., et al. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: individual differences in change scores as a function of age. Psychology & Aging, 14(3), 365-379.	No IIV
108	Christensen, H., Mackinnon, A., Jorm, A. F., Henderson, A. S., Scott, L. R., & Korten, A. E. (1994). Age-Differences and Interindividual Variation in Cognition in Community-Dwelling Elderly. Psychology and Aging, 9(3), 381-390.	No IIV
109	Christensen, H., Mackinnon, A., Jorm, A. F., Korten, A., Jacomb, P., Hofer, S. M., et al. (2004). The Canberra longitudinal study: Design, aims, methodology, outcomes and recent empirical investigations. Aging, Neuropsychology, and Cognition, 11(2), 169-195.	No IIV

110	Christina, R. W. (1992). The 1991 Mccloy,C.H. Research Lecture - Unraveling the Mystery of the Response Complexity Effect in Skilled Movements. Research Quarterly for Exercise and Sport, 63(3), 218-230.	No empirical data
111	Christrup, L. L., Sjogren, P., Jensen, N. H., Banning, A. M., Elbaek, K., & Ersboll, A. K. (1999). Steady-state kinetics and dynamics of morphine in cancer patients: is sedation related to the absorption rate of morphine? Journal of Pain & Symptom Management, 18(3), 164-173.	No IIV
112	Chudley, A. E., de von Flindt, R., & Hagerman, R. J. (1987). Cognitive variability in the fragile X syndrome. American Journal of Medical Genetics, 28(1), 13-15.	No empirical data
113	Cicerone, K. D. (1997). Clinical sensitivity of four measures of attention to mild traumatic brain injury. Clinical Neuropsychologist, 11(3), 266-272.	No IIV
114	Cicerone, K. D., Smith, L. C., Ellmo, W., Mangel, H. R., Nelson, P., Chase, R. F., et al. (1996). Neuropsychological rehabilitation of mild traumatic brain injury. Brain Injury, 10(4), 277-286.	No IIV
115	Clodore, M., Benoit, O., Foret, J., & Bouard, G. (1990). The Multiple Sleep Latency Test: individual variability and time of day effect in normal young adults. Sleep, 13(5), 385-394.	No RT
116	Coffin, S., & Ganz, L. (1977). Perceptual correlates of variability in the duration of the cortical excitability cycle. Neuropsychologia, 15(2), 231-241.	Wrong task
117	Cohen, R. A., Barnes, H. J., Jenkins, M., & Albers, H. E. (1997). Disruption of short-duration timing associated with damage to the suprachiasmatic region of the hypothalamus. Neurology, 48(6), 1533-1539.	No RT
118	Cohen, R., & Rist, F. (Eds.). (1992). The Modality Shift Effect: Further explorations at the crossroads. New York, NY,US: New York Academy of Sciences.	No IIV
119	Coimbra, E. R., Rezek, K., Escorsi-Rosset, S., Landemberger, M. C., Castro, R. M. R. P. S., Valadao, M. N., et al. (2006). Cognitive performance of patients with mesial temporal lobe epilepsy is not associated with human prion protein gene variant allele at codons 129 and 171. Epilepsy and Behavior, 8(3), 635-642.	No RT
120	Collerton, D., Burn, D., McKeith, I., & O'Brien, J. (2003). Systematic review and meta-analysis show that dementia with Lewy bodies is a visual-perceptual and attentional-executive dementia. Dementia and Geriatric Cognitive Disorders, 16(4), 229-237.	No empirical data
121	Colosio, C., Tiramani, M., & Maroni, M. (2003). Neurobehavioral effects of pesticides: state of the art. Neurotoxicology, 24(4-5), 577-591.	No empirical data
122	Connor, P. D., Streissguth, A. P., Sampson, P. D., Bookstein, F. L., & Barr, H. M. (1999). Individual differences in auditory and visual attention among fetal alcohol-affected adults. Alcoholism-Clinical and Experimental Research, 23(8), 1395-1402.	Wrong task
123	Conte, S., Ferlazzo, F., & D'Olimpio, F. (1995). Rhythmic variations in reaction times and evoked potentials. International Journal of Neuroscience, 82(1), 127-133.	No IIV
124	Contin, M., Provini, F., Martinelli, P., Riva, R., Albani, F., Vetrugno, R., et al. (2003). Excessive daytime sleepiness and levodopa in Parkinson's disease: Polygraphic, placebo-controlled monitoring. Clinical Neuropharmacology, 26(3), 115-118.	No IIV
125	Contrerasvidal, J. L., & Stelmach, G. E. (1995). Effects of Parkinsonism on Motor Control. Life Sciences, 58(3), 165-176.	No empirical data
126	Conway Greig, T., Nicholls, S. S., Wexler, B. E., & Bell, M. D. (2004). Test-retest stability of neuropsychological testing and individual differences in variability in schizophrenia outpatients. Psychiatry Research, 129(3), 241-247.	No RT
127	Cook, M. R., Graham, C., Sastre, A., & Gerkovich, M. M. (2002). Physiological and performance effects of pyridostigmine bromide in healthy volunteers: A dose-response study. Psychopharmacology, 162(2), 186- 192.	No IIV
128	Cools, R., & Robbins, T. W. (2004). Chemistry of the adaptive mind. Philosophical Transactions of the Royal Society London, Series a, 362(1825), 2871-2888.	No empirical data
129	Cooney, J. B., & Troyer, R. (1994). A dynamic model of reaction time in a short-term memory task. Journal of Experimental Child Psychology, 58(2), 200-226.	Wrong task
130	Coppens, P. (1995). Subpopulations in closed-head injury: Preliminary results. Brain Injury, 9(2), 195-208.	No RT
131	Coppens, P., & Hungerford, S. (2001). Crossed aphasia: Two new cases. Aphasiology, 15(9), 827-854.	No RT
132	Correa, A., Lupianez, J., & Tudela, P. (2006). The attentional mechanism of temporal orienting: determinants and attributes. Experimental Brain Research, 169(1), 58-68.	Wrong task
133	Corsi-Cabrera, M., Arce, C., del Rio-Portilla, I. Y., PÃ©rez-Garci, E., & Guevara, M. A. (1999). Amplitude reduction in visual event-related potentials as a function of sleep deprivation. Sleep: Journal of Sleep Research & Sleep Medicine, 22(2), 181-189.	Wrong task
134	Costa, L. D., Horwitz, M., & Vaughan, H. G., Jr. (1965). Choice reaction time as a function of stimulus uncertainty in patients with brain lesions. Perceptual and Motor Skills, 21(3), 885-886.	Wrong task
135	Cousineau, D. (Ed.). (2004). Fast fitting of convolutions using rational approximations. Mahwah, NJ,US: Lawrence Erlbaum Associates Publishers.	No participants - simulated data
136	Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. Intelligence, 31(6), 567-587.	No empirical data
137	Cramer, D. (1977). Reaction time as a function of arousal. Irish Journal of Psychology, 3(3), 206-214.	No IIV
138	Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B.-K., et al. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence.[see comment]. Brain, 126(Pt 10), 2139-2152.	No RT

139	Crockett, D., Clark, C., Spreen, O., & Klonoff, H. (1981). Severity of impairment or specific types of aphasia: An empirical investigation. Cortex, 17(1), 83-95.	No RT
140	Cronister, A., Schreiner, R., Wittenberger, M., Amiri, K., Harris, K., & Hagerman, R. J. (1991). Heterozygous fragile X female: historical, physical, cognitive, and cytogenetic features. American Journal of Medical Genetics, 38(2-3), 269-274.	No RT
141	Crystal, H. A., Schaumburg, H. H., Grober, E., Fuld, P. A., & Lipton, R. B. (1988). Cognitive impairment and sensory loss associated with chronic low-level ethylene oxide exposure. Neurology, 38(4), 567-569.	Wrong task
142	Cudmore, L. J., & Segalowitz, S. J. (2000). Signal-to-noise ratio sensitivity in ERPs to stimulus and task complexity: Different effects for early and late components. Brain and Cognition, 43(1), 130-134.	Wrong task
143	Cuesta, M. J., Peralta, V., & Zarzuela, A. (2007). Empirical validation of competing definitions of schizophrenia: a poly-diagnostic study of cognitive impairment in non-affective psychosis. Schizophrenia Research, 95(1-3), 39-47.	No RT
144	Cummings, J. L. (2000). Cognitive and behavioral heterogeneity in Alzheimer's disease: seeking the neurobiological basis.[see comment]. Neurobiology of Aging, 21(6), 845-861.	No empirical data
145	Cummings, J. L., Schneider, E., Tariot, P. N., & Graham, S. M. (2006). Behavioral effects of memantine in Alzheimer disease patients receiving donepezil treatment. Neurology, 67(1), 57-63.	No RT
146	Curley, M. D., & Butler, F. K., Jr. (1987). Visual reaction time performance preceding CNS oxygen toxicity. Undersea Biomedical Research, 14(4), 301-310.	No IIV
147	Curtin, J. J., Patrick, C. J., Lang, A. R., Cacioppo, J. T., & Birbaumer, N. (2001). Alcohol affects emotion through cognition. Psychological Science, 12(6), 527-531.	Wrong task
148	Dagtekin, O., Gerbershagen, H. J., Wagner, W., Petzke, F., Radbruch, L., & Sabatowski, R. (2007). Assessing cognitive and psychomotor performance under long-term treatment with transdermal buprenorphine in chronic noncancer pain patients. Anesthesia & Analgesia, 105(5), 1442-1448.	No IIV
149	Dalbokova, D., & Kolev, P. (1994). Mental capacity changes in school teachers. Reviews on Environmental Health, 10(1), 51-56.	Wrong task
150	Daniele, A., Albanese, A., Contarino, M. F., Zinzi, P., Barbier, A., Gasparini, F., et al. (2003). Cognitive and behavioural effects of chronic stimulation of the subthalamic nucleus in patients with Parkinson's disease. Journal of Neurology, Neurosurgery & Psychiatry, 74(2), 175-182.	No RT
151	Day, B. L., Thompson, P. D., Harding, A. E., & Marsden, C. D. (1998). Influence of vision on upper limb reaching movements in patients with cerebellar ataxia. Brain: A Journal of Neurology, 121(2), 357-372.	No IIV
152	De Geus, E. J. C., Wright, M. J., Martin, N. G., & Boomsma, D. I. (2001). Genetics of brain function and cognition. Behavior Genetics, 31(6), 489-495.	No empirical data
153	De Luca, V., Vincent, J. B., Muller, D. J., Hwang, R., Shinkai, T., Volavka, J., et al. (2005). Identification of a naturally occurring 21 bp deletion in alpha2c noradrenergic receptor gene and cognitive correlates to antipsychotic treatment. Pharmacological Research, 51(4), 381-384.	No RT
154	de Vries, P. J., & Howe, C. J. (2007). The tuberous sclerosis complex proteins - a GRIPP on cognition and neurodevelopment. Trends in Molecular Medicine, 13(8), 319-326.	No empirical data
155	Deary, I. J. (2003). Ten things I hated about intelligence research. The Psychologist, 16(10), 534-537.	No empirical data
156	Deary, I. J., Whiteman, M. C., Pattie, A., Starr, J. M., Hayward, C., Wright, A. F., et al. (2004). Apolipoprotein e gene variability and cognitive functions at age 79: a follow-up of the Scottish mental survey of 1932. Psychology & Aging, 19(2), 367-371.	No RT
157	Deary, I. J., Wright, A. F., Harris, S. E., Whalley, L. J., & Starr, J. M. (2004). Searching for genetic influences on normal cognitive ageing. Trends in Cognitive Sciences, 8(4), 178-184.	No empirical data
158	Decary, A., Vendette, M. l., Massicotte-Marquez, J., & Mathieu, A. (2005). A Preliminary Study of the Impact of Age-related Pathologies on Cognitive Functioning and Waking EEG. North American Journal of Psychology, 7(3), 469-480.	No RT
159	Decety, J. (1993). Analysis of actual and mental movement times in graphic tasks. Acta Psychologica, 82(1-3), 367-372.	Wrong task
160	Deixelberger-Fritz, D., Tischler, M. A., & Kallus, K. W. (2003). Changes in Performance, Mood State and Workload Due to Energy Drinks in Pilots. International Journal of Applied Aviation Studies, 3(2), 195-205.	No IIV
161	Dembowski, J., & Watson, B. C. (1991). Preparation time and response complexity effects on stutterers' and nonstutterers' acoustic LRT. Journal of Speech & Hearing Research, 34(1), 49-59.	No IIV
162	Demiralp, T., Ademoglu, A., Schurmann, M., Basar-Eroglu, C., & Basar, E. (1999). Detection of P300 waves in single trials by the wavelet transform (WT). Brain and Language, 66(1), 108-128.	Wrong task
163	Depp, C. A., Glatt, S. J., & Jeste, D. V. (2007). Recent advances in research on successful or heatlhy aging. Current Psychiatry Reports, 9(1), 7-13.	No empirical data
164	Derakhshan, I. (2006). Nonconvulsive status epilepticus with an unusual EEG: A fresh look at lateralities of motor control and awareness. Epilepsy & Behavior, 9(1), 204-210.	No RT
165	Desmurget, M., Gaveau, V., Vindras, P., Turner, R. S., Broussolle, E., & Thobois, S. (2004). On-line motor control in patients with Parkinson's disease. Brain, 127(Pt 8), 1755-1773.	No IIV
166	D'Esposito, M., Zarahn, E., Aguirre, G. K., & Rypma, B. (1999). The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. Neuroimage, 10(1), 6-14.	No IIV
167	Despres, C., Lamoureux, D., & Beuter, A. (2000). Standardization of a neuromotor test battery: the CATSYS system. Neurotoxicology, 21(5), 725-735.	No IIV
168	Devinsky, O. (1995). Cognitive and behavioral effects of antiepileptic drugs. Epilepsia, 36 Suppl 2, S46-65.	No empirical data

169	Devinsky, O., & Vazquez, B. (1993). Behavioral changes associated with epilepsy. Neurologic Clinics, 11(1), 127-149.	No empirical data
170	Diefenbach, K., Donath, F., Maurer, A., Bravo, S. Q., Wernecke, K. D., Schwantes, U., et al. (2003). Randomised, double-blind study of the effects of oxybutynin, tolterodine, trospium chloride and placebo on sleep in healthy young volunteers. Clinical Drug Investigation, 23(6), 395-404.	Wrong task
171	Diekmann, V., Reinke, W., & Grozinger, B. (1985). Diminished order in the EEG of schizophrenic patients. Naturwissenschaften, 72(10), 541-542.	No RT
172	Dise, J. E., & Lohr, M. E. (1998). Examination of deficits in conceptual reasoning abilities associated with spina bifida. American Journal of Physical Medicine & Rehabilitation, 77(3), 247-251.	No RT
173	Dixon, R. A. (Ed.). (2003). Themes in the aging of intelligence: Robust decline with intriguing possibilities. Washington, DC,US: American Psychological Association.	No empirical data
174	Dockree, P. M., Bellgrove, M. A., O'Keeffe, F. M., Moloney, P., Aimola, L., Carton, S., et al. (2006). Sustained attention in traumatic brain injury (TBI) and healthy controls: enhanced sensitivity with dual-task load. Experimental Brain Research, 168(1-2), 218-229.	No RT
175	Dohrn, C. S., Lichtor, J. L., Finn, R. S., Uitvlugt, A., Coalson, D. W., Rupani, G., et al. (1992). Subjective and Psychomotor Effects of Nitrous-Oxide in Healthy-Volunteers. Behavioural Pharmacology, 3(1), 19-30.	No IIV
176	Drane, D. L., Williamson, D. J., Stroup, E. S., Holmes, M. D., Jung, M., Koerner, E., et al. (2006). Cognitive impairment is not equal in patients with epileptic and psychogenic nonepileptic seizures. Epilepsia, 47(11), 1879-1886.	No RT
177	Duarte, M., & Latash, M. L. (2007). Effects of postural task requirements on the speed-accuracy trade-off. Experimental Brain Research, 180(3), 457-467.	Wrong task
178	Dubal, S. p., & Jouvent, R. (2004). Time-on-task effect in trait anhedonia. European Psychiatry, 19(5), 285-291.	Wrong task
179	Duclos, Y., Schmied, A., Burle, B., Burnet, H., & Rossi-Durand, C. (2008). Anticipatory changes in human motoneuron discharge patterns during motor preparation. Journal of Physiology-London, 586(4), 1017-1028.	Wrong task
180	Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. Biological Psychology, 14(1), 1-52.	Wrong task
181	Dunham, M. D., & Johnstone, B. (1999). Variability of neuropsychological deficits associated with carbon monoxide poisoning: four case reports. Brain Injury, 13(11), 917-925.	No RT
182	Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. Seminars in Neurology, 25(1), 117-129.	No empirical data
183	Dzierzewski, J., McCrae, C., Marsiske, M., McCoy, K., McNamara, J., Dautovich, N., et al. (2007). Daily variability in objective and subjective sleep and cognition in older adults with insomnia. Sleep, 30, A112-A112.	No RT
184	Echeverria, D., Heyer, N. J., Martin, M. D., Naleway, C. A., Woods, J. S., & Bittner, A. C., Jr. (1995). Behavioral effects of low-level exposure to elemental Hg among dentists. Neurotoxicology & Teratology, 17(2), 161-168.	No IIV
185	Egan, M. F., Straub, R. E., Goldberg, T. E., Yakub, I., Callicott, J. H., Hariri, A. R., et al. (2004). Variation in GRM3 affects cognition, prefrontal glutamate, and risk for schizophrenia. Proceedings of the National Academy of Sciences of the United States of America, 101(34), 12604-12609.	No IIV
186	Egner, T., & Gruzelier, J. H. (2004). EEG Biofeedback of low beta band components: Frequency-specific effects on variables of attention and event-related brain potentials. Clinical Neurophysiology, 115(1), 131-139.	Wrong task
187	Elbel, E. R. (1939). A study in variation in response time. Research Quarterly of the American Association for Health, Physical Education, & Recreation, 10, 35-50.	No IIV
188	Elbert, J. C., & Seale, T. W. (1988). Complexity of the cognitive phenotype of an inherited form of learning disability. Developmental Medicine and Child Neurology, 30(2), 181-189.	No RT
189	Elliott, D., Roy, E. A., Goodman, D., Carson, R. G., Chua, R., & Maraj, B. K. V. (1993). Asymmetries in the Preparation and Control of Manual Aiming Movements. Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 47(3), 570-589.	Matter not relevant
190	Elsass, P., & Henriksen, L. (1984). Acute cerebral dysfunction after open-heart surgery. A reaction-time study. Scandinavian Journal of Thoracic & Cardiovascular Surgery, 18(2), 161-165.	No IIV
191	Elsesser, K., Sartory, G., & Tackenberg, A. (2005). Initial symptoms and reactions to trauma-related stimuli and the development of posttraumatic stress disorder. Depression and Anxiety, 21(2), 61-70.	No IIV
192	Englund, C. E., Ryman, D. H., Naitoh, P., & Hodgdon, J. A. (1985). Cognitive performance during successive sustained physical work episodes. Behavior Research Methods, Instruments & Computers, 17(1), 75-85.	No IIV
193	Engman, E., Andersson-Roswall, L., & Malmgren, K. (2001). Pre--and postoperative general neurocognitive status and memory in 70 epilepsy surgery patients. Acta Neurologica Scandinavica, 103(6), 351-359.	No RT
194	Erker, G. J., Searight, H. R., & Peterson, P. (1995). Patterns of neuropsychological functioning among patients with multi-infarct and Alzheimer's dementia: A comparative analysis. International Psychogeriatrics, 7(3), 393-406.	No RT
195	Eston, R. G. (1984). The regular menstrual cycle and athletic performance. Sports Medicine, 1(6), 431-445.	No empirical data
196	Evins, A. E., Deckersbach, T., Cather, C., Freudenreich, O., Culhane, M. A., Henderson, D. C., et al. (2005). Independent Effects of Tobacco Abstinence and Bupropion on Cognitive Function in Schizophrenia. Journal of Clinical Psychiatry, 66(9), 1184-1190.	Wrong task

197	Extermann, M., Chen, H., Booth-Jones, M., Meyer, J., Balducci, L., & Jacobsen, P. (2005). Pilot testing of the computerized cognitive test Microcog in chemotherapy-treated older cancer patients. Critical Reviews in Oncology-Hematology, 54(2), 137-143.	No IIV
198	Fabbrini, M., Frittelli, C., Bonanni, E., Maestri, M., Manca, M. L., & Iudice, A. (2005). Psychomotor performance in healthy young adult volunteers receiving lormetazepam and placebo: a single-dose, randomized, double-blind, crossover trial. Clinical Therapeutics, 27(1), 78-83.	No IIV
199	Fallgatter, A. J., Herrmann, M. J., Hohoff, C., Ehlis, A. C., Jarczok, T. A., Freitag, C. M., et al. (2006). DTNBP1 (dysbindin) gene variants modulate prefrontal brain function in healthy individuals. Neuropsychopharmacology, 31(9), 2002-2010.	Wrong task
200	Fan, J., Fossella, J., Sommer, T., Wu, Y., & Posner, M. I. (2003). Mapping the genetic variation of executive attention onto brain activity. Proceedings of the National Academy of Sciences of the United States of America, 100(12), 7406-7411.	Wrong task
201	Fan, J., Wu, Y., Fossella, J. A., & Posner, M. I. (2001). Assessing the heritability of attentional networks. BMC Neuroscience, 2(14).	Wrong task
202	Fargo, J. D., Schefft, B. K., Szaflarski, J. P., Dulay, M. F., Testa, S. M., Privitera, M. D., et al. (2004). Accuracy of self-reported neuropsychological functioning in individuals with epileptic or psychogenic nonepileptic seizures. Epilepsy and Behavior, 5(2), 143-150.	No RT
203	Ferrand, C. T., Gilbert, H. R., & Blood, G. W. (1991). Selected Aspects of Central Processing and Vocal Motor Function in Stutterers and Nonstutterers - P300, Laryngeal Shift, and Vibratory Onset. Journal of Fluency Disorders, 16(2-3), 101-115.	Wrong task
204	Ferrand, C., Blood, G. W., & Gilbert, H. R. (1991). A Continuous-Flow Model for Phonatory Reaction-Time. Journal of Speech and Hearing Research, 34(3), 517-525.	Wrong task
205	Ferraro, F. R., Wonderlich, S., & Jocic, Z. (1997). Performance variability as a new theoretical mechanism regarding eating disorders and cognitive processing. Journal of Clinical Psychology, 53(2), 117-121.	No RT
206	Fiedorowicz, J. G., Moser, D. J., Hynes, S. M., Beglinger, L. J., Schultz, S. K., & Ellingrod, V. L. (2007). LA allelic heterozygosity of the 5HTTLPR polymorphism is associated with higher cognitive function and lower interpersonal sensitivity [2]. Psychiatric Genetics, 17(1), 3-4.	No RT
207	Fields, S. D., MacKenzie, C. R., Charlson, M. E., & Sax, F. L. (1986). Cognitive impairment. Can it predict the course of hospitalized patients? Journal of the American Geriatrics Society, 34(8), 579-585.	No RT
208	Finkel, D., & Pedersen, N. L. (2004). Processing speed and longitudinal trajectories of change for cognitive abilities: The Swedish Adoption/Twin Study of Aging. Aging, Neuropsychology, and Cognition, 11(2-3), 325-345.	No RT
209	Finkel, D., Pedersen, N. L., & Harris, J. R. (2000). Genetic mediation of the associations among motor and perceptual speed and adult cognitive abilities. Aging, Neuropsychology, and Cognition, 7(3), 141-155.	No RT
210	Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2005). The longitudinal relationship between processing speed and cognitive ability: Genetic and environmental influences. Behavior Genetics, 35(5), 535-549.	No RT
211	Finkelstein, Y., Milatovic, D., & Aschner, M. (2007). Modulation of cholinergic systems by manganese. Neurotoxicology, 28(5), 1003-1014.	No empirical data
212	Fisher, B. E., Winstein, C. J., & Velicki, M. R. (2000). Deficits in compensatory trajectory adjustments after unilateral sensorimotor stroke.[erratum appears in Exp Brain Res 2000 Jun;132(3):417]. Experimental Brain Research, 132(3), 328-344.	No RT
213	Fisk, A. D., Rogers, W. A., & Giambra, L. M. (1990). Consistent and varied memory/visual search: Is there an interaction between age and response-set effects? Journals of Gerontology, 45(3), P81-PP87.	Wrong task
214	Fjell, A. M., & Walhovd, K. B. (2003). Effects of auditory stimulus intensity and hearing threshold on the relationship among P300, age, and cognitive function. Clinical Neurophysiology, 114(5), 799-807.	No RT
215	Fjell, A. M., & Walhovd, K. B. (2007). Stability of brain potentials, mental abilities, and cortical thickness. Neuroreport, 18(8), 725-728.	Wrong task
216	Fjell, A. M., Ostby, Y., & Walhovd, K. B. (2007). Intraindividual variability in cognitive performance and its neural foundation--An important approach in studying the developing and aging mind. In M. A. Lange, (Ed). Leading-edge psychological tests and testing research. (pp. 193-200). xii, NY, US: Nova Science Publishers.	No empirical data
217	Flehmig, H. C., Steinborn, M., Langner, R., Scholz, A., & Westhoff, K. (2007). Assessing intraindividual variability in sustained attention: Reliability, relation to speed and accuracy, and practice effects. Psychology Science, 49(2), pp.	Wrong task
218	Fleischman, D. A., Wilson, R. S., Bienias, J. L., & Bennett, D. A. (2005). Parkinsonian signs and cognitive function in old age. Journal of the International Neuropsychological Society, 11(5), 591-597.	No RT
219	Flint, R. W., Jr., & Turek, C. (2003). Glucose effects on a continuous performance test of attention in adults. Behavioural Brain Research, 142(1), 217-228.	Wrong task
220	Floras, J. S., Hassan, M. O., Jones, J. V., Osikowska, B. A., Sever, P. S., & Sleight, P. (1988). Consequences of impaired arterial baroreflexes in essential hypertension: effects on pressor responses, plasma noradrenaline and blood pressure variability. Journal of Hypertension, 6(7), 525-535.	No IIV
221	Flouris, A. D., Westwood, D. A., & Cheung, S. S. (2007). Thermal balance effects on vigilance during 2-hour exposures to -20 degrees C. Aviation, Space, and Environmental Medicine, 78(7), pp.	Wrong task
222	Ford, J. M., & Pfefferbaum, A. (1991). Event-Related Potentials and Eyeblink Responses in Automatic and Controlled Processing - Effects of Age. Electroencephalography and Clinical Neurophysiology, 78(5), 361-377.	Wrong task

223	Fossella, J., Sommer, T., Fan, J., Wu, Y., Swanson, J. M., Pfaff, D. W., et al. (2002). Assessing the molecular genetics of attention networks. BMC Neuroscience, 3(04).	No IIV
224	Foster, J. K., Behrmann, M., & Stuss, D. T. (1995). Aging and visual search: Generalized cognitive slowing or selective deficit in attention? Aging & Cognition, 2(4), 279-299.	Wrong task
225	Frank, M. J., Santamaria, A., O'Reilly, R. C., & Willcutt, E. (2007). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. Neuropsychopharmacology, 32(7), 1583-1599.	Wrong task
226	Frank, M. J., Scheres, A., & Sherman, S. J. (2007). Understanding decision-making deficits in neurological conditions: insights from models of natural action selection. Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 362(1485), 1641-1654.	No empirical data
227	Frey, D. J., Badia, P., & Wright, K. P., Jr. (2004). Inter- and intra-individual variability in performance near the circadian nadir during sleep deprivation.[see comment]. Journal of Sleep Research, 13(4), 305-315.	Wrong IIIV
228	Friedrich, D., & Hawkins, W. F. (1975). Response-stimulus interval performance of moderately retarded institutionalized subjects. American Journal of Mental Deficiency, 80(3), 281-285.	No IIV
229	Frisoni, G. B., Gozzetti, A., Bignamini, V., Vellas, B. J., Berger, A. K., Bianchetti, A., et al. (1998). Special care units for dementia in nursing homes: A controlled study of effectiveness. Archives of Gerontology and Geriatrics, 215-224.	No RT
230	Fuller, R., & Jahanshahi, M. (1999). Impairment of willed actions and use of advance information for movement preparation in schizophrenia. Journal of Neurology, Neurosurgery & Psychiatry, 66(4), 502-509.	No IIV
231	Gaal, Z. A., Csuhaj, R., & Molnar, M. (2007). Age-dependent changes of auditory evoked potentials - Effect of task difficulty. Biological Psychology, 76(3), 196-208.	Wrong task
232	Gabriel, D., Veuillet, E., Bonnet, C., Vesson, J. F., & Collet, L. (2006). Simple reaction times in subjects with steeply sloping hearing loss: Is there an alteration at the edge of the loss? International Journal of Audiology, 45(8), 454-462.	No IIV
233	Gadea, M., Martinez-Bisbal, M. C., Marti-Bonmati, L., Espert, R., Casanova, B., Coret, F., et al. (2004). Spectroscopic axonal damage of the right locus coeruleus relates to selective attention impairment in early stage relapsing-remitting multiple sclerosis. Brain, 127(Pt 1), 89-98.	No RT
234	Galliaudi, E., Taillard, J., Sagaspei, P., Valtati, C., Bioulac, B., & Philip, P. (2008). Sharp and sleepy: evidence for dissociation between sleep pressure and nocturnal performance. Journal of Sleep Research, 17(1), 11-15.	No IIV
235	Garcia, C., Philippot, P., Jouve, E., Bruguerolle, B., & Blin, O. (1998). Effects of anxiety on emotional reactivity, performance and vigilance, in healthy volunteers. Therapie, 53(3), 301-308.	No IIV
236	Gates, D. H., & Dingwell, J. B. (2008). The effects of neuromuscular fatigue on task performance during repetitive goal-directed movements. Experimental Brain Research, 187(4), 573-585.	No RT
237	Georgopoulos, A. P., Kalaska, J. F., & Massey, J. T. (1981). Spatial trajectories and reaction times of aimed movements: Effects of practice, uncertainty, and change in target location. Journal of Neurophysiology, 46(4), 725-743.	No IIV
238	Gerson, A. D., Parra, L. C., & Sajda, P. (2005). Cortical origins of response time variability during rapid discrimination of visual objects. Neuroimage, 28(2), 342-353.	Wrong task
239	Gerstorf, D., Smith, J., & Baltes, P. B. (2006). A systemic-wholistic approach to differential aging: longitudinal findings from the Berlin Aging Study. Psychology & Aging, 21(4), 645-663.	No RT
240	Ghajar, J., & Ivry, R. B. (2008). The predictive brain state: Timing deficiency in traumatic brain injury? Neurorehabilitation and Neural Repair, 22(3), 217-227.	No empirical data
241	Ghez, C., Gordon, J., & Ghilardi, M. F. (1993). Programming of extent and direction in human reaching movements. Biomedical Research, 14(SUPPL. 1), 1-5.	No empirical data
242	Ghoneim, M. M., Hinrichs, J. V., O'Hara, M. W., Mehta, M. P., Pathak, D., Kumar, V., et al. (1988). Comparison of psychologic and cognitive functions after general or regional anesthesia. Anesthesiology, 69(4), 507-515.	No IIV
243	Gibbon, J., & Rutschmann, R. (1969). Temporal order judgement and reaction time. Science, 165(891), 413-415.	Matter not relevant
244	Gibbons, H., & Stahl, J. (2007). Response-time corrected averaging of event-related potentials. Clinical Neurophysiology, 118(1), 197-208.	Wrong task
245	Gilden, D. L. (1997). Fluctuations in the time required for elementary decisions. Psychological Science, 8(4), 296-301.	Wrong task
246	Gillberg, M., Kecklund, G., & Akerstedt, T. (1996). Sleepiness and performance of professional drivers in a truck simulator - Comparisons between day and night driving. Journal of Sleep Research, 5(1), 12-15.	No IIV
247	Giordani, B., Berent, S., Boivin, M. J., Penney, J. B., Lehtinen, S., Markel, D. S., et al. (1995). Longitudinal neuropsychological and genetic linkage analysis of persons at risk for Huntington's disease. Archives of Neurology, 52(1), 59-64.	No IIV
248	Giubilei, F., Bastianello, S., Paolillo, A., Gasperini, C., Tisei, P., Casini, A. R., et al. (1997). Quantitative magnetic resonance analysis in vascular dementia. Journal of Neurology, 244(4), 246-251.	No IIV
249	Glanz, B. I., Slonim, D., Urowitz, M. B., Gladman, D. D., Gough, J., & MacKinnon, A. (1997). Pattern of neuropsychologic dysfunction in inactive systemic lupus erythematosus. Neuropsychiatry, Neuropsychology, & Behavioral Neurology, 10(4), 232-238.	No RT
250	Glass, J. M. (2007). Visual function and cognitive aging: Differential role of contrast sensitivity in verbal versus spatial tasks. Psychology and Aging, 22(2), 233-238.	No RT

251	Glazebrook, C. M., Elliott, D., & Lyons, J. (2006). A kinematic analysis of how young adults with and without autism plan and control goal-directed movements. Motor Control, 10(3), 244-264.	No IIV
252	Glencross, D. J., & Koreman, M. M. (1979). The processing of proprioceptive signals. Neuropsychologia, 17(6), 683-687.	No IIV
253	Glisky, E. L. (Ed.). (2007). Changes in cognitive function in human aging. Boca Raton, FL,US: CRC Press.	No empirical data
254	Goldman, N., Turra, C. M., Glei, D. A., Lin, Y.-H., & Weinstein, M. (2006). Physiological dysregulation and changes in health in an older population. Experimental Gerontology, 41(9), 862-870.	No RT
255	Goldstein, I. B., Shapiro, D., La Rue, A., & Guthrie, D. (1998). Relationship between 24-hour ambulatory blood pressure and cognitive function in healthy elderly people. Aging, Neuropsychology, and Cognition, 5(3), 215-224.	No RT
256	Goodin, D. S., & Aminoff, M. J. (1995). Nonrandom Variability in a Reaction-Time Time-Series. Neurology, 45(4), A212-A212.	Wrong task
257	Gorman, M. F., & Fisher, D. L. (1998). Visual search tasks: Slowing of strategic and nonstrategic processes in the nonlexical domain. The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 53(3), P189-PP200.	Wrong task
258	Gourovitch, M. L., Torrey, E. F., Gold, J. M., Randolph, C., Weinberger, D. R., & Goldberg, T. E. (1999). Neuropsychological performance of monozygotic twins discordant for bipolar disorder. Biological Psychiatry, 45(5), 639-646.	No IIV
259	Grace, J., Amick, M. M., D'Abreu, A., Festa, E. K., Heindel, W. C., & Ott, B. R. (2005). Neuropsychological deficits associated with driving performance in Parkinson's and Alzheimer's disease. Journal of the International Neuropsychological Society, 11(6), 766-775.	No RT
260	Grea, H., Desmurget, M., & Prablanc, C. (2000). Postural invariance in three-dimensional reaching and grasping movements. Experimental Brain Research, 134(2), 155-162.	No RT
261	Greenwood, P. M., Fossella, J. A., & Parasuraman, R. (2005). Specificity of the effect of a nicotinic receptor polymorphism on individual differences in visuospatial attention. Journal of Cognitive Neuroscience, 17(10), 1611-1620.	Wrong task
262	Gribova, A., Donchin, O., Bergman, H., Vaadia, E., & de Oliveira, S. C. (2002). Timing of bimanual movements in human and non-human primates in relation to neuronal activity in primary motor cortex and supplementary motor area. Experimental Brain Research, 146(3), 322-335.	Matter not relevant
263	Gron, G., Friess, E., Herpers, M., & Rupprecht, R. (1997). Assessment of cognitive performance after progesterone administration in healthy male volunteers. Neuropsychobiology, 35(3), 147-151.	Wrong task
264	Grunberger, J., & et al. (1978). Pharmacodynamic investigations with WE 941, a new triazolodiazepine, by means of psychometric analyses. Current Therapeutic Research, 24(4), 427-440.	No IIV
265	Grunberger, J., Saletu, B., Linzmayer, L., & Stohr, H. (1982). Objective measures in determining the central effectiveness of a new antihypoxidotic SL-76188: Pharmaco-EEG, psychometric and pharmacokinetic analyses in the elderly. Archives of Gerontology and Geriatrics, 1(3), 261-285.	No IIV
266	Guerin, F., Belleville, S., & Ska, B. (2002). Characterization of visuoconstructional disabilities in patients with probable dementia of Alzheimer's type. Journal of Clinical & Experimental Neuropsychology: Official Journal of the International Neuropsychological Society, 24(1), 1-17.	No RT
267	Gunal, D. I., Nurichalichi, K., Tuncer, N., Bekiroglu, N., & Aktan, S. (2000). The clinical profile of nonmotor fluctuations in Parkinson's disease patients. Marmara Medical Journal, 13(3), 148-152.	No RT
268	Gunstad, J., Paul, R. H., Brickman, A. M., Cohen, R. A., Arns, M., Roe, D., et al. (2006). Patterns of cognitive performance in middle-aged and older adults: A cluster analytic examination. Journal of Geriatric Psychiatry & Neurology, 19(2), 59-64.	No RT
269	Gupta, A., Lind, S., Eklund, A., & Lennmarken, C. (1997). The effects of midazolam and flumazenil on psychomotor function. Journal of Clinical Anesthesia, 9(1), 21-25.	No IIV
270	Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., et al. (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. Neuropsychopharmacology, 25(5), 766-776.	Wrong task
271	Hagberg, B., Bauer Alfredson, B., Poon, L. W., & Homma, A. (2001). Cognitive functioning in centenarians: a coordinated analysis of results from three countries. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 56(3), P141-151.	No IIV
272	Hahn, B., Ross, T. J., & Stein, E. A. (2007). Cingulate activation increases dynamically with response speed under stimulus unpredictability. Cerebral Cortex, 17(7), 1664-1671	Wrong task
273	Hahn, B., Ross, T. J., Yang, Y., Kim, I., Huestis, M. A., & Stein, E. A. (2007). Nicotine enhances visuospatial attention by deactivating areas of the resting brain default network. Journal of Neuroscience, 27(13), 3477-3489.	Wrong task
274	Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. Psychophysiology, 40(6), 895-903.	No IIV
275	Haley, A. P., Forman, D. E., Poppas, A., Hoth, K. F., Gunstad, J., Jefferson, A. L., et al. (2007). Carotid artery intima-media thickness and cognition in cardiovascular disease. International Journal of Cardiology, 121(2), 148-154.	No RT
276	Hallgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. Journal of the American Academy of Audiology, 12(7), 357-370.	Wrong task
277	Hammeke, T. A., & Hastings, J. E. (1988). Neuropsychologic alterations after cardiac operation.[see comment]. Journal of Thoracic & Cardiovascular Surgery, 96(2), 326-331.	No RT

278	Hanks, R. A., Ricker, J. H., & Millis, S. R. (Eds.). (2004). Empirical Evidence Regarding the Neuropsychological Assessment of Moderate and Severe Traumatic Brain Injury. New York, NY,US: Springer Publishing Co.	No empirical data
279	Hansell, N. K., Wright, M. J., Luciano, M., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2005). Genetic covariation between event-related potential (ERP) and behavioral non-ERP measures of working-memory, processing speed, and IQ. Behavior Genetics, 35(6), 695-706.	Sample consisted mainly of teenagers
280	Hansen, A. L., Johnsen, B. H., Sollers, J. J., 3rd, Stenvik, K., & Thayer, J. F. (2004). Heart rate variability and its relation to prefrontal cognitive function: the effects of training and detraining. European Journal of Applied Physiology, 93(3), 263-272.	No IIV
281	Hansen, A. L., Johnsen, B. H., Thornton, D., Waage, L., & Thayer, J. F. (2007). Facets of psychopathy, heart rate variability and cognitive function. Journal of Personality Disorders, 21(5), 568-582.	No IIV
282	Harrington, C., Salloway, S., & Malloy, P. (1997). Dramatic neurobehavioral disorder in two cases following bilateral anteromedial frontal lobe injury: Delayed psychosis and marked change in personality. Neurocase, 3(2), 137-149.	No RT
283	Harrison, M. J., & Ravdin, L. D. (2002). Cognitive dysfunction in neuropsychiatric systemic lupus erythematosus. Current Opinion in Rheumatology, 14(5), 510-514.	No empirical data
284	Harsh, J., Stone, P., Leiker, M., & Badia, P. (1990). Ultradian rhythms in responsiveness to stimuli presented during sleep. Biological Psychology, 31(3), 245-256.	Wrong task
285	Harvey, P. D., Lombardi, J., Kincaid, M. M., Parrella, M., White, L., Powchik, P., et al. (1995). Cognitive functioning in chronically hospitalized schizophrenic patients: age-related changes and age disorientation as a predictor of impairment. Schizophrenia Research, 17(1), 15-24.	No RT
286	Hasbroucq, T., Kornblum, S., & Osman, A. (1988). A new look at reaction time estimates of interhemispheric transmission time. Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 8(3), 207-221.	No empirical data
287	Haskell, S. G., Richardson, E. D., & Horwitz, R. I. (1997). The effect of estrogen replacement therapy on cognitive function in women: a critical review of the literature. Journal of Clinical Epidemiology, 50(11), 1249-1264.	No empirical data
288	Hatsukami, D. K., Huber, M., Callies, A., & Skoog, K. (1993). Physical dependence on nicotine gum: Effect of duration of use. Psychopharmacology, 111(4), 449-456.	Wrong task
289	Hatsukami, D., Anton, D., Keenan, R., & Callies, A. (1992). Smokeless tobacco abstinence effects and nicotine gum dose. Psychopharmacology, 106(1), 60-66.	Wrong task
290	Haynes, E. T., & Phillips, J. P. (1973). Inconsistency, loose construing and schizophrenic thought disorder. British Journal of Psychiatry, 123(573), 209-217.	No RT
291	Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. Neuropsychology, 20(5), 511-528.	No RT
292	Heil, P., Neubauer, H., Tiefenau, A., & von Specht, H. (2006). Comparison of absolute thresholds derived from an adaptive forced-choice procedure and from reaction probabilities and reaction times in a simple reaction time paradigm. Jaro-Journal of the Association for Research in Otolaryngology, 7(3), 279-298.	Matter not relevant
293	Hemsley, D. R. (1977). What have cognitive deficits to do with schizophrenic symptoms? British Journal of Psychiatry, 130, 167-173.	No empirical data
294	Hening, W., Vicario, D., & Ghez, C. (1988). Trajectory control in targeted force impulses. IV. Influences of choice, prior experience and urgency. Experimental Brain Research, 71(1), 103-115.	Matter not relevant
295	Hermann, B. P., Seidenberg, M., Schoenfeld, J., Peterson, J., Leveroni, C., & Wyler, A. R. (1996). Empirical techniques for determining the reliability, magnitude, and pattern of neuropsychological change after epilepsy surgery. Epilepsia, 37(10), 942-950.	No RT
296	Herold, K. C., Polonsky, K. S., Cohen, R. M., Levy, J., & Douglas, F. (1985). Variable deterioration in cortical function during insulin-induced hypoglycemia. Diabetes, 34(7), 677-685.	No IIV
297	Herrmann, W. M., & Stephan, K. (1991). Efficacy and clinical relevance of cognition enhancers. Alzheimer Disease & Associated Disorders, 5 Suppl 1, S7-12.	No RT
298	Heuer, H., Kleinsorge, T., Klein, W., & Kohlisch, O. (2004). Total sleep deprivation increases the costs of shifting between simple cognitive tasks. Acta Psychologica, 117(1), 29-64.	Wrong task
299	Hilton, G. (1994). Behavioral and cognitive sequelae of head trauma. Orthopaedic Nursing, 13(4), 25-32.	No empirical data
300	Hinchliffe, F. J., Murdoch, B. E., Chenery, H. J., Baglioni, A. J., Jr., & Harding-Clark, J. (1998). Cognitive-linguistic subgroups in closed-head injury. Brain Injury, 12(5), 369-398.	No RT
301	Hinton, S. C., Paulsen, J. S., Hoffmann, R. G., Reynolds, N. C., Zimbelman, J. L., & Rao, S. M. (2007). Motor timing variability increases in preclinical Huntington's disease patients as estimated onset of motor symptoms approaches. Journal of the International Neuropsychological Society, 13(3), 539-543.	No RT
302	Hinton, V. J., Ho, Y. Y., Wang, D., Kranz-Eble, P., Yang, H., Englestadt, K., et al. (2001). Glut-1 deficiency syndrome: Phenotypic variability in cognition and behavior. Annals of Neurology, 50(3), S124-S124.	No RT
303	Hinz, A., HÃ¼ber, B., & Schreinicke, G. (1996). Covariation of cardiovascular parameters across subjects, situations, and replications. Journal of Psychophysiology, 10(2), 115-124.	No IIV
304	Hirdes, J. P., Frijters, D. H., & Teare, G. F. (2003). The MDS-CHESS Scale: A New Measure to Predict Mortality in Institutionalized Older People. Journal of the American Geriatrics Society, 51(1), 96-100.	No RT
305	Hoerr, R. (2003). Behavioural and psychological symptoms of dementia (BPSD): Effects of EGb 761. Pharmacopsychiatry, 36(SUPPL. 1), S56-S61.	No empirical data
306	Hogan, M. J. (2004). The cerebellum in thought and action: A fronto-cerebellar aging hypothesis. New Ideas in Psychology, 22(2), 97-125.	No empirical data

307	Hogan, M. J., Carolan, L., Roche, R. A. P., Dockree, P. M., Kaiser, J., Bunting, B. P., et al. (2006). Electrophysiological and information processing variability predicts memory decrements associated with normal age-related cognitive decline and Alzheimer's disease (AD). Brain Research, 1119(1), 215-226.	Wrong task
308	Holdnack, J. A., Moberg, P. J., Arnold, S. E., Gur, R. C., & Gur, R. E. (1995). Speed of Processing and Verbal-Learning Deficits in Adults Diagnosed with Attention-Deficit Disorder. Neuropsychiatry Neuropsychology and Behavioral Neurology, 8(4), 282-292.	Wrong task
309	Holdstock, L., & de Wit, H. (2001). Individual differences in responses to ethanol and d-amphetamine: a within-subject study. Alcoholism: Clinical & Experimental Research, 25(4), 540-548.	No IIV
310	Hollister, L. E. (1990). Problems in the search for cognition enhancers. Pharmacopsychiatry, 23(2), 33-36.	No empirical data
311	Holm, A., Ranta-aho, P. O., Sallinen, M., Karjalainen, P. A., & MÃ¼ller, K. (2006). Relationship of P300 single-trial responses with reaction time and preceding stimulus sequence. International Journal of Psychophysiology, 61(2), 244-252.	Wrong task
312	Honey, G. D., Sharma, T., Suckling, J., Giampietro, V., Soni, W., Williams, S. C. R., et al. (2003). The functional neuroanatomy of schizophrenic subsyndromes. Psychological Medicine, 33(6), 1007-1018.	No IIV
313	Hot, P., Naveteur, J., Leconte, P., & Sequeira, H. (1999). Diurnal variations of tonic electrodermal activity. International Journal of Psychophysiology, 33(3), 223-230.	No IIV
314	Huber, S. J., Freidenberg, D. L., Shuttleworth, E. C., Paulson, G. W., & Clapp, L. E. (1989). Neuropsychological similarities in lateralized parkinsonism. Cortex, 25(3), 461-470.	No RT
315	Hudnell, H. K., Otto, D. A., & House, D. E. (1996). The influence of vision on computerized neurobehavioral test scores: a proposal for improving test protocols. Neurotoxicology & Teratology, 18(4), 391-400.	No IIV
316	Hughes, J. R. , Keenan, R. M., & Yellin, A. (1989). Effect of tobacco withdrawal on sustained attention. Addictive Behaviors, 14(5), 577-580.	Wrong task
317	Hultsch, D. F., Strauss, E., Hunter, M. A., & MacDonald, S. W. S. (Eds.). (2008). Intraindividual variability, cognition, and aging. New York, NY,US: Psychology Press.	No empirical data
318	Huntley, M. S. (1972). Influences of alcohol and S-R uncertainty upon spatial localization time. Psychopharmacologia, 27(2), 131-140.	No IIV
319	Hutter, B. O., Gilsbach, J. M., & Kreitschmann, I. (1995). Quality of life and cognitive deficits after subarachnoid haemorrhage. British Journal of Neurosurgery, 9(4), 465-475.	No IIV
320	Iacoboni, M., & Zaidel, E. (2000). Cross-uncrossed difference in simple reaction times to lateralized flashes: Between- and within-subjects variability. Neuropsychologia, 38(5), 535-541.	Matter not relevant
321	Ingum, J., Pettersen, G., Sager, G., & Morland, J. (1994). Relationship between unbound plasma concentrations and various psychomotor and subjective effects after intakes of diazepam and flunitrazepam. International Clinical Psychopharmacology, 9(2), 115-121.	No IIV
322	Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (1993). Effects of Aging on Event-Related Brain Potentials and Reaction-Times in an Auditory Oddball Task. Psychophysiology, 30(1), 10-22.	Wrong task
323	Isotalo, E., Pyykko, I., Juhola, M., & Aalto, H. (1995). Predictable and pseudo random saccades in patients with acoustic neuroma. Acta Oto-Laryngologica, Supplement(520 I), 22-24.	No IIV
324	Ito, J., & Takamatsu, N. (1997). Somatosensory event-related potentials in healthy subjects: Single trial analysis and averages of reaction time terciles. Journal of Psychophysiology, 11(1), 2-11.	Wrong task
325	Ivnik, R. J., Smith, G. E., Petersen, R. C., Boeve, B. F., Kokmen, E., & Tangalos, E. G. (2000). Diagnostic accuracy of four approaches to interpreting neuropsychological test data. Neuropsychology, 14(2), 163-177.	No RT
326	Ivry, R. B., Keele, S. W., & Diener, H. C. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. Experimental Brain Research, 73(1), 167-180.	No RT
327	Jagannathan, V., & Venitz, J. (1997). Pharmacokinetics and CNS pharmacodynamics of the 5-HT1A agonist buspirone in humans following acute L-tryptophan depletion challenge. Methods & Findings in Experimental & Clinical Pharmacology, 19(5), 351-362.	No IIV
328	Jansma, J. M., Ramsey, N. F., Slagter, H. A., & Kahn, R. S. (2001). Functional anatomical correlates of controlled and automatic processing. Journal of Cognitive Neuroscience, 13(6), 730-743.	Matter not relevant
329	Jasiewicz, J., & Simmons, R. W. (1996). Response timing accuracy as a function of movement velocity and distance. Journal of Motor Behavior, 28(3), 224-232.	Matter not relevant
330	Jason, G. W., Suchowersky, O., Pajurkova, E. M., Graham, L., Klimek, M. L., Garber, A. T., et al. (1997). Cognitive manifestations of Huntington disease in relation to genetic structure and clinical onset. Archives of Neurology, 54(9), 1081-1088.	No RT
331	Jensen, A. R. (2002). Galton's legacy to research on intelligence. Journal of Biosocial Science, 34(2), 145-172.	No empirical data
332	Jensen, A. R., & Reed, T. E. (1992). The correlation between reaction time and the ponderal index. Perceptual and Motor Skills, 75(3), 843-846.	Matter not relevant
333	Jensen, H. H., Hansen, H. C., & Drenck, N. E. (1993). Response variability to sedative effects of diazepam in normal subjects. Scandinavian Journal of Behaviour Therapy, 22(3), 171-178.	No empirical data
334	Johnels, B., Ingvarsson, P. E., Thorselius, M., Valls, M., & Steg, G. (1989). Disability profiles and objective quantitative assessment in Parkinson's disease. Acta Neurologica Scandinavica, 79(3), 227-238.	No RT
335	Johnson, B. F., Hoch, K., & Johnson, J. (1991). Variability in psychometric test scores: The importance of the practice effect in patient study design. Progress in Neuro-Psychopharmacology & Biological Psychiatry, 15(5), 625-635.	Matter not relevant

336	Johnston, D. W., Schmidt, T. F. H., Albus, C., Vagt, S., & et al. (1994). The relationship between cardiovascular reactivity in the laboratory and heart rate response in real life: Active coping and beta blockade. Psychosomatic Medicine, 56(4), 369-376.	No IIV
337	Judd, P. H. (2005). Neurocognitive impairment as a moderator in the development of borderline personality disorder. Development & Psychopathology, 17(4), 1173-1196.	No empirical data
338	Juhel, J. (1993). Should we take the shape of reaction time distributions into account when studying the relationship between RT and psychometric intelligence? Personality and Individual Differences, 15(3), 357-360.	Wrong task
339	Kahlon, B., Sundbarg, G., & Rehncrona, S. (2002). Comparison between the lumbar infusion and CSF tap tests to predict outcome after shunt surgery in suspected normal pressure hydrocephalus.[see comment]. Journal of Neurology, Neurosurgery & Psychiatry, 73(6), 721-726.	No IIV
340	Kaiser, S., Roth, A., Rentrop, M., Friederich, H.-C., Bender, S., & Weisbrod, M. (2008). Intra-individual reaction time variability in schizophrenia, depression and borderline personality disorder. Brain and Cognition, 66(1), 73-82.	Wrong task
341	Kalkanis, S. N., Blumenfeld, H., Sherman, J. C., Krebs, D. E., Irizarry, M. C., Parker, S. W., et al. (1996). Delayed complications thirty-six years after hemispherectomy: a case report. Epilepsia, 37(8), 758-762.	No RT
342	Kalra, L., Jackson, S. H., & Swift, C. G. (1993). Assessment of changes in psychomotor performance of elderly subjects. British Journal of Clinical Pharmacology, 36(5), 383-389.	Wrong IIIV
343	Kapur, N., Abbott, P., Lowman, A., & Will, R. G. (2003). The neuropsychological profile associated with variant Creutzfeldt-Jakob disease. Brain, 126(Pt 12), 2693-2702.	No RT
344	Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. Psychophysiology, 44(1), 128-144.	Matter not relevant
345	Kato, A., Kato, M., Ishii, H., Ichimiya, Y., Suzuki, K., Kawasaki, H., et al. (2004). Development of quantitative neuropsychological tests for diagnosis of subclinical hepatic encephalopathy in liver cirrhosis patients and establishment of diagnostic criteria - multicenter collaborative study in Japanese. Hepatology Research, 30(2), 71-78.	No IIV
346	Kay, G. G., & Wesnes, K. A. (2005). Pharmacodynamic effects of darifenacin, a muscarinic M selective receptor antagonist for the treatment of overactive bladder, in healthy volunteers. BJU International, 96(7), 1055-1062.	No IIV
347	Kaye, J. A., DeCarli, C., Luxenberg, J. S., & Rapoport, S. I. (1992). The significance of age-related enlargement of the cerebral ventricles in healthy men and women measured by quantitative computed X-ray tomography. Journal of the American Geriatrics Society, 40(3), 225-231.	No RT
348	Keizer, A. M. A., Hijman, R., Kalkman, C. J., Kahn, R. S., van Dijk, D., & Octopus Study, G. (2005). The incidence of cognitive decline after (not) undergoing coronary artery bypass grafting: the impact of a controlled definition. Acta Anaesthesiologica Scandinavica, 49(9), 1232-1235.	No RT
349	Kelly, A. M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2008). Competition between functional brain networks mediates behavioral variability. Neuroimage, 39(1), 527-537.	Wrong task
350	Kent, T. A., Gunn, W. H., & Goodwin, D. W. (1986). Individual differences in state-dependent retrieval effects of alcohol intoxication. Journal of Studies on Alcohol, 47(3), 241-243.	No RT
351	Keri, S., & Janka, Z. (2004). Critical evaluation of cognitive dysfunctions as endophenotypes of schizophrenia. Acta Psychiatrica Scandinavica, 110(2), 83-91.	No empirical data
352	Kerr, J. S., & Hindmarch, I. (1998). The effects of alcohol alone or in combination with other drugs on information processing, task performance and subjective responses. Human Psychopharmacology: Clinical and Experimental, 13(1), 1-9.	No empirical data
353	Ketcham, C. J., Seidler, R. D., Van Gemmert, A. W. A., & Stelmach, G. E. (2002). Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 57(1), P54-64.	No IIV
354	Keyser, L., Witol, A. D., Kreutzer, J. S., & Rosenthal, M. (1995). A Multicenter Investigation of Neurobehavioral Outcome after Traumatic Brain Injury. Neurorehabilitation, 5(3), 255-267.	No RT
355	Kiesel, A., & Miller, J. (2007). Impact of contingency manipulations on accessory stimulus effects. Perception & Psychophysics, 69(7), pp.	Matter not relevant
356	Kilburn, K. H. (1994). Neurobehavioral impairment and seizures from formaldehyde. Archives of Environmental Health, 49(1), 37-44.	No IIV
357	Kim, H., Guilleminault, C., Hong, S., Kim, D., Kim, S., Go, H., et al. (2001). Pattern analysis of sleep-deprived human EEG. Journal of Sleep Research, 10(3), 193-201.	No RT
358	Kindermann, S. S., Karimi, A., Symonds, L., Brown, G. G., & Jeste, D. V. (1997). Review of functional magnetic resonance imaging in schizophrenia. Schizophrenia Research, 27(2-3), 143-156.	No empirical data
359	King, D. J. (1990). The effect of neuroleptics on cognitive and psychomotor function. British Journal of Psychiatry, 157, 799-811.	No empirical data
360	Kitagawa, H., Takenouchi, T., Azuma, R., Wesnes, K. A., Kramer, W. G., Clody, D. E., et al. (2003). Safety, pharmacokinetics, and effects on cognitive function of multiple doses of GTS-21 in healthy, male volunteers. Neuropsychopharmacology, 28(3), 542-551.	No IIV
361	Kixmiller, J. S., Briggs, J. R., Hartlage, L. C., & Dean, R. S. (1993). Factor structure of emotional and cognitive behaviors for normals and neurologically impaired patients. Journal of Clinical Psychology, 49(2), 233-241.	No RT
362	Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-Subject Variability in Attention-Deficit Hyperactivity Disorder. Biological Psychiatry, 60(10), 1088-1097.	Sample consisted of children
363	Kleykamp, B. A., Jennings, J. M., Blank, M. D., & Eissenberg, T. (2005). The effects of nicotine on attention and working memory in never-smokers. Psychology of Addictive Behaviors, 19(4), 433-438.	Wrong task

364	Klimovitch Lofthus, G., & Hanson, C. J. (1981). Fractionated reaction times and performance times as a function of practice. American Journal of Physical Medicine, 60(4), 190-197.	Matter not relevant
365	Klonowicz, T., Zawadzka, G., & Zawadzki, B. (1987). Reactivity, arousal, and coping with stress. Personality and Individual Differences, 8(6), 793-798.	Sample consisted of high school students
366	Kobrick, J. L. (1972). Effects of hypoxia on voluntary response time to peripheral stimuli during central target monitoring. Ergonomics, 15(2), 147-156.	No IIV
367	Kohler, M., Pavy, A., & Van Den Heuvel, C. (2006). The effects of chewing versus caffeine on alertness, cognitive performance and cardiac autonomic activity during sleep deprivation. Journal of Sleep Research, 15(4), 358-368.	No IIV
368	Kopp, M. r. S., MihÃ¡ly, K., Linka, E., & Bitter, I. n. (1987). Electrodermally differentiated subgroups of anxiety patients: I Automatic and vigilance characteristics. International Journal of Psychophysiology, 5(1), 43-51.	No IIV
369	Kozma, C. (1998). On cognitive variability in velocardiofacial syndrome: profound mental retardation and autism. American Journal of Medical Genetics, 81(3), 269-270.	No RT
370	Kranzler, J. H. (1992). A test of Larson and Alderton's (1990) worst performance rule of reaction time variability. Personality and Individual Differences, 13(3), 255-261.	Matter not relevant
371	Kranzler, J. H. (1992). The Skewness of the Distribution of Rt Trials Does Not Correlate with Psychometric-G. Personality and Individual Differences, 13(8), 945-946.	Matter not relevant
372	Kruglikov, S. Y., & Schiff, S. J. (2003). Interplay of electroencephalogram phase and auditory-evoked neural activity. Journal of Neuroscience, 23(31), 10122-10127.	No RT
373	Kubova, Z., Chlubnova, J., Szanyi, J., Kuba, M., & Kremlacek, J. (2005). Influence of physiological changes of glycemia on VEPs and visual ERPs. Physiological Research, 54(2), 245-250.	Wrong task
374	Kugler, C. F. A., Taghavy, A., & Platt, D. (1993). The Event-Related P300 Potential Analysis of Cognitive Human Brain Aging - a Review. Gerontology, 39(5), 280-303.	No empirical data
375	Kuningas, M., De Rijk, R. H., Westendorp, R. G. J., Jolles, J., Eline Slagboom, P., & Van Heemst, D. (2007). Mental performance in old age dependent on cortisol and genetic variance in the mineralocorticoid and glucocorticoid receptors. Neuropsychopharmacology, 32(6), 1295-1301.	No RT
376	Kuningas, M., Slagboom, P. E., Westendorp, R. G. J., & Van Heemst, D. (2006). Impact of genetic variations in the WRN gene on age related pathologies and mortality. Mechanisms of Ageing and Development, 127(3), 307-313.	No RT
377	Kupper, Z., & Hoffmann, H. (2000). Course patterns of psychosocial functioning in schizophrenia patients attending a vocational rehabilitation program. Schizophrenia Bulletin, 26(3), 681-698.	No RT
378	Laessoe, U., Hoeck, H. C., Simonsen, O., & Voigt, M. (2008). Residual attentional capacity amongst young and elderly during dual and triple task walking. Human Movement Science, 27(3), 496-512.	No RT
379	Lambert, J. C., Ferreira, S., Gussekloo, J., Christiansen, L., Brysbaert, G., Slagboom, E., et al. (2007). Evidence for the association of the S100beta gene with low cognitive performance and dementia in the elderly. Molecular Psychiatry, 12(9), 870-880.	No RT
380	Landgraf, R. (2006). Neuropeptides in anxiety and depression. Amino Acids, 31(3), 211-213.	No empirical data
381	Lannoo, E., De Paepe, A., Leroy, B., & Thiery, E. (1996). Neuropsychological aspects of Marfan syndrome. Clinical Genetics, 49(2), 65-69.	No RT
382	Lansing, R. W., & Meyerink, L. (1981). Load compensating responses of human abdominal muscles. Journal of Physiology, Vol.320, 253-268.	Wrong task
383	Lansing, R. W., Meyerink, L. H., & Thomas, J. E. (Eds.). (1988). Human respiratory reaction times. Hillsdale, NJ,England: Lawrence Erlbaum Associates, Inc.	No IIV
384	Laures, J. S. (2005). Reaction time and accuracy in individuals with aphasia during auditory vigilance tasks. Brain and Language, 95(2), 353-357.	Wrong task
385	Laws, K. R., & Kondel, T. K. (1998). Variability in cognitive deterioration in schizophrenia.[comment]. British Journal of Psychiatry, 172, 187-188.	No empirical data
386	Le Ber, I., Moreira, M.-C., Rivaud-Pechoux, S., Chamayou, C., Ochsner, F., Kuntzer, T., et al. (2003). Cerebellar ataxia with oculomotor apraxia type 1: clinical and genetic studies. Brain, 126(Pt 12), 2761-2772.	No RT
387	Lecerf, T., Ghisletta, P., & Jouffray, C. (2004). Intraindividual variability and level of performance in four visuo-spatial working memory tasks. Swiss Journal of Psychology, 63(4), 261-272.	No RT
388	Lee, S. B., & Berg, W. D. (1998). Development of safety-based level-of-service parameters for two-way stop-controlled intersections. In Highway Safety Modeling, Analysis, and Design (pp. 127-132).	No participants - simulated data
389	Lee, V. L. (1996). Variability of human latencies with and without a criterion for reinforcement. Psychological Record, 46(3), 519-538.	Matter not relevant
390	Lee, W. A. (1980). Anticipatory control of postural and task muscles during rapid arm flexion. Journal of Motor Behavior, 12(3), 185-196.	Matter not relevant
391	Leegaard, O. F. (1983). Diffuse cerebral symptoms in convalescents from cerebral infarction and myocardial infarction. Acta Neurologica Scandinavica, 67(6), 348-355.	No RT
392	Leibold, L. J., & Werner, L. A. (2002). Relationship between intensity and reaction time in normal-hearing infants and adults.[erratum appears in Ear Hear 2002 Jun;23(3):267]. Ear & Hearing, 23(2), 92-97.	Matter not relevant
393	Leibovici, D., Ritchie, K., Ledesert, B., & Touchon, J. (1999). The effects of wine and tobacco consumption on cognitive performance in the elderly: a longitudinal study of relative risk. International Journal of Epidemiology, 28(1), 77-81.	Wrong task

394	Leis, B. C., Rand, M. K., Van Gemmert, A. W. A., Longstaff, M. G., Lou, J. S., & Stelmach, G. E. (2005). Movement precues in planning and execution of aiming movements in Parkinson's disease. Experimental Neurology, 194(2), 393-409.	No IIV
395	Lemmer, B., Kern, R. I., Nold, G., & Lohrer, H. (2002). Jet lag in athletes after eastward and westward time-zone transition. Chronobiology International, 19(4), 743-764.	No RT
396	Leonard, C. M., Eckert, M. A., & Kuldau, J. M. (2006). Exploiting human anatomical variability as a link between genome and cognome. Genes, Brain, & Behavior, 5 Suppl 1, 64-77.	No empirical data
397	Lessov-Schlaggar, C. N., Swan, G. E., Reed, T., Wolf, P. A., & Carmelli, D. (2007). Longitudinal genetic analysis of executive function in elderly men. Neurobiology of Aging, 28(11), 1759-1768.	No RT
398	Leung, W. W., McClure, M. M., Siever, L. J., Barch, D. M., & Harvey, P. D. (2007). Catechol-O-methyltransferase Val158Met genotype in healthy and personality disorder individuals: Preliminary results from an examination of cognitive tests hypothetically differentially sensitive to dopamine functions. Neuropsychiatric Disease and Treatment, 3(6), 925-934.	Wrong task
399	Levin, E. D., Conners, C. K., Sparrow, E., Hinton, S. C., Erhardt, D., Meck, W. H., et al. (1996). Nicotine effects on adults with attention-deficit/hyperactivity disorder. Psychopharmacology, 123(1), 55-63.	Wrong task
400	Levin, E. D., Wilson, W., Rose, J. E., & McEvoy, J. (1996). Nicotine-haloperidol interactions and cognitive performance in schizophrenics. Neuropsychopharmacology, 15(5), 429-436.	No IIV
401	Levine, A. J., Hardy, D. J., Miller, E., Castellon, S. A., Longshore, D., & Hinkin, C. H. (2006). The Effect of Recent Stimulant Use on Sustained Attention in HIV-Infected Adults. Journal of Clinical and Experimental Neuropsychology, 28(1), 29-42.	Wrong task
402	Levine, B., Fujiwara, E., O'Connor, C., Richard, N., Kovacevic, N., Mandic, M., et al. (2006). In vivo characterization of traumatic brain injury neuropathology with structural and functional neuroimaging. Journal of Neurotrauma, 23(10), 1396-1411.	No empirical data
403	Levinson, B., Wright, P., & Barklem, S. (1985). Effect of buflomedil on behaviour, memory, and intellectual capacity in patients with dementia. A placebo-controlled study. South African Medical Journal, Suid-Afrikaanse Tydskrif Vir Geneeskunde. 68(5), 302-307.	No RT
404	Levy, J. (1983). Is cerebral asymmetry of function a dynamic process? Implications for specifying degree of lateral differentiation. Neuropsychologia, 21(1), 3-11.	Wrong task
405	Levy, J. A., & Chelune, G. J. (2007). Cognitive-behavioral profiles of neurodegenerative dementias: beyond Alzheimer's disease. Journal of Geriatric Psychiatry & Neurology, 20(4), 227-238.	No empirical data
406	Lewis, G. N., Byblow, W. D., & Walt, S. E. (2000). Stride length regulation in Parkinson's disease: the use of extrinsic, visual cues. Brain, 123, 2077-2090.	No IIV
407	Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. Psychological Science, 15(3), 155-163.	Wrong task
408	Li, S.-C., Lindenberger, U., & Frensch, P. A. (2000). Unifying cognitive aging: From neuromodulation to representation to cognition. Neurocomputing: An International Journal, 32, 879-890.	No empirical data
409	Liao, M.-J., Jagacinski, R. J., & Greenberg, N. (1997). Quantifying the performance limitations of older and younger adults in a target acquisition task. Journal of Experimental Psychology: Human Perception and Performance, 23(6), 1644-1664.	Wrong task
410	Liberman, J. N., Stewart, W. F., Wesnes, K., & Troncoso, J. (2002). Apolipoprotein E epsilon 4 and short-term recovery from predominantly mild brain injury. Neurology, 58(7), 1038-1044.	No IIV
411	Lim, J., Choo, W.-C., & Chee, M. W. L. (2007). Reproducibility of changes in behaviour and fMRI activation associated with sleep deprivation in a working memory task. Sleep: Journal of Sleep and Sleep Disorders Research, 30(1), 61-70.	Wrong task
412	Lindenberger, U., & von Oertzen, T. (Eds.). (2006). Variability in Cognitive Aging: From Taxonomy to Theory. New York, NY,US: Oxford University Press.	No empirical data
413	Lindenberger, U., Li, S. C., & Backman, L. (2006). Delineating brain-behavior mappings across the lifespan: Substantive and methodological advances in developmental neuroscience. Neuroscience and Biobehavioral Reviews, 30(6), 713-717.	No empirical data
414	Line, W., Griffin, J. D. M., & Anderson, G. W. (1935). The objective measurement of mental stability. Journal of Mental Science, 81, 61-106.	Wrong task
415	Linkenkaer-Hansen, K., Nikulin, V. V., Palva, S., Ilmoniemi, R. J., & Palva, J. M. (2004). Prestimulus oscillations enhance psychophysical performance in humans. Journal of Neuroscience, 24(45), 10186-10190.	Wrong task
416	Linnoila, M., Erwin, C. W., & Brendle, A. (1982). Efficacy and side effects of flunitrazepam and pentobarbital in severely insomniac patients. Journal of Clinical Pharmacology, 22(1), 14-19.	Wrong IIIV
417	Lisper, H. O., Dureman, I., Ericsson, S., & Karlsson, N. G. (1971). Effects of sleep deprivation and prolonged driving on a subsidiary auditory reaction time. Accident Analysis & Prevention, 2(4), 335-341.	Wrong task
418	Little, B. C., Matta, R. J., & Zahn, T. P. (1974). Physiological and psychological effects of progesterone in man. Journal of Nervous and Mental Disease, 159(4), 256-262.	No IIV
419	Litvan, I., Grimes, D. A., Lang, A. E., Jankovic, J., McKee, A., Verny, M., et al. (1999). Clinical features differentiating patients with postmortem confirmed progressive supranuclear palsy and corticobasal degeneration. Journal of Neurology, 246 Suppl 2, II1-5.	No RT
420	Liu, X., Ingram, H. A., Palace, J. A., & Miall, R. C. (1999). Dissociation of 'on-line' and 'off-line' visuomotor control of the arm by focal lesions in the cerebellum and brainstem. Neuroscience Letters, 264(1-3), 121-124.	No IIV

421	Loewenstein, D. A., Acevedo, A., Agron, J., Issacson, R., Strauman, S., Crocco, E., et al. (2006). Cognitive profiles in Alzheimer's disease and in mild cognitive impairment of different etiologies. Dementia & Geriatric Cognitive Disorders, 21(5-6), 309-315.	No RT
422	Loh, M., Rolls, E. T., & Deco, C. (2007). Statistical fluctuations in attractor networks related to schizophrenia. Pharmacopsychiatry, 40(SUPPL. 1), S78-S84.	No participants - simulated data
423	Lolas, F., Araya, A., de Andraca, I., & Pinto-Hamuy, T. (1976). Vertex slow potentials and classical reaction time task: Specificity and 'directional' anticipatory factors. Neuropsychologia, 14(4), 499-504.	No IIV
424	Lorenzo-Lopez, L., Amenedo, E., Pazo-Alvarez, P., & Cadaveira, F. (2007). Visual target processing in high- and low-performing older subjects indexed by P3 component. Neurophysiologie Clinique-Clinical Neurophysiology, 37(2), 53-61.	Wrong task
425	Lorist, M. M., Kernell, D., Meijman, T. F., & Zijdewind, I. (2002). Motor fatigue and cognitive task performance in humans. Journal of Physiology, 545(Pt 1), 313-319.	No IIV
426	Lou, H. C. (1994). Dopamine precursors and brain function in phenylalanine hydroxylase deficiency. Acta Paediatrica Supplement, 407, 86-88.	No empirical data
427	Lovden, M., Li, S.-C., Shing, Y. L., & Lindenberger, U. (2007). Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: Longitudinal data from the Berlin Aging Study. Neuropsychologia, 45(12), 2827-2838.	Wrong task
428	Luciano, M., Wright, M. J., Smith, G. A., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. Behavior Genetics, 31(6), 581-592.	Sample consisted mainly of teenagers
429	Lundervold, A. J., & Reinvang, I. (1995). Variability in cognitive function among persons at high genetic risk of Huntington's disease. Acta Neurologica Scandinavica, 91(6), 462-469.	No RT
430	Lundervold, A. J., Bergmann, N., & Wootton, C. (2005). Visual neglect in the first weeks after a stroke in the right hemisphere. Scandinavian Journal of Psychology, 46(3), 297-303.	No IIV
431	Luo, D. S., Thompson, L. A., & Detterman, D. K. (2003). The causal factor underlying the correlation between psychometric g and scholastic performance. Intelligence, 31(1), 67-83.	Sample consisted of children
432	Luo, D., Thompson, L. A., & Detterman, D. K. (2003). Phenotypic and behavioral genetic covariation between elemental cognitive components and scholastic measures. Behavior Genetics, 33(3), 221-246.	Sample consisted of children
433	Lupien, S. J., & Meaney, M. J. (Eds.). (1998). Stress, glucocorticoids, and hippocampal aging in rat and human. San Diego, CA,US: Academic Press.	No RT
434	Lussier, I., & Stip, E. (2001). Memory and attention deficits in drug naive patients with schizophrenia. Schizophrenia Research, 48(1), 45-55.	No IIV
435	Lutz, A., Lachaux, J.-P., Martinerie, J., & Varela, F. J. (2002). Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task. Proceedings of the National Academy of Sciences of the United States of America, 99(3), 1586-1591.	Wrong task
436	Lyros, E., Messinis, L., & Papathanasopoulos, P. (2008). Does motor subtype influence neurocognitive performance in Parkinson's disease without dementia? European Journal of Neurology, 15(3), 262-267.	No RT
437	MacDonald, S. W. S., Hultsch, D. F., & Bunce, D. (2006). Intraindividual Variability in Vigilance Performance: Does Degrading Visual Stimuli Mimic Age-Related 'Neural Noise?'. Journal of Clinical and Experimental Neuropsychology, 28(5), 655-675.	Wrong task
438	MacDonald, S. W. S., Nyberg, L., & Backman, L. (2006). Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. Trends in Neurosciences, 29(8), 474-480.	No empirical data
439	MacDonald, S. W. S., Nyberg, L., Sandblom, J., Fischer, H., & Backman, L. (2008). Increased response-time variability is associated with reduced inferior parietal activation during episodic recognition in aging. Journal of Cognitive Neuroscience, 20(5), 779-786.	Wrong task
440	Madden, D. J., Nebes, R. D., & Allen, P. A. (1992). Cognitive slowing in Alzheimer's disease as a function of task type and response type. Developmental Neuropsychology, 8(4), 459-471.	Wrong task
441	Madelain, L., Champrenaut, L., & Chauvin, A. (2007). Control of sensorimotor variability by consequences. Journal of Neurophysiology, 98(4), 2255-2265.	Matter not relevant
442	Majewska, M. D. (1996). Cocaine addiction as a neurological disorder: implications for treatment. NIDA Research Monograph, 163, 1-26.	No empirical data
443	Malisza, K. L. (2007). Neuroimaging cognitive function in fetal alcohol spectrum disorders. International Journal on Disability and Human Development, 6(2), 171-188.	No empirical data
444	Manni, R., Ratti, M. T., Perucca, E., Galimberti, C. A., & Tartara, A. (1993). A multiparametric investigation of daytime sleepiness and psychomotor functions in epileptic patients treated with phenobarbital and sodium valproate: a comparative controlled study. Electroencephalography & Clinical Neurophysiology, 86(5), 322-328.	No IIV
445	Manning, J. J., & Hammond, G. R. (1990). The effect of motor preparation on changes in H reflex amplitude during the response latency of a warned reaction time task. Journal of Motor Behavior, 22(2), 292-314.	No IIV
446	Marra, C., Silveri, M. C., & Gainotti, G. (2000). Predictors of cognitive decline in the early stage of probable Alzheimer's disease. Dementia & Geriatric Cognitive Disorders, 11(4), 212-218.	No RT
447	Martin del Campo, A. F., McMurray, R. G., Besser, G. M., & Grossman, A. (1992). Effect of 12-hour infusion of naloxone on mood and cognition in normal male volunteers. Biological Psychiatry, 32(4), 344-353.	No IIV
448	Martin, F., Delpont, E., Suisse, G., & Dolisi, C. (1993). Brain-Stem Auditory-Evoked Potentials - Differences Related to Physical-Activity. International Journal of Sports Medicine, 14(8), 427-432.	Wrong task

449	Martin, M., & Hofer, S. M. (2004). Intraindividual variability, change, and aging: conceptual and analytical issues. Gerontology, 50(1), 7-11.	No empirical data
450	Martin, T., Egly, R., Houck, J. M., Bish, J. P., Barrera, B. D., Lee, D. C., et al. (2005). Chronometric evidence for entrained attention. Perception & Psychophysics, 67(1), 168-184.	Matter not relevant
451	Martino, D. J., Strejilevich, S. A., Scapola, M., Igoa, A., Marengo, E., Ais, E. D., et al. (2008). Heterogeneity in cognitive functioning among patients with bipolar disorder. Journal of Affective Disorders, 109(1-2), 149-156.	Wrong task
452	Maruff, P., Malone, V., & Currie, J. (1995). Asymmetries in the covert orienting of visual spatial attention to spatial and non-spatial cues in Alzheimer's disease. Brain, 118(Pt 6), 1421-1435.	Wrong task
453	Mascord, D. J., & Heath, R. A. (1992). Behavioral and physiological indices of fatigue in a visual tracking task. Journal of Safety Research, 23(1), 19-25.	Wrong task
454	Maxson, P. J., Berg, S., & McClearn, G. (1997). Multidimensional patterns of aging: a cluster-analytic approach. Experimental Aging Research, 23(1), 13-31.	No RT
455	Mazzocco, M. M. (2000). Advances in research on the fragile X syndrome. Mental Retardation & Developmental Disabilities Research Reviews, 6(2), 96-106.	No empirical data
456	McAllister, T. W. (1992). Neuropsychiatric sequelae of head injuries. Psychiatric Clinics of North America, 15(2), 395-413.	No empirical data
457	McCann, J. C., & Ames, B. N. (2008). Is there convincing biological or behavioral evidence linking vitamin D deficiency to brain dysfunction? Faseb Journal, 22(4), 982-1001.	No empirical data
458	McCarthy, M. E., & Waters, W. F. (1997). Decreased attentional responsivity during sleep deprivation: Orienting response latency, amplitude, and habituation. Sleep: Journal of Sleep Research & Sleep Medicine, 20(2), 115-123.	Wrong task
459	McClernon, F. J., Kollins, S. H., Lutz, A. M., Fitzgerald, D. P., Murray, D. W., Redman, C., et al. (2008). Effects of smoking abstinence on adult smokers with and without attention deficit hyperactivity disorder: results of a preliminary study. Psychopharmacology, 197(1), 95-105.	Wrong task
460	McCrea, M., Guskiewicz, K. M., Marshall, S. W., Barr, W., Randolph, C., Cantu, R. C., et al. (2003). Acute effects and recovery time following concussion in collegiate football players: the NCAA Concussion Study.[see comment]. JAMA, 290(19), 2556-2563.	No RT
461	McDonald, W. I., & Ron, M. A. (1999). Multiple sclerosis: the disease and its manifestations. Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 354(1390), 1615-1622.	No empirical data
462	McInnes, G. T., Bunting, E. A., Ings, R. M., Robinson, J., & Ankier, S. I. (1985). Pharmacokinetics and pharmacodynamics following single and repeated nightly administrations of loprazolam, a new benzodiazepine hypnotic. British Journal of Clinical Pharmacology, 19(5), 649-656.	No IIV
463	McKay, C. L., Strauss, M. S., Kinney, J. A., & Luria, S. M. (1977). Visual evoked responses, EEG's and reaction time during a normoxic saturation dive, NISAT I. Undersea Biomedical Research, 4(2), 131-145.	No IIV
464	McKay, K. E., & Halperin, J. M. (2001). ADHD, aggression, and antisocial behavior across the lifespan. Interactions with neurochemical and cognitive function. Annals of the New York Academy of Sciences, 931, 84-96.	No empirical data
465	McKee, P. J., Blacklaw, J., Butler, E., Gillham, R. A., & Brodie, M. J. (1992). Variability and clinical relevance of the interaction between sodium valproate and carbamazepine in epileptic patients. Epilepsy Research, 11(3), 193-198.	No IIV
466	McKeith, I. G., & Mosimann, U. P. (2004). Dementia with Lewy bodies and Parkinson's disease. Parkinsonism & Related Disorders, 10, S15-S18.	No RT
467	McNair, D. M., Kahn, R. J., Frankenthaler, L. M., & Faldetta, L. L. (1984). Amoxapine and amitriptyline. II. Specificity of cognitive effects during brief treatment of depression. Psychopharmacology, 83(2), 134-139.	Wrong task
468	Medina, K. L., Shear, P. K., Schafer, J., Armstrong, T. G., & Dyer, P. (2004). Cognitive functioning and length of abstinence in polysubstance dependent men. Archives of Clinical Neuropsychology, 19(2), 245-258.	No RT
469	Mega, M. S., Thompson, P. M., Cummings, J. L., Back, C. L., Xu, M. L., Zohoori, S., et al. (1998). Sulcal variability in the Alzheimer's brain - Correlations with cognition. Neurology, 50(1), 145-151.	No RT
470	Meijers, L. M., & Eijkman, E. G. (1974). The motor system in simple reaction time experiments. Acta Psychologica, 38(5), 367-377.	No human participants
471	Meng, L. F., Lu, C. P., & Chan, H. L. (2007). Time-varying brain potentials and interhemispheric coherences of anterior and posterior regions during repetitive unimanual finger movements. Sensors, 7(6), 960-978.	No IIV
472	Menz, H. B., Lord, S. R., & Fitzpatrick, R. C. (2003). Age-related differences in walking stability. Age and Ageing, 32(2), 137-142.	No IIV
473	Mercer, V. S., & Sahrmann, S. A. (1999). Postural synergies associated with a stepping task. Physical Therapy, 79(12), 1142-1152.	No IIV
474	Meric, C., & Collet, L. (1994). Differential effects of visual attention on spontaneous and evoked otoacoustic emissions. International Journal of Psychophysiology, 17(3), 281-289.	No IIV
475	Merry, A. F., & Warman, G. R. (2006). Fatigue and the anaesthetist. Anaesthesia and Intensive Care, 34(5), 577-578.	No empirical data
476	Meyer-Lindenberg, A., Straub, R. E., Lipska, B. K., Verchinski, B. A., Goldberg, T., Callicott, J. H., et al. (2007). Genetic evidence implicating DARPP-32 in human frontostriatal structure, function, and cognition. Journal of Clinical Investigation, 117(3), 672-682.	Wrong task

477	Michiels, V., & Cluydts, R. (2001). Neuropsychological functioning in chronic fatigue syndrome: A review. Acta Psychiatrica Scandinavica, 103(2), 84-93.	No empirical data
478	Michiels, V., Cluydts, R., Fischler, B., Hoffmann, G., Le Bon, O., & De Meirleir, K. (1996). Cognitive functioning in patients with chronic fatigue syndrome. Journal of Clinical & Experimental Neuropsychology: Official Journal of the International Neuropsychological Society, 18(5), 666-677.	No RT
479	Middleton, H. C., Coull, J. T., Sahakian, B. J., & Robbins, T. W. (1994). Clonidine-induced changes in the spectral distribution of heart rate variability correlate with performance on a test of sustained attention. Journal of Psychopharmacology, 8(1), 1-7.	Wrong task
480	Mikhelashvili-Browner, N., Yousem, D. M., Mandir, A. S., Calhoun, V. D., Wu, C., Oguz, K. K., et al. (2002). Correlation of reaction time in and out of the functional MR unit. Academic Radiology, 9(5), 513-519.	No IIV
481	Mileson, B. E., & Ferenc, S. A. (2001). Methods to identify and characterize developmental neurotoxicity for human health risk assessment: Overview. Environmental Health Perspectives, 109(SUPPL. 1), 77-78.	No empirical data
482	Miller, E. M. (1994). Intelligence and brain myelination: A hypothesis. Personality and Individual Differences, 17(6), 803-832.	No empirical data
483	Miller, J., & Ulrich, R. (1998). Locus of the effect of the number of alternative responses: Evidence from the lateralized readiness potential. Journal of Experimental Psychology: Human Perception and Performance, 24(4), 1215-1231.	No IIV
484	Miller, L. S., & Miller, S. E. (1996). Caffeine enhances initial but not extended learning of a proprioceptive-based discrimination task in nonsmoking moderate users. Perceptual & Motor Skills, 82(3 Pt 1), 891-898.	No IIV
485	Mitnitski, A. B., Graham, J. E., Mogilner, A. J., & Rockwood, K. (1999). The rate of decline in function in Alzheimer's disease and other dementias. Journals of Gerontology Series A-Biological Sciences & Medical Sciences, 54(2), M65-69.	No RT
486	Mitsutake, G., Otsuka, K., Oinuma, S., Ferguson, I., Cornelissen, G., Wanliss, J., et al. (2004). Does exposure to an artificial ULF magnetic field affect blood pressure, heart rate variability and mood? Biomedicine & Pharmacotherapy, 58 Suppl 1, S20-27.	No IIV
487	Moffat, S. D. (2005). Effects of testosterone on cognitive and brain aging in elderly men. Annals of the New York Academy of Sciences, 1055, 80-92.	No empirical data
488	Mohamed, M. A., Yousem, D. M., Tekes, A., Browner, N., & Calhoun, V. D. (2004). Correlation between the amplitude of cortical activation and reaction time: A functional MRI study. American Journal of Roentgenology, 183(3), 759-765.	No IIV
489	Monnet, F. P. (2003). Behavioural disturbances following Japanese B encephalitis. European Psychiatry: the Journal of the Association of European Psychiatrists, 18(6), 269-273.	No empirical data
490	Montgomery, G. K. (1995). A Multifactor Account of Disability after Brain Injury - Implications for Neuropsychological Counseling. Brain Injury, 9(5), 453-469.	No empirical data
491	Mooijaart, S. P., Van Vliet, P., Van Heemst, D., Rensen, P. C. N., Berbee, J. F. P., Jolles, J., et al. (2007). Plasma levels of apolipoprotein E and cognitive function in old age. Biogerontology: Mechanisms and InterventionsAnnals of the New York Academy of Sciences, 1100, 148-161.	No RT
492	Moretti, R., Torre, P., Vilotti, C., Antonello, R. M., & Pizzolato, G. (2007). Rivastigmine and Parkinson dementia complex. Expert Opinion on Pharmacotherapy, 8(6), 817-829.	No empirical data
493	Morgan, K. D., & David, A. S. (Eds.). (2004). Neuropsychological studies of insight in patients with psychotic disorders. New York, NY,US: Oxford University Press.	No empirical data
494	Morley, K. I., & Montgomery, G. W. (2001). The genetics of cognitive processes: Candidate genes in humans and animals. Behavior Genetics, 31(6), 511-531	No empirical data
495	Morris, A. M., Yuen, S., Lee, K. A., Lash, A. A., & Becker, C. E. (1992). The P300 Event-Related Potential - the Effects of Sleep-Deprivation. Journal of Occupational and Environmental Medicine, 34(12), 1143-1152.	Wrong task
496	Morris, J. C., & McManus, D. Q. (1991). The Neurology of Aging - Normal Versus Pathological Change. Geriatrics, 46(8), 47-54.	No empirical data
497	Morrison, C. E., Borod, J. C., Perrine, K., Beric, A., Brin, M. F., Rezai, A., et al. (2004). Neuropsychological functioning following bilateral subthalamic nucleus stimulation in Parkinson's disease. Archives of Clinical Neuropsychology, 19(2), 165-181.	Wrong task
498	Morse, C. K. (1993). Does variability increase with age? An archival study of cognitive measures. Psychology and Aging, 8(2), 156-164.	No empirical data
499	Morse, C. K. (1999). Age and variability in Francis Galton's data. Journal of Genetic Psychology, 160(1), 99-104.	No IIV
500	Mortel, K. F., Pavol, M. A., Wood, S., Meyer, J. S., Terayama, Y., Rexer, J. L., et al. (1994). Prospective studies of cerebral perfusion and cognitive testing among elderly normal volunteers and patients with ischemic vascular dementia and Alzheimer's disease. Angiology, 45(3), 171-180.	No RT
501	Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. Journal of Autism & Developmental Disorders, 36(1), 27-43.	No empirical data
502	Muir, K. W., Grosset, D. G., & Lees, K. R. (1995). Clinical pharmacology of CNS 1102 in volunteers. Annals of the New York Academy of Sciences, 765, 279-289; discussion 298.	No IIV
503	Mulder, G., & Mulder, L. J. M. (1981). Information processing and cardiovascular control. Psychophysiology, 18(4), 392-402.	Wrong task
504	Mulder, G., & van der Meulen, M.-H. (1972). Heart rate variability in a binary choice reaction task: an evaluation of some scoring methods. Acta Psychologica, 36(3), 239-251.	No IIV

505	Mulholland, T. B. (1990). Alpha-contingent stimulation with binocular and monocular viewing. International Journal of Psychophysiology, 9(3), 249-256.	No RT
506	Muller, U., Suckling, J., Zelaya, F., Honey, G., Faessel, H., Williams, S. C. R., et al. (2005). Plasma level-dependent effects of methylphenidate on task-related functional magnetic resonance imaging signal changes. Psychopharmacology, 180(4), 624-633.	No IIV
507	Munte, T. F., Heinze, H. J., Kunkel, H., & Scholz, M. (1984). Personality traits influence the effects of diazepam and caffeine on CNV magnitude. Neuropsychobiology, 12(1), 60-67.	No IIV
508	Murata, A. (2005). An attempt to evaluate mental workload using wavelet transform of EEG. Human Factors, 47(3), 498-508.	Wrong task
509	Murian, A., Deschamps, T., & Temprado, J. J. (2008). Effects of force production and trial duration on bimanual performance and attentional demands in a rhythmic coordination task. Motor Control, 12(1), 21-37.	No IIV
510	Murian, A., Deschamps, T., Bourbousson, J., & Temprado, J. J. (2008). Influence of an exhausting muscle exercise on bimanual coordination stability and attentional demands. Neuroscience Letters, 432(1), 64-68.	Wrong task
511	Myerson, J., & Hale, S. (Eds.). (1993). General slowing and age invariance in cognitive processing: The other side of the coin. San Diego, CA,US: Academic Press.	No empirical data
512	Myerson, J., Robertson, S., & Hale, S. (2007). Aging and intraindividual variability in performance: Analyses of response time distributions. Journal of the Experimental Analysis of Behavior, 88(3), 319-337.	Wrong task
513	Nacmias, B., Bessi, V., Bagnoli, S., Tedde, A., Cellini, E., Piccini, C., et al. (2008). KIBRA gene variants are associated with episodic memory performance in subjective memory complaints. Neuroscience Letters, 436(2), 145-147.	No RT
514	Nacmias, B., Piccini, C., Bagnoli, S., Tedde, A., Cellini, E., Bracco, L., et al. (2004). Brain-derived neurotrophic factor, apolipoprotein E genetic variants and cognitive performance in Alzheimer's disease. Neuroscience Letters, 367(3), 379-383.	No RT
515	Nakano, T., Araki, K., Michimori, A., Inbe, H., Hagiwara, H., & Koyama, E. (2000). Temporal order of sleepiness, performance and physiological indices during 19-h sleep deprivation. Psychiatry and Clinical Neurosciences, 54(3), 280-282.	No IIV
516	Namihira, T., Hirayasu, Y., & Koga, Y. (2004). The assessment of cognitive function in a Williams syndrome patient: A case report. Psychiatry and Clinical Neurosciences, 58(1), 99.	No RT
517	Naruse, K., Sakuma, H., & Hirai, T. (2004). Effect of slow movement execution on cognitive function. Perceptual and Motor Skills, 98(1), 35-43.	Matter not relevant
518	Nash, E. L., Phelan, J. G., Demas, G., & Bittner, A. (1966). Effects of manifest and induced anxiety and experimenter variability on simple reaction time. Perceptual and Motor Skills, 22(2), 483-487.	No IIV
519	Navalta, C. P., Polcari, A., Webster, D. M., Boghossian, A., & Teicher, M. H. (2006). Effects of childhood sexual abuse on neuropsychological and cognitive function in college women. Journal of Neuropsychiatry & Clinical Neurosciences, 18(1), 45-53.	Wrong task
520	Neargarder, S. A., Carvalho, J. O., & Sullivan, K. D. (2007). Neuropsychologic profile of a high-functioning family with a mitochondrial cytopathy. Cognitive & Behavioral Neurology, 20(3), 193-201.	No RT
521	Nederhof, E., Lemmink, K., Visscher, C., Meeusen, R., & Mulder, T. (2006). Psychomotor speed - Possibly a new marker for overtraining syndrome. Sports Medicine, 36(10), 817-828.	No empirical data
522	Nelson, C. A., Collins, P. F., & Torres, F. (1991). P300 Brain Activity in Seizure Patients Preceding Temporal Lobectomy. Archives of Neurology, 48(2), 141-147.	Wrong task
523	Nesselroade, J. R., & Salthouse, T. A. (2004). Methodological and theoretical implications of intraindividual variability in perceptual-motor performance. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 59(2), P49-55.	No RT
524	Nettelbeck, T., & Wilson, C. (2005). Uncertainty about the biology of intelligence: A role for a marker task. Cortex, 41(2), 234-235.	No empirical data
525	Neubauer, A. C., Riemann, R., Mayer, R., & Angleitner, A. (1997). Intelligence and reaction times in the Hick, Sternberg and Posner paradigms. Personality and Individual Differences, 22(6), 885-894.	Matter not relevant
526	Neumann, C. S., & Walker, E. F. (1999). Motor dysfunction in schizotypal personality disorder. Schizophrenia Research, 38(2-3), 159-168.	Wrong task
527	Ng, B., Camacho, A., Lara, D. R., Brunstein, M. G., Pinto, O. C., & Akiskal, H. S. (2008). A case series on the hypothesized connection between dementia and bipolar spectrum disorders: Bipolar type VI? Journal of Affective Disorders, 107(1-3), 307-315.	No RT
528	Nielen, M. M. A., & Den Boer, J. A. (2003). Neuropsychological performances of OCD patients before and after treatment with fluoxetine: Evidence for persistent cognitive deficits. Psychological Medicine, 33(5), 917-925.	Wrong task
529	Nigg, J. T., Blaskey, L. G., Stawicki, J. A., & Sachek, J. (2004). Evaluating the Endophenotype Model of ADHD Neuropsychological Deficit: Results for Parents and Siblings of Children With ADHD Combined and Inattentive Subtypes. Journal of Abnormal Psychology, 113(4), 614-625.	Wrong task
530	Nio, Y. K., Jansonius, N. M., Lamers, P., Mager, A., Zeinstra, J., & Kooijman, A. C. (2005). Influence of the rate of contrast change on the quality of contrast sensitivity assessment: a comparison of three psychophysical methods. Ophthalmic and Physiological Optics, 25(1), 18-26.	Wrong task
531	Nishimura, A., & Yokosawa, K. (2007). Preparation for horizontal or vertical dimensions affects the right-left prevalence effect. Perception & Psychophysics, 69(7), pp.	Matter not relevant
532	Noble, C. E., Baker, B. L., & Jones, T. A. (1964). Age and sex parameters in psychomotor learning. Perceptual and Motor Skills, 19(3), 935-945.	No IIV

533	Norrie, M. L. (1967). Effect of practice on the true score and intra-individual variability for reaction and movement times for simple and complex movements. Research Quarterly, 38(3), 457-467.	Matter not relevant
534	Norris, S. L., Lee, C.-T., Burshteyn, D., & Cea-Aravena, J. (2001). The effects of Performance Enhancement Training on hypertension, human attention, stress, and brain wave patterns: A case study. Journal of Neurotherapy, 4(3), 29-44.	Wrong task
535	Norris, S. L., Lee, C.-t., Cea, J., & Burshteyn, D. (1998). Performance enhancement training effects on attention: A case study. Journal of Neurotherapy, 3(1), 19-25.	Wrong task
536	O'Connell, R. G., Bellgrove, M. A., Dockree, P. M., Lau, A., Fitzgerald, M., & Robertson, I. H. (2008). Self-Alert Training: Volitional modulation of autonomic arousal improves sustained attention. Neuropsychologia, 46(5), 1379-1390.	Wrong task
537	Olafsdottir, H., Yoshida, N., Zatsiorsky, V. M., & Latash, M. L. (2007). Elderly show decreased adjustments of motor synergies in preparation to action. Clinical Biomechanics, 22(1), 44-51.	No IIV
538	Olde Rikkert, M. G. M., Rigaud, A. S., van Hoeyweghen, R. J., & de Graaf, J. (2003). Geriatric syndromes: medical misnomer or progress in geriatrics? Netherlands Journal of Medicine, 61(3), 83-87.	No empirical data
539	Ownby, R. L., Loewenstein, D. A., Schram, L., & Acevedo, A. (2004). Assessing the cognitive abilities that differentiate patients with Alzheimer's disease from normals: single and multiple factor models. International Journal of Geriatric Psychiatry, 19(3), 232-242.	No RT
540	Paleacu, D., Shutzman, A., Giladi, N., Herman, T., Simon, E. S., & Hausdorff, J. M. (2007). Effects of pharmacological therapy on gait and cognitive function in depressed patients. Clinical Neuropharmacology, 30(2), 63-71.	Wrong task
541	Palmer, B. W., Dunn, L. B., Appelbaum, P. S., & Jeste, D. V. (2004). Correlates of treatment-related decision-making capacity among middle-aged and older patients with schizophrenia. Archives of General Psychiatry, 61(3), 230-236.	No RT
542	Panzeri, S., Biella, G., Rolls, E. T., Skaggs, W. E., & Treves, A. (1996). Speed, noise, information and the graded nature of neuronal responses. Network-Computation in Neural Systems, 7(2), 365-370.	No empirical data
543	Papiol, S., Molina, V., Rosa, A., Sanz, J., Palomo, T., & Fananas, L. (2007). Effect of interleukin-1beta gene functional polymorphism on dorsolateral prefrontal cortex activity in schizophrenic patients. American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics, 144(8), 1090-1093.	Wrong task
544	Parmenter, B. A., Denney, D. R., & Lynch, S. G. (2003). The cognitive performance of patients with multiple sclerosis during periods of high and low fatigue. Multiple Sclerosis, 9(2), 111-118.	No RT
545	Parrott, A. (2003). Cognitive deficits and cognitive normality in recreational cannabis and Ecstacy/MDMA users. Human Psychopharmacology, 18(2), 89-90.	No empirical data
546	Parsons, O. A. (1987). Intellectual impairment in alcoholics: persistent issues. Acta Medica Scandinavica - Supplementum, 717, 33-46.	No RT
547	Patterson, J. V., Jin, Y., Gierczak, M., Hetrick, W. P., Potkin, S., Bunney, W. E., et al. (2000). Effects of temporal variability on P50 and the gating ratio in schizophrenia - A frequency domain adaptive filter single-trial analysis. Archives of General Psychiatry, 57(1), 57-64.	No RT
548	Patterson, J. V., Michalewski, H. J., & Starr, A. (1988). Latency variability of the components of auditory event-related potentials to infrequent stimuli in aging, Alzheimer-type dementia, and depression. Electroencephalography & Clinical Neurophysiology: Evoked Potentials, 71(6), 450-460.	Wrong task
549	Pattyn, N., Neyt, X., Heridericlcx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? Physiology & Behavior, 93(1-2), 369-378.	Wrong task
550	Paule, M. G., Rowland, A. S., Ferguson, S. A., Chelonis, J. J., Tannock, R., Swanson, J. M., et al. (2000). Attention deficit/hyperactivity disorder: characteristics, interventions and models. Neurotoxicology & Teratology, 22(5), 631-651.	No empirical data
551	Peeke, S. C., & Grings, W. W. (1968). Magnitude of UCR as a function of variability in the CS-UCS relationship. Journal of Experimental Psychology, 77(1), 64-69.	No RT
552	Pelosi, L., Holly, M., Slade, T., Hayward, M., & et al. (1992). Wave form variations in auditory event-related potentials evoked by a memory-scanning task and their relationship with tests of intellectual function. Electroencephalography & Clinical Neurophysiology: Evoked Potentials, 84(4), 344-352.	Wrong task
553	Perera, T. D., Luber, B., Nobler, M. S., Prudic, J., Anderson, C., & Sackeim, H. A. (2004). Seizure expression during electroconvulsive therapy: relationships with clinical outcome and cognitive side effects. Neuropsychopharmacology, 29(4), 813-825.	No RT
554	Perls, T. (2004). Centenarians who avoid dementia. Trends in Neurosciences, 27(10), 633-636.	No empirical data
555	Perry, R. J., Watson, P., & Hodges, J. R. (2000). The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: relationship to episodic and semantic memory impairment. Neuropsychologia, 38(3), 252-271.	No RT
556	Pfefferbaum, A., Ford, J. M., Roth, W. T., & Kopell, B. S. (1980). Age differences in P3-reaction time associations. Electroencephalography & Clinical Neurophysiology, 49(3-4), 257-265.	Wrong task
557	Philippsen, C., Hahn, M., Schwabe, L., Richter, S., Drewe, J., & Schachinger, H. (2007). Cardiovascular reactivity to mental stress is not affected by alpha2-adrenoreceptor activation or inhibition. Psychopharmacology, 190(2), 181-188.	No IIV
558	Pinelli, P. (2007). Trials and training with 100 self-rated verbal reactions and a theory of sequences. Functional Neurology, 22(1), 29-37.	Wrong task
559	Plattsmier, H. S., & McFadden, D. (1988). Temporary threshold shift measured with two psychophysical procedures. Audiology, 27(6), 334-343.	Matter not relevant
560	Plomin, R., Pedersen, N. L., Lichtenstein, P., & McClearn, G. E. (1994). Variability and Stability in Cognitive-Abilities Are Largely Genetic Later in Life. Behavior Genetics, 24(3), 207-215.	No RT
561	Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. Acta Psychologica, 86(2), 227-272.	No empirical data

562	Plummer-D'Amato, P., Altmann, L. J. P., Saracino, D., Fox, E., Behrman, A. L., & Marsiske, M. (2008). Interactions between cognitive tasks and gait after stroke: A dual task study. Gait & Posture, 27(4), 683-688.	Wrong task
563	Polich, J., & Gloria, R. (2001). Cognitive effects of a Ginkgo biloba/vinpocetine compound in normal adults: Systematic assessment of perception, attention and memory. Human Psychopharmacology, 16(5), 409-416.	No IIV
564	Pollak, M. H. (1991). Heart-Rate Reactivity to Laboratory Tasks and Ambulatory Heart-Rate in Daily Life. Psychosomatic Medicine, 53(1), 25-35.	No IIV
565	Pomerantz, J. M. (2001). Remarkable findings from brain imaging studies. Drug Benefit Trends, 13(1), 2BH+11BH.	No empirical data
566	Porges, S. W. (1973). Heart-Rate Variability - Autonomic Correlate of Reaction-Time Performance. Bulletin of the Psychonomic Society, 1(4), 270-272.	No IIV
567	Porter, R. J., Bourke, C., & Gallagher, P. (2007). Neuropsychological impairment in major depression: its nature, origin and clinical significance. Australian & New Zealand Journal of Psychiatry, 41(2), 115-128.	No empirical data
568	Price, G. W. (2000). Interactive ERP recording increases the amplitude of the endogenous P300 peak in schizophrenia. Schizophrenia Research, 41(3), 463-472.	No RT
569	Proctor, E. K., Morrow-Howell, N., Chadiha, L., Braverman, A. C., Darkwa, O., & Dore, P. (1997). Physical and cognitive functioning among chronically ill African-American and white elderly in home care following hospital discharge. Medical Care, 35(8), 782-791.	No RT
570	Purdon, S. E., Vallakalayil, A., Hanstock, C. C., Seres, P., & Tibbo, P. (2008). Elevated 3T proton MRS glutamate levels associated with poor Continuous Performance Test (CPT-0X) scores and genetic risk for schizophrenia. Schizophrenia Research, 99(1-3), 218-224.	Wrong task
571	Putman, J. A., Othmer, S. F., Othmer, S., & Pollock, V. E. (2005). TOVA Results Following Inter-Hemispheric Bipolar EEG Training. Journal of Neurotherapy, 9(1), 37-52.	Wrong task
572	Rabbitt, P. M. A., McInnes, L., Diggle, P., Holland, F., Bent, N., Abson, V., et al. (2004). The University of Manchester longitudinal study of cognition in normal healthy old age, 1983 through 2003. Aging, Neuropsychology, and Cognition, 11(2), 245-279.	No RT
573	Rabbitt, P., Clancy, M., & Vyas, S. (1977). After-effects of responding and of withholding responses in CRT tasks. The Quarterly Journal of Experimental Psychology, 29(3), 425-436.	Matter not relevant
574	Rabbitt, P., Osman, P., Moore, B., & Stollery, B. (2001). There are stable individual differences in performance variability, both from moment to moment and from day to day. The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 54(4), 981-1003.	Wrong task
575	Rakheja, S., Romero, J. A., Lozano, A., Liu, P. J., & Ahmed, A. K. W. (2002). Assessment of open-loop rollover control of articulated vehicles under different manoeuvres. Heavy Vehicle Systems-International Journal of Vehicle Design, 9(3), 204-222.	No participants - simulated data
576	Rakitin, B. C. (2005). The Effects of Spatial Stimulus-Response Compatibility on Choice Time Production Accuracy and Variability. Journal of Experimental Psychology: Human Perception and Performance, 31(4), 685-702.	Matter not relevant
577	Ramaekers, J. G., Muntjewerff, N. D., Uiterwijk, M. M. C., van Veggel, L. M. A., & et al. (1996). A study of the pharmacodynamic interaction between befloxatone and ethanol on performance and mood in healthy volunteers. Journal of Psychopharmacology, 10(4), 288-294.	No IIV
578	Randall, D. C., Viswanath, A., Bharania, P., Elsabagh, S. M., Hartley, D. E., Shneerson, J. M., et al. (2005). Does modafinil enhance cognitive performance in young volunteers who are not sleep-deprived? Journal of Clinical Psychopharmacology, 25(2), 175-179.	No IIV
579	Ranjitkar, P., Nakatsuji, T., Azuta, Y., & Gurusinghe, G. S. (2003). Stability analysis based on instantaneous driving behavior using car-following data. In Traffic Flow Theory and Highway Capacity 2003 (pp. 140-151).	Wrong task
580	Rao, S. L., Mukundan, C. R., Jamuna, N., Das, B. S., Shastry, K. V. R., Hegde, T., et al. (1997). Patterns of association between symptoms and neuropsychological deficits in post traumatic syndrome. Nimhans Journal, 15(2), 157-167.	No IIV
581	Rapp, M. A., Schnaider-Beeri, M., Sano, M., Silverman, J. M., & Haroutunian, V. (2005). Cross-Domain Variability of Cognitive Performance in Very Old Nursing Home Residents and Community Dwellers: Relationship to Functional Status. Gerontology, 51(3), 206-212.	No RT
582	Rasmussen, L. S., Siersma, V. D., & Ispocd, G. (2004). Postoperative cognitive dysfunction: true deterioration versus random variation. Acta Anaesthesiologica Scandinavica, 48(9), 1137-1143.	No RT
583	Ratcliff, R. (2006). Modeling response signal and response time data. Cognitive Psychology, 53(3), 195-237.	Wrong task
584	Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. Psychonomic Bulletin & Review, 9(3), 438-481.	No empirical data
585	Ratcliff, R., Schmiedek, F., & McKoon, G. (2008). A diffusion model explanation of the worst performance rule for reaction tune and IQ. Intelligence, 36(1), 10-17.	No participants - simulated data
586	Ratcliff, R., Spieler, D., & McKoon, G. (2004). Analysis of group differences in processing speed: Where are the models of processing? Psychonomic Bulletin & Review, 11(4), 755-769.	No participants - simulated data
587	Rauch, S. L., Wedig, M. M., Wright, C. I., Martis, B., McMullin, K. G., Shin, L. M., et al. (2007). Functional Magnetic Resonance Imaging Study of Regional Brain Activation During Implicit Sequence Learning in Obsessive-Compulsive Disorder. Biological Psychiatry, 61(3), 330-336.	Wrong task

588	Ravdin, L. D., Barr, W. B., Jordan, B., Lathan, W. E., & Relkin, N. R. (2003). Assessment of cognitive recovery following sports related head trauma in boxers. Clinical Journal of Sport Medicine, 13(1), 21-27.	No RT
589	Ray, E. C., Engum, E. S., Lambert, E. W., Bane, G. F., Nash, M. R., & Bracy, O. L. (1997). Ability of the cognitive behavioral driver's inventory to distinguish malingerers from brain-damaged subjects. Archives of Clinical Neuropsychology, 12(5), 491-503.	No IIV
590	Reed, T. E. (1998). Causes of intraindividual variability in reaction times: A neurophysiologically oriented review and a new suggestion. Personality and Individual Differences, 25(5), 991-998.	No empirical data
591	Register-Mihalik, J., Guskiewicz, K. M., Mann, J. D., & Shields, E. W. (2007). The effects of headache on clinical measures of neurocognitive function. Clinical Journal of Sport Medicine, 17(4), 282-288.	Sample consisted of teenagers
592	Reimers, S., & Maylor, E. A. (2006). Gender Effects on Reaction Time Variability and Trial-to-Trial Performance: Reply to Deary and Der (2005). Aging, Neuropsychology, and Cognition, 13(3), 479-489	Wrong task
593	Reinberg, A. E., Ashkenazi, I., & Smolensky, M. H. (2007). Euchronism, allochronism, and dyschronism: Is internal desynchronization of human circadian rhythms a sign of illness? Chronobiology International, 24(4), pp.	No empirical data
594	Reinvang, I., Espeseth, T., & Gjerstad, L. (2005). Cognitive ERPs are related to ApoE allelic variation in mildly cognitively impaired patients. Neuroscience Letters, 382(3), 346-351.	No RT
595	Reisberg, B., Franssen, E. H., Souren, L. E. M., Auer, S., & Kenowsky, S. (1998). Progression of Alzheimer's disease: variability and consistency: ontogenic models, their applicability and relevance. Journal of Neural Transmission, 9-20.	No empirical data
596	Rekkas, P. V. (2006). Interference Resolution in the Elderly: Evidence Suggestive of Differences in Strategy on Measures of Prepotent Inhibition and Dual Task Processing. Aging, Neuropsychology, and Cognition, 13(3), 341-365.	No IIV
597	Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. Journal of Motor Behavior, 35(4), 355-370.	No RT
598	Reynolds, C. A., Finkel, D., McArdle, J. J., Gatz, M., Berg, S., & Pedersen, N. L. (2005). Quantitative genetic analysis of latent growth curve models of cognitive abilities in adulthood. Developmental Psychology, 41(1), 3-16.	No RT
599	Reynolds, J. D., Amory, D. W., Grocott, H. P., White, W. D., & Newman, M. F. (2002). Change in plasma glutamate concentration during cardiac surgery is a poor predictor of cognitive outcome. Journal of Cardiothoracic & Vascular Anesthesia, 16(4), 431-436.	No RT
600	Richards, M., Sano, M., Goldstein, S., Mindry, D., Todak, G., & Stern, Y. (1992). The stability of neuropsychological test performance in a group of parenteral drug users. Journal of Substance Abuse Treatment, 9(4), 371-377.	No RT
601	Richer, F. o., & Boulet, C. (1999). Frontal lesions and fluctuations in response preparation. Brain and Cognition, 40(1), 234-238.	Wrong task
602	Rikli, R. E. (2000). Reliability, validity, and methodological issues in assessing physical activity in older adults. Research Quarterly for Exercise and Sport, 71(2), S89-S96.	No empirical data
603	Rist, F., & Thurm, I. (1984). Effects of intramodal and crossmodal stimulus diversity on the reaction time of chronic schizophrenics. Journal of Abnormal Psychology, 93(3), 331-338.	No IIV
604	Ritchie, K., Ledesert, B., & Touchon, J. (1993). The Eugeria Study of Cognitive Aging - Who Are the Normal Elderly. International Journal of Geriatric Psychiatry, 8(12), 969-977.	No IIV
605	Ritter, W., Rotkin, L., & Vaughan, H. G. (1980). The modality specificity of the slow negative wave. Psychophysiology, 17(3), 222-227.	No IIV
606	Robbe, H. (1998). Marijuana's impairing effects on driving are moderate when taken alone but severe when combined with alcohol. Human Psychopharmacology-Clinical and Experimental, 13, S70-S78.	No IIV
607	Roberts, R. D. (1997). Fitts' law, movement time and intelligence. Personality and Individual Differences, 23(2), 227-246.	No RT
608	Robertson, S., Myerson, J., & Hale, S. (2006). Are There Age Differences in Intraindividual Variability in Working Memory Performance? Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 61(1), P18-P24.	Wrong task
609	Robichaud, J. A., & Corcos, D. M. (2005). Motor deficits, exercise, and Parkinson's disease. Quest, 57(1), 79-101.	No empirical data
610	Robichaud, J. A., Pfann, K. D., Vaillancourt, D. E., Comella, C. L., & Corcos, D. M. (2005). Force control and disease severity in Parkinson's disease. Movement Disorders, 20(4), 441-450.	No RT
611	Robinson, M. D., & Tamir, M. (2005). Neuroticism as Mental Noise: A Relation Between Neuroticism and Reaction Time Standard Deviations. Journal of Personality and Social Psychology, 89(1), 107-114.	Wrong task
612	Robinson, M. D., Wilkowski, B. M., & Meier, B. P. (2006). Unstable in More Ways Than One: Reaction Time Variability and the Neuroticism/Distress Relationship. Journal of Personality, 74(2), 311-343.	Wrong task
613	Rockwood, K. (2002). Natural history of vascular cognitive impairment and its subtypes. CPD Bulletin Old Age Psychiatry, 3(2), 38-40.	No empirical data
614	Rockwood, K., & Middleton, L. (2007). Physical activity and the maintenance of cognitive function. Alzheimer's and Dementia, 3(2 SUPPL.), S38-S44.	No empirical data
615	Roiser, J. P., McLean, A., Ogilvie, A. D., Blackwell, A. D., Bamber, D. J., Goodyer, I., et al. (2005). The subjective and cognitive effects of acute phenylalanine and tyrosine depletion in patients recovered from depression. Neuropsychopharmacology, 30(4), 775-785.	Wrong task
616	Romero, L. J., Schuyler, M., Kamboh, M. I., Qualls, C., LaRue, A., Liang, H. C., et al. (2002). The APO E4 allele and cognition in New Mexico Hispanic elderly. Ethnicity & Disease, 12(2), 235-241.	No RT

617	Ropacki, S. A., Bert, A. A., Ropacki, M. T., Rogers, B. L., & Stern, R. A. (2007). The influence of cognitive reserve on neuropsychological functioning following coronary artery bypass grafting (CABG). Archives of Clinical Neuropsychology, 22(1), 73-85.	No RT
618	Rosen, B., Lundborg, G., Dahlin, L. B., Holmberg, J., & Karlson, B. (1994). Nerve repair: Correlation of restitution of functional sensibility with specific cognitive capacities. Journal of Hand Surgery, 19 B(4), 452-458.	No IIV
619	Ross, J. L., Stefanatos, G. A., Kushner, H., Bondy, C., Nelson, L., Zinn, A., et al. (2004). The effect of genetic differences and ovarian failure: Intact cognitive function in adult women with premature ovarian failure versus turner syndrome. Journal of Clinical Endocrinology and Metabolism, 89(4), 1817-1822.	No RT
620	Ross, J., Zinn, A., & McCauley, E. (2000). Neurodevelopmental and psychosocial aspects of Turner syndrome. Mental Retardation & Developmental Disabilities Research Reviews, 6(2), 135-141.	No empirical data
621	Rossetti, Y., Revol, P., McIntosh, R., Pisella, L., Rode, G., Danckert, J., et al. (2005). Visually guided reaching: bilateral posterior parietal lesions cause a switch from fast visuomotor to slow cognitive control. Neuropsychologia, 43(2), 162-177.	No RT
622	Rossit, S., & Harvey, M. (2008). Age-related differences in corrected and inhibited pointing movements. Experimental Brain Research, 185(1), 1-10.	No IIV
623	Rossiter, T. (2004). The Effectiveness of Neurofeedback and Stimulant Drugs in Treating AD/HD: Part II Replication. Applied Psychophysiology and Biofeedback, 29(4), 233-243.	Wrong task
624	Rossiter, T. R. (1998). Patient-directed neurofeedback for AD/HD. Journal of Neurotherapy, 2(4), 54-63.	Wrong task
625	Rossy, L. A., & Thayer, J. F. (1998). Fitness and gender-related differences in heart period variability. Psychosomatic Medicine, 60(6), 773-781.	Wrong task
626	Roth, T., White, D., Schmidt-Nowara, W., Wesnes, K. A., Niebler, G., Arora, S., et al. (2006). Effects of Armodafinil in the Treatment of Residual Excessive Sleepiness Associated with Obstructive Sleep Apnea/Hypopnea Syndrome: A 12-Week, Multicenter, Double-Blind, Randomized, Placebo-Controlled Study in nCPAP-Adherent Adults. Clinical Therapeutics, 28(5), 689-706.	No IIV
627	Rothwell, J. C., Traub, M. M., & Marsden, C. D. (1982). Automatic and 'voluntary' responses compensating for disturbances of human thumb movements. Brain Research, 248(1), 33-41.	No IIV
628	Royall, D. R., Chiodo, L. K., Mouton, C., & Polk, M. J. (2007). Cognitive predictors of mortality in elderly retirees: Results from the Freedom House Study. American Journal of Geriatric Psychiatry, 15(3), 243-251.	No RT
629	Royall, D. R., Palmer, R., Chiodo, L. K., & Polk, M. J. (2005a). Executive Control Mediates Memory's Association with Change in Instrumental Activities of Daily Living: The Freedom House Study. Journal of the American Geriatrics Society, 53(1), 11-17.	No RT
630	Royall, D. R., Palmer, R., Chiodo, L. K., & Polk, M. J. (2005b). Normal rates of cognitive change in successful aging: The Freedom House study. Journal of the International Neuropsychological Society, 11(7), 899-909.	No RT
631	Rupp, T. L., Acebo, C., Seifer, R., & Carskadon, M. A. (2007). Effects of a moderate evening alcohol dose II: Performance. Alcoholism: Clinical and Experimental Research, 31(8), 1365-1371.	No IIV
632	Russell, V. A., Oades, R. D., Tannock, R., Killeen, P. R., Auerbach, J. G., Johansen, E. B., et al. (2006). Response variability in Attention-Deficit/Hyperactivity Disorder: A neuronal and glial energetics hypothesis. Behavioral and Brain Functions, 2(30).	No empirical data
633	Ruzicka, E., & Elmassioui, F. (1993). Event-Related Potentials in Parkinsons-Disease - a Review. Behavioural Neurology, 6(1), 15-26.	No empirical data
634	SÃ©vigny, M.-C., Everett, J., & Grondin, S. (2003). Depression, attention, and time estimation. Brain and Cognition, 53(2), 351-353.	Wrong task
635	Sacco, K. A., Termine, A., Dudas, M. M., Seyal, A. A., Allen, T. M., Vessicchio, J. C., et al. (2006). Neuropsychological deficits in nonsmokers with schizophrenia: Effects of a nicotinic antagonist. Schizophrenia Research, 85(1-3), 213-221.	Wrong task
636	Saccuzzo, D. P., Larson, G. E., & Rimland, B. (1986). Visual, auditory and reaction time approaches to the measurement of speed of information processing and individual differences in intelligence. Personality and Individual Differences, 7(5), 659-667.	No IIV
637	Saletu, B. (1977). The evoked potential in pharmacopsychiatry. Neuropsychobiology, 3(2-3), 75-104.	No empirical data
638	Saletu, B., & GrÃ¼nberger, J. (1983). Cerebral hypoxic hypoxidosis: Neurophysiological, psychometric and pharmacotherapeutic aspects. Advances in Biological Psychiatry, 13, 146-164.	No IIV
639	Saletu, B., Grunberger, J., Saletu, M., Mader, R., & Volavka, J. (1978). Treatment of the alcoholic organic brain syndrome with EMD 21657. A derivative of a pyritinolmetabolite: double-blind clinical, quantitative EEG and psychometric studies. International Pharmacopsychiatry, 13(3), 177-192.	No IIV
640	Salmon, D., & Hodges, J. R. (2005). Introduction: Mild cognitive impairment - Cognitive, behavioral, and biological factors. Neurocase, 11(1), 1-2.	No empirical data
641	Salmoni, A. W., & McIlwain, J. S. (1979). Fitts' reciprocal tapping task, a measure of motor capacity? Perceptual & Motor Skills, 49(2), 403-413.	No RT
642	Salthouse, T. A. (1998). Relation of successive percentiles of reaction time distributions to cognitive variables and adult age. Intelligence, 26(2), 153-166.	Wrong task
643	Salthouse, T. A. (2004). What and when of cognitive aging. Current Directions in Psychological Science, 13(4), 140-144.	No empirical data
644	Salthouse, T. A. (2007). Implications of within-person variability in cognitive and neuropsychological functioning for the interpretation of change. Neuropsychology, 21(4), 401-411.	Wrong task
645	Salthouse, T. A., & Berish, D. E. (2005). Correlates of Within-Person (Across-Occasion) Variability in Reaction Time. Neuropsychology, 19(1), 77-87.	Wrong IIIV

646	Salthouse, T. A., Fristoe, N. M., Lineweaver, T. T., & Coon, V. E. (1995). Aging of attention: does the ability to divide decline? Memory & Cognition, 23(1), 59-71.	Wrong task
647	Salthouse, T. A., Nesselroade, J. R., & Berish, D. E. (2006). Short-term variability in cognitive performance and the calibration of longitudinal change. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 61(3), P144-151.	No RT
648	Sanders, R. D., Schuepbach, D., Goldstein, G., Haas, G. L., Sweeney, J. A., & Keshavan, M. S. (2004). Relationships between cognitive and neurological performance in neuroleptic-naive psychosis. Journal of Neuropsychiatry & Clinical Neurosciences, 16(4), 480-487.	No RT
649	Sandman, C. A., & Patterson, J. V. (2000). The auditory event-related potential is a stable and reliable measure in elderly subjects over a 3 year period. Clinical Neurophysiology, 111(8), 1427-1437.	Wrong task
650	Sannita, W. G., Waldeier, H., & Dolce, G. (1979). Individual variability in the effects of alcohol on the contingent negative variation (CNV) in human subjects. Research Communications in Psychology, Psychiatry & Behavior, 4(2), 97-108.	No IIV
651	Satava, R. M. (2005). Mental workload: A new parameter for objective assessment? Surgical Innovation, 12(1), 79.	No empirical data
652	Sauerwein, H. C., & Lassonde, M. (1994). Cognitive and sensori-motor functioning in the absence of the corpus callosum: neuropsychological studies in callosal agenesis and callosotomized patients. Behavioural Brain Research, 64(1-2), 229-240.	No IIV
653	Savage, C. R., & Thomas, D. G. (1993). Information-Processing and Interhemispheric-Transfer in Left-Handed and Right-Handed Adults. International Journal of Neuroscience, 71(1-4), 201-219.	No IIV
654	Savettieri, G., Messina, D., Andreoli, V., Bonavita, S., Caltagirone, C., Cittadella, R., et al. (2004). Gender-related effect of clinical and genetic variables on the cognitive impairment in multiple sclerosis. Journal of Neurology, 251(10), 1208-1214.	No RT
655	Savitz, J., Van Der Merwe, L., Solms, M., & Ramesar, R. (2007). A linkage and family-based association analysis of a potential neurocognitive endophenotype of bipolar disorder. NeuroMolecular Medicine, 9(2), 101-116.	No RT
656	Savoy, R. L. (2006). Using small of subjects numbers in fMRI-based research - Searching for universal truths in human brain function and organization. Ieee Engineering in Medicine and Biology Magazine, 25(2), 52-59.	No empirical data
657	Sawyer, E., Mauro, L. S., & Ohlinger, M. J. (2008). Amantadine enhancement of arousal and cognition after traumatic brain injury. Annals of Pharmacotherapy, 42(2), 247-252.	No empirical data
658	Schaie, K. W. (1988). Variability in cognitive function in the elderly: implications for societal participation. Basic Life Sciences, 43, 191-211.	No RT
659	Schall, J. D. (2002). Decision-making: Neural correlates of response time. Current Biology, 12(23), R800-R801.	No empirical data
660	Schall, J. D., & Bichot, N. P. (1998). Neural correlates of visual and motor decision processes. Current Opinion in Neurobiology, 8(2), 211-217.	No empirical data
661	Schaper, K., Kolsch, H., Popp, J., Wagner, M., & Jessen, F. (2008). KIBRA gene variants are associated with episodic memory in healthy elderly. Neurobiology of Aging, 29(7), 1123-1125.	No RT
662	Schellekens, J. M. H., Sijtsma, G. J., Vegter, E., & Meijman, T. F. (2000). Immediate and delayed after-effects of long lasting mentally demanding work. Biological Psychology, 53(1), 37-56.	Wrong task
663	Schmaling, K. B., DiClementi, J. D., Cullum, C. M., & Jones, J. F. (1994). Cognitive functioning in chronic fatigue syndrome and depression: a preliminary comparison. Psychosomatic Medicine, 56(5), 383-388.	No RT
664	Schofield, W. N. (Ed.). (1987). Estimating cerebral asymmetry and interhemispheric transmission time from individual differences in bimanual response to lateralized stimuli. New York, NY,US: Plenum Press.	Sample consisted of children
665	Schretlen, D. J., Munro, C. A., Anthony, J. C., & Pearlson, G. D. (2003). Examining the range of normal intraindividual variability in neuropsychological test performance. Journal of the International Neuropsychological Society, 9(6), 864-870.	Wrong task
666	Schroder, J., Essig, M., Baudendistel, K., Jahn, T., Gerdsen, I., Stockert, A., et al. (1999). Motor dysfunction and sensorimotor cortex activation changes in schizophrenia: A study with functional magnetic resonance imaging. Neuroimage, 9(1), 81-87.	No RT
667	Schumacher, E. H., Hendricks, M. J., & D'Esposito, M. (2005). Sustained involvement of a frontal-parietal network for spatial response selection with practice of a spatial choice-reaction task. Neuropsychologia, 43(10), 1444-1455.	No IIV
668	Schwabe, L., Szinnai, G., Keller, U., & Schachinger, H. (2007). Dehydration does not influence cardiovascular reactivity to behavioural stress in young healthy humans. Clinical Physiology and Functional Imaging, 27(5), 291-297.	No IIV
669	Schwartz, J. R. L., Nelson, M. T., Schwartz, E. R., & Hughes, R. J. (2004). Effects of modafinil on wakefulness and executive function in patients with narcolepsy experiencing late-day sleepiness. Clinical Neuropharmacology, 27(2), 74-79.	No RT
670	Schweinberger, S. R., Buse, C., & Sommer, W. (1993). Reaction time improvements with practice in brain-damaged patients. Cortex, 29(2), 333-340.	Wrong task
671	Schweizer, K. (1993). The Contribution of Access to External Information, Stimulus Complexity, and Variability to Cognitive-Abilities. Personality and Individual Differences, 14(1), 87-95.	Wrong task
672	Seeber, A., Schaper, M., Zupanic, M., Blaszkewicz, M., Demes, P., Kiesswetter, E., et al. (2004). Toluene exposure below 50 ppm and cognitive function: a follow-up study with four repeated measurements in rotogravure printing plants. International Archives of Occupational & Environmental Health, 77(1), 1-9.	No IIV

673	Segalowitz, S. J., Dywan, J., & Unsal, A. (1997). Attentional factors in response time variability after traumatic brain injury: An ERP study. Journal of the International Neuropsychological Society, 3(2), 95-107.	Wrong task
674	Seidler, R. D., & Stelmach, G. E. (1995). Reduction in Sensorimotor Control with Age. Quest, 47(3), 386-394.	No empirical data
675	Seidler, R. D., & Stelmach, G. E. (2000). Trunk-assisted prehension: specification of body segments with imposed temporal constraints. Journal of Motor Behavior, 32(4), 379-389.	No RT
676	Sellal, F., Wolff, V., & Marescaux, C. (2004). The cognitive pattern of vascular dementia and its assessment. Seminars in Cerebrovascular Diseases and Stroke, 4(2 SPEC. ISS.), 79-86.	No RT
677	Selnes, O. A., & Zeger, S. L. (2007). Coronary Artery Bypass Grafting Baseline Cognitive Assessment: Essential Not Optional. Annals of Thoracic Surgery, 83(2), 374-376.	No empirical data
678	Selnes, O. A., Pham, L., Zeger, S., & McKhann, G. M. (2006). Defining cognitive change after CABG: decline versus normal variability. Annals of Thoracic Surgery, 82(2), 388-390.	Wrong task
679	Sengel, R. a., & Lovallo, W. R. (1980). A random process model of cognitive deficit in schizophrenia. Schizophrenia Bulletin, 6(3), 526-535.	No RT
680	Shagass, C., Straumanis, J. J., & Overton, D. A. (1972). Electrophysiological recordings in the reaction time experiment: Exploratory studies for possible psychiatric research application. Biological Psychiatry, 5(3), 271-287.	No IIV
681	Shah, J., Wesnes, K. A., Kovelesky, R. A., & Henney, H. R., 3rd. (2006). Effects of food on the single-dose pharmacokinetics/pharmacodynamics of tizanidine capsules and tablets in healthy volunteers. Clinical Therapeutics, 28(9), 1308-1317.	No IIV
682	Shalev, S. A., Chervinski, E., Weiner, E., Mazor, G., Friez, M. J., & Schwartz, C. E. (2006). Clinical variation of Aarskog syndrome in a large family with 2189delA in the FGD1 gene. American Journal of Medical Genetics, Part A. 140(2), 162-165.	No RT
683	Sheppard, J. M., & Young, W. B. (2006). Agility literature review: classifications, training and testing. Journal of Sports Sciences, 24(9), 919-932.	No empirical data
684	Sheridan, P. L., & Hausdorff, J. M. (2007). The role of higher-level cognitive function in gait: Executive dysfunction contributes to fall risk in Alzheimer's disease. Dementia and Geriatric Cognitive Disorders, 24(2), 125-137.	No empirical data
685	Shieh, K. K., Kuo, W. M., & Lai, C. J. (1997). Effects of duration of stimulus and variability of foreperiod on the identification of multidimensional stimuli. Perceptual and Motor Skills, 84(3), 1379-1393.	Matter not relevant
686	Sidaway, B., Sekiya, H., & Fairweather, M. (1995). Movement Variability as a Function of Accuracy Demand in Programmed Serial Aiming Responses. Journal of Motor Behavior, 27(1), 67-76.	Matter not relevant
687	Sidaway, B., Yook, D., & Russell, D. (1999). Distributed control in rapid sequential aiming responses. Journal of Motor Behavior, 31(4), 367-379.	Matter not relevant
688	Siepmann, M., Grossmann, J., MÃ¼ck-Weymann, M., & Kirch, W. (2003). Effects of sertraline on autonomic and cognitive functions in healthy volunteers. Psychopharmacology, 168(3), 293-298.	Wrong task
689	Siepmann, M., Handel, J., Mueck-Weymann, M., & Kirch, W. (2004). The effects of moclobemide on autonomic and cognitive functions in healthy volunteers. Pharmacopsychiatry, 37(2), 81-87.	Wrong task
690	Siepmann, M., Krause, S., Joraschky, P., Muck-Weymann, M., & Kirch, W. (2002). The effects of St John's wort extract on heart rate variability, cognitive function and quantitative EEG: a comparison with amitriptyline and placebo in healthy men. British Journal of Clinical Pharmacology, 54(3), 277-282.	Wrong task
691	Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. Plos Biology, 3(2), 334-349.	Wrong task
692	Simonen, R. L., Videman, T., BattiÃ©, M. C., & Gibbons, L. E. (1998). Determinants of psychomotor speed among 61 pairs of adult male monozygotic twins. Journals of Gerontology: Series A: Biological Sciences and Medical Sciences, 53(3), M228-M234.	No IIV
693	Sinotte, M. P., & Coelho, C. A. (2007). Attention training for reading impairment in mild aphasia: A follow-up study. NeuroRehabilitation, 22(4), 303-310.	No RT
694	Sirevaag, E. J., Rohrbaugh, J. W., Stern, J. A., Vedeniapin, A. B., Packingham, K. D., & LaJonchere, C. M. (1999). Multi-dimensional characterizations of operator state: A validation of oculomotor metrics. FAA Office of Aviation Medicine Reports(99), 1-29.	Wrong task
695	Skeel, R. L., Johnstone, B., Yangco, D. T., Jr., Walker, S. E., & Komatireddy, G. R. (2000). Neuropsychological deficit profiles in systemic lupus erythematosus. Applied Neuropsychology, 7(2), 96-101.	No RT
696	Slanska, J. (1978). Time estimation and reaction time in relation to the personality trait stability-lability. Activitas Nervosa Superior, 20(1), 36-37.	No IIV
697	Sleegers, K., de Koning, I., Aulchenko, Y. S., van Rijn, M. J. E., Houben, M. P. W. A., Croes, E. A., et al. (2007). Cerebrovascular risk factors do not contribute to genetic variance of cognitive function. The ERF study. Neurobiology of Aging, 28(5), 735-741.	No RT
698	Sliwinski, M. J., Smyth, J. M., Hofer, S. M., & Stawski, R. S. (2006). Intraindividual Coupling of Daily Stress and Cognition. Psychology and Aging, 21(3), 545-557.	Wrong task
699	Smeding, H. M. M., Speelman, J. D., Koning-Haanstra, M., Schuurman, P. R., Nijssen, P., van Laar, T., et al. (2006). Neuropsychological effects of bilateral STN stimulation in Parkinson disease: a controlled study.[see comment]. Neurology, 66(12), 1830-1836.	No RT
700	Smelson, D. A., Davis, C. W., Eisenstein, N., Engelhart, C., Williams, J., Losonczy, M. F., et al. (2003). Cognitive disparity in schizophrenics with and without cocaine dependency. Journal of Substance Abuse Treatment, 24(1), 75-79.	No RT

701	Smiley, A. (1986). Marijuana: On-road and driving simulator studies. Alcohol, Drugs & Driving, 2(3), 121-134.	No empirical data
702	Smiley-Oyen, A. L., Lowry, K. A., & Kerr, J. P. (2007). Planning and control of sequential rapid aiming in adults with Parkinson's disease. Journal of Motor Behavior, 39(2), 103-114.	No IIV
703	Smith, B. D., Osborne, A., Mann, M., Jones, H., & White, T. (Eds.). (2004). Arousal and behavior: Biopsychological effects of caffeine. Boca Raton, FL,US: CRC Press.	No empirical data
704	Smyrnis, N., Avramopoulos, D., Evdokimidis, I., Stefanis, C. N., Tsekou, H., & Stefanis, N. C. (2007). Effect of Schizotypy on Cognitive Performance and Its Tuning by COMT val¹⁵⁸ Met Genotype Variations in a Large Population of Young Men. Biological Psychiatry, 61(7), 845-853.	Wrong task
705	Socan, G., & Bucik, V. (1998). Relationship between speed of information-processing and two major personality dimensions - extraversion and neuroticism. Personality and Individual Differences, 25(1), 35-48.	Matter not relevant
706	Soloff, P. H. (1990). What's new in personality disorders?: An update on pharmacologic treatment. Journal of Personality Disorders, 4(3), 233-243.	No empirical data
707	Somsen, R., Jennings, J. R., & Van der Molen, M. (2002). Human vagal heart rate responses in warned reaction time tasks: A cross-validation with simulated vagus-sino-atrial node interactions. Journal of Electrocardiology, 35(SUPPL.), 231-237.	No IIV
708	Sonuga-Barke, E. J. S., & Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: A neurobiological hypothesis. Neuroscience & Biobehavioral Reviews, 31(7), 977-986.	No empirical data
709	Sosnoff, J. J., & Newell, K. M. (2006). The generalization of perceptual-motor intra-individual variability in young and old adults. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 61(5), P304-310.	No RT
710	Sosnoff, J. J., & Newell, K. M. (2007). Are visual feedback delays responsible for aging-related increases in force variability? Experimental Aging Research, 33(4), 399-415.	No RT
711	Sosnoff, J. J., Broglio, S. P., Hillman, C. H., & Ferrara, M. S. (2007). Concussion does not impact intraindividual response time variability. Neuropsychology, 21(6), 796-802.	Wrong task
712	Spencer, R. M. C., Ivry, R. B., & Zelaznik, H. N. (2005). Role of the cerebellum in movements: control of timing or movement transitions? Experimental Brain Research, 161(3), 383-396.	No IIV
713	Spoletini, I., Caltagirone, C., Bria, P., & Spalletta, G. (2006). ApolipoproteinE epsilon-4 allelic variant, cognitive decline and psychosis in Alzheimer disease: a review of the literature and suggestions for upcoming studies. European Journal of Psychiatry, 20(2), 74-87.	No empirical data
714	Springer, S., Giladi, N., Peretz, C., Yogev, G., Simon, E. S., & Hausdorff, J. M. (2006). Dual-tasking effects on gait variability: The role of aging, falls, and executive function. Movement Disorders, 21(7), 950-957.	Wrong task
715	Starbuck, V. N., Bleiberg, J., & Kay, G. G. (1995). D-Amphetamine-Mediated Enhancement of the P300 Erp - a Placebo-Crossover Double-Blind Case-Study. Neuropsychiatry Neuropsychology and Behavioral Neurology, 8(3), 189-192.	No RT
716	Starkes, J., Helsen, W., & Elliott, D. (2002). A menage a trois: the eye, the hand and on-line processing. Journal of Sports Sciences, 20(3), 217-224.	No empirical data
717	Starkstein, S. E., Sabe, L., Vazquez, S., Teson, A., Petracca, G., Chemerinski, E., et al. (1996). Neuropsychological, psychiatric, and cerebral blood flow findings in vascular dementia and Alzheimer's disease. Stroke, 27(3), 408-414.	No RT
718	Staude, G. H., Wolf, W. M., Appel, U., & Dengler, R. (1996). Methods for onset detection of voluntary motor responses in tremor patients. IEEE Transactions on Biomedical Engineering, 43(2), 177-188.	Wrong task
719	Staude, G., Wolf, W., Ott, M., Oertel, W. H., & Dengler, R. (1995). Tremor as a factor in prolonged reaction times of parkinsonian patients. Movement Disorders, 10(2), 153-162.	No IIV
720	Stefanis, N. C., van Os, J., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., & Stefanis, C. N. (2005). Effect of COMT ValÂ¹-sup-5-sup-8Met Polymorphism on the Continuous Performance Test, Identical Pairs Version: Tuning Rather Than Improving Performance. American Journal of Psychiatry, 162(9), 1752-1754.	Wrong task
721	Stein, A. C., Allen, R. W., & Cook, M. L. (1985). The interaction of alcohol and fatigue on driver simulator performance. Proceedings of the American Association for Automotive Medicine, 29, 91-104.	Wrong task
722	Stelmach, G. E., Teasdale, N., Phillips, J., & Worringham, C. J. (1989). Force production characteristics in Parkinson's disease. Experimental Brain Research, 76(1), 165-172.	Wrong task
723	Stelzel, C., Schumacher, E. H., Schubert, T., & D'Esposito, M. (2006). The neural effect of stimulus-response modality compatibility on dual-task performance: an fMRI study. Psychological Research, 70(6), 514-525.	No IIV
724	Stevens, A., Lutzenberger, W., Bartels, D. M., Strik, W., & Lindner, K. (1997). Increased duration and altered topography of EEG microstates during cognitive tasks in chronic schizophrenia. Psychiatry Research, 66(1), 45-57.	Wrong task
725	Stevens, J. R., Lonsbury, B., & Goel, S. (1972). Electroencephalographic spectra and reaction time in disorders of higher nervous function. Science, 176(4041), 1346-1349.	Wrong task
726	Stewart, R. M., Tunell, G., & Ehle, A. (1981). Familial spastic paraplegia, peroneal neuropathy, and crural hypopigmentation: a new neurocutaneous syndrome. Neurology, 31(6), 754-757.	No RT
727	Strauss, E., Slick, D. J., Levy-Bencheton, J., Hunter, M., MacDonald, S. W. S., & Hultsch, D. F. (2002). Intraindividual variability as an indicator of malingering in head injury. Archives of Clinical Neuropsychology, 17(5), 423-444.	Wrong task
728	Strobel, A., Dreisbach, G., Muller, J., Goschke, T., Brocke, B., & Lesch, K. P. (2007). Genetic variation of serotonin function and cognitive control. Journal of Cognitive Neuroscience, 19(12), 1923-1931.	Wrong task

729	Stuss, D. T., Pogue, J., Buckle, L., & Bondar, J. (1994). Characterization of stability of performance in patients with traumatic brain injury: Variability and consistency on reaction time tests. Neuropsychology, 8(3), 316-324.	Wrong task
730	Swanwick, G. R., Coen, R. F., Coakley, D., & Lawlor, B. A. (1998). Assessment of progression and prognosis in 'possible' and 'probable' Alzheimer's disease. International Journal of Geriatric Psychiatry, 13(5), 331-335.	No RT
731	Swift, C. G., Haythorne, J. M., Clarke, P., & Stevenson, I. H. (1981). Cardiovascular, sedative and anticholinergic effects of amitriptyline and zimelidine in young and elderly volunteers. Acta Psychiatrica Scandinavica, 63(Suppl. 290), 425-432.	No IIV
732	Swift, C. G., Swift, M. R., Ankier, S. I., Pidgen, A., & Robinson, J. (1985). Single dose pharmacokinetics and pharmacodynamics of oral loprazolam in the elderly. British Journal of Clinical Pharmacology, 20(2), 119-128.	No IIV
733	Sylvester, C. M., d'Avossa, G., & Corbetta, M. (2006). Models of human visual attention should consider trial-by-trial variability in preparatory neural signals. Neural Networks, 19(9), 1447-1449.	No empirical data
734	Szelag, E., Rymarczyk, K., & Poppel, E. (2001). Conscious control of movements: increase of temporal precision in voluntarily delayed actions. Acta Neurobiologiae Experimentalis, 61(3), 175-179.	Matter not relevant
735	Tabares-Seisdedos, R., Escamez, T., Martinez-Gimenez, J. A., Balanza, V., Salazar, J., Selva, G., et al. (2006). Variations in genes regulating neuronal migration predict reduced prefrontal cognition in schizophrenia and bipolar subjects from mediterranean Spain: A preliminary study. Neuroscience, 139(4), 1289-1300.	No IIV
736	Takeyama, H., Itani, T., Tachi, N., Sakamura, O., Murata, K., Inoue, T., et al. (2005). Effects of shift schedules on fatigue and physiological functions among firefighters during night duty. Ergonomics, 48(1), 1-11.	No IIV
737	Takeyama, H., Matsumoto, S., Murata, K., Ebara, T., Kubo, T., Tachi, N., et al. (2004). Effects of the length and timing of nighttime naps on task performance and physiological function. Revista de Saude Publica, 38 Suppl, 32-37.	No IIV
738	Tang, A. C., Pearlmutter, B. A., Malaszenko, N. A., & Phung, D. B. (2002). Independent components of magnetoencephalography: single-trial response onset times. Neuroimage, 17(4), 1773-1789.	Wrong task
739	Tariot, P. N., Patel, S. V., Cox, C., & Henderson, R. E. (1996). Age-related decline in central cholinergic function demonstrated with scopolamine. Psychopharmacology, 125(1), 50-56.	No IIV
740	Tassi, P., Nicolas, A., Seegmuller, C., Dewasmes, G., Libert, J. P., & Muzet, A. (1993). Interaction of the alerting effect of noise with partial sleep deprivation and circadian rhythmicity of vigilance. Perceptual & Motor Skills, 77(3 Pt 2), 1239-1248.	No IIV
741	Tate, R. L., Fenelon, B., Manning, M. L., & Hunter, M. (1991). Patterns of neuropsychological impairment after severe blunt head injury. Journal of Nervous & Mental Disease, 179(3), 117-126.	No RT
742	Ter Maaten, J. C., & Strack van Schijndel, R. J. M. (1996). Cyclic variability of blood pressure and heart rate in tetanus [4]. Intensive Care Medicine, 22(11), 1275-1276.	No RT
743	The Isle Is Full of Noises. (2005). Pharmaceutical Journal, 275(7357), 64.	No empirical data
744	Thorndike, E. L. (1903). Sex differences in mental traits: New York, NY, Lemcke & Buechner.	Sample consisted of children
745	Tiippana, K., Paramei, G. V., & Alshuth, E. (2002). Contrast discrimination and choice reaction times at near-threshold pedestals. Spatial Vision, 15(1), 25-43.	Wrong task
746	Tinius, T. P. (2003a). The Integrated Visual and Auditory Continuous Performance Test as a neuropsychological measure. Archives of Clinical Neuropsychology, 18(5), 439-454.	Wrong task
747	Tinius, T. P. (2003b). The Intermediate Visual and Auditory Continuous Performance Test as a neuropsychological measure. Archives of Clinical Neuropsychology, 18(2), 199-214.	Wrong task
748	Tomaiuolo, F., Scapin, M., Di Paola, M., Le Nezet, P., Fadda, L., Musicco, M., et al. (2007). Gross anatomy of the corpus callosum in Alzheimer's disease: regions of degeneration and their neuropsychological correlates. Dementia & Geriatric Cognitive Disorders, 23(2), 96-103.	No RT
749	Torre, K., Delignieres, D., & Lemoine, L. (2007). 1/f(beta) fluctuations in bimanual coordination: an additional challenge for modeling. Experimental Brain Research, 183(2), 225-234.	No RT
750	Toulopoulou, T., Picchioni, M., Rijsdijk, F., Hua-Hall, M., Ettinger, U., Sham, P., et al. (2007). Substantial genetic overlap between neurocognition and schizophrenia - Genetic modeling in twin samples. Archives of General Psychiatry, 64(12), 1348-1355.	No RT
751	Townsend, J., Harris, N. S., & Courchesne, E. (1996). Visual attention abnormalities in autism: Delayed orienting to location. Journal of the International Neuropsychological Society, 2(6), 541-550.	No IIV
752	Townsend, R. E., & Johnson, L. C. (1979). Relation of frequency-analyzed EEG to monitoring behavior. Electroencephalography & Clinical Neurophysiology, 47(3), 272-279.	Wrong task
753	Tractenberg, R. E., Aisen, P. S., Weiner, M. F., Cummings, J. L., & Hancock, G. R. (2006). Independent contributions of neural and 'higher-order' deficits to symptoms in Alzheimer's disease: A latent variable modeling approach. Alzheimer's & Dementia, 2(4), 303-313.	No RT
754	Travis, L. E. (1924). A test for distinguishing between schizophrenoses and psychoneuroses. The Journal of Abnormal Psychology and Social Psychology, 19(3), 283-298.	No IIV
755	Treat, T. A., McFall, R. M., Viken, R. J., Nosofsky, R. M., MacKay, D. B., & Kruschke, J. K. (2002). Assessing clinically relevant perceptual organization with multidimensional scaling techniques. Psychological Assessment, 14(3), 239-252.	No empirical data
756	TrommershÃ¤user, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. (2005). Optimal Compensation for Changes in Task-Relevant Movement Variability. Journal of Neuroscience, 25(31), 7169-7178.	Wrong task

757	Trommershauser, J., Mattis, J., Maloney, L. T., & Landy, M. S. (2006). Limits to human movement planning with delayed and unpredictable onset of needed information. Experimental Brain Research, 175(2), 276-284.	No IIV
758	Trompet, S., De Craen, A. J. M., Slagboom, P., Shepherd, J., Blauw, G. J., Murphy, M. B., et al. (2008). Genetic variation in the interleukin-1beta-converting enzyme associates with cognitive function. The PROSPER study. Brain, 131(4), 1069-1077.	No RT
759	Uc, E. Y., Rizzo, M., Anderson, S. W., Qian, S., Rodnitzky, R. L., & Dawson, J. D. (2005). Visual dysfunction in Parkinson disease without dementia. Neurology, 65(12), 1907-1913.	No IIV
760	Ulrich, R., & Wing, A. M. (1991). A Recruitment Theory of Force Time Relations in the Production of Brief Force Pulses - the Parallel Force Unit Model. Psychological Review, 98(2), 268-294.	Matter not relevant
761	Unverzagt, F. W., Gao, S., Lane, K. A., Callahan, C., Ogunniyi, A., Baiyewu, O., et al. (2007). Mild cognitive dysfunction: An epidemiological perspective with an emphasis on African Americans. Journal of Geriatric Psychiatry and Neurology, 20(4), 215-226.	No empirical data
762	Vaillancourt, D. E., Slifkin, A. B., & Newell, K. M. (2001). Intermittency in the visual control of force in Parkinson's disease. Experimental Brain Research, 138(1), 118-127.	No RT
763	Vallesi, A., Shallice, T., & Walsh, V. (2007). Role of the prefrontal cortex in the foreperiod effect: TMS evidence for dual mechanisms in temporal preparation. Cerebral Cortex, 17(2), 466-474.	No IIV
764	van den Berg, J., & Neely, G. (2006). Performance on a simple reaction time task while sleep deprived. Perceptual and Motor Skills, 102(2), 589-599.	No IIV
765	van den Bosch, R. J., Rombouts, R. P., & van Asma, M. J. (1996). What determines continuous performance task performance?[erratum appears in Schizophr Bull 1997;23(2):177]. Schizophrenia Bulletin, 22(4), 643-651.	Wrong task
766	Van Der Hiele, K., Vein, A. A., Van Der Welle, A., Van Der Grond, J., Westendorp, R. G. J., Bollen, E. L. E. M., et al. (2007). EEG and MRI correlates of mild cognitive impairment and Alzheimer's disease. Neurobiology of Aging, 28(9), 1322-1329.	No RT
767	Van Der Linden, D., Keijsers, G. P. J., Eling, P., & Van Schaijk, R. (2005). Work stress and attentional difficulties: An initial study on burnout and cognitive failures. Work and Stress, 19(1), 23-36.	Wrong task
768	Van der Lubbe, R. H. J., Los, S. A., Jaskowski, P., & Verleger, R. (2004). Being prepared on time: on the importance of the previous foreperiod to current preparation, as reflected in speed, force and preparation-related brain potentials. Acta Psychologica, 116(3), 245-262.	No IIV
769	van der Steen, J. T., Kruse, R. L., Mehr, D. R., Ribbe, M. W., & van der Wal, G. (2007). Dementia severity, decline and improvement after a lower respiratory tract infection. Journal of Nutrition, Health & Aging, 11(6), 502-506.	No RT
770	Van Dongen, H. P. A., Baynard, M. D., Maislin, G., & Dinges, D. F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. Sleep, 27(3), 423-433.	No IIV
771	van Duinen, H., Renken, R., Maurits, N., & Zijdewind, I. (2007). Effects of motor fatigue on human brain activity, an fMRI study. Neuroimage, 35(4), 1438-1449.	No IIV
772	van Galen, G. P., & van Huygevoort, M. (2000). Error, stress and the role of neuromotor noise in space oriented behaviour. Biological Psychology, 51(2-3), 151-171.	No IIV
773	Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. Journal of Experimental Psychology: General, 132(3), 331-350.	Matter not relevant
774	Van Thiel, E., Meulenbroek, R. G., Hulstijn, W., & Steenbergen, B. (2000). Kinematics of fast hemiparetic aiming movements toward stationary and moving targets. Experimental Brain Research, 132(2), 230-242.	No RT
775	van Zomeren, A. H., ten Duis, H.-J., Minderhoud, J. M., & Sipma, M. (1998). Lightning stroke and neuropsychological impairment: Cases and questions. Journal of Neurology, Neurosurgery & Psychiatry, 64(6), 763-770.	No IIV
776	Vandijk, J. G., Caekebeke, J. F. V., Jennekensschinkel, A., & Zwinderman, A. H. (1992). Background Eeg Reactivity in Auditory Event-Related Potentials. Electroencephalography and Clinical Neurophysiology, 83(1), 44-51.	No IIV
777	Varkevisser, M., Van Dongen, H. P. A., Van Amsterdam, J. G. C., & Kerkhof, G. A. (2007). Chronic insomnia and daytime functioning: an ambulatory assessment. Behavioral Sleep Medicine, 5(4), 279-296.	No IIV
778	Varley, C. K., Holm, V. A., & Eren, M. O. (1985). Cognitive and psychiatric variability in three brothers with fragile X syndrome. Journal of Developmental and Behavioral Pediatrics, 6(2), 87-90.	Sample consisted of children
779	Veggiotti, P., De Agostini, G., Muzio, C., Termine, C., Baldi, P. L., Ferrari Ginevra, O., et al. (1999). Vigabatrin use in psychotic epileptic patients: Report of a prospective pilot study. Acta Neurologica Scandinavica, 99(3), 142-146.	No IIV
780	Veiga, H., Deslandes, A. a., Cagy, M. c., McDowell, K., Pompeu, F., Piedade, R., et al. (2004). Visual event-related potential (P300): A normative study. Arquivos de Neuro-PsiquiatrÃa, 62(3), 575-581.	Wrong task
781	Verbaan, D., Marinus, J., Visser, M., van Rooden, S. M., Stiggelbout, A. M., Middelkoop, H. A. M., et al. (2007). Cognitive impairment in Parkinson's disease. Journal of Neurology, Neurosurgery & Psychiatry, 78(11), 1182-1187.	No RT
782	Verleger, R., Neukater, W., Kompf, D., & Vieregge, P. (1991). On the Reasons for the Delay of P3-Latency in Healthy Elderly Subjects. Electroencephalography and Clinical Neurophysiology, 79(6), 488-502.	No IIV
783	Vernon, P. A., Nador, S., & Kantor, L. (1985). Group differences in intelligence and speed of information-processing. Intelligence, 9(2), 137-148.	Matter not relevant

784	Versavel, M., van Laack, D., Evertz, C., Meier, F., & Kuhlmann, J. (1996). Effect of educational level, gender and age on the performance in a multi-user computerized psychometric test system for use in clinical pharmacological studies.[see comment]. Arzneimittel-Forschung, 46(12), 1179-1185.	No IIV
785	Versavel, M., vanLaack, D., Evertz, C., Unger, S., Meier, F., & Kuhlmann, J. (1997). Test-retest reliability and influence of practice effects on performance in a multi-user computerized psychometric test system for use in clinical pharmacological studies. Arzneimittel-Forschung/Drug Research, 47(6), 781-786.	Wrong task
786	Verschuren, P. M. (2002). Functional foods: Scientific and global perspectives. British Journal of Nutrition, 88(SUPPL. 2), S125-S130.	No empirical data
787	Vesey, R., Birrell, J. M., Bolton, C., Chipperfield, R. S., Blackwell, A. D., Dening, T. R., et al. (2002). Cholinergic nicotinic systems in Alzheimer's disease: prospects for pharmacological intervention. CNS Drugs, 16(7), 485-500.	No empirical data
788	Visser, P. J. (2006). Role of cognitive testing in disease modifying AD trials. Journal of Nutrition, Health & Aging, 10(2), 131-132; discussion 132-133.	No empirical data
789	Voelcker-Rehage, C., Stronge, A. J., & Alberts, J. L. (2006). Age-related Differences in Working Memory and Force Control under Dual-task Conditions. Aging, Neuropsychology, and Cognition, 13(3), 366-384.	No RT
790	Vogel, F., Schalt, E., Kruger, J., Propping, P., & Lehnert, K. F. (1979). The electroencephalogram (EEG) as a research tool in human behavior genetics: psychological examinations in healthy males with various inherited EEG variants. I. Rationale of the study. Material. Methods. Heritability of test parameters. Human Genetics, 47(1), 1-45.	No IIV
791	von Bierbrauer, A., & Weissenborn, K. (1998). P300 after minor head injury (a follow-up examination). Acta Neurologica Belgica, 98(1), 21-26.	Wrong task
792	Vossel, G. (1988). Electrodermal lability, errors, and reaction times: An examination of the motor impulsivity hypothesis. International Journal of Psychophysiology, 6(1), 15-24.	No IIV
793	Voyer, D., & Flight, J. I. (2001). Reliability and magnitude of auditory laterality effects: The influence of attention. Brain and Cognition, 46(3), 397-413.	No RT
794	Vuurman, E. F., Uiterwijk, M. M., Rosenzweig, P., & O'Hanlon, J. F. (1994). Effects of mizolastine and clemastine on actual driving and psychomotor performance in healthy volunteers. European Journal of Clinical Pharmacology, 47(3), 253-259.	Wrong task
795	Wachs, T. D., McCabe, G., Moussa, W., Yunis, F., Kirksey, A., Galal, O., et al. (1996). Cognitive performance of Egyptian adults as a function of nutritional intake and sociodemographic factors. Intelligence, 22(2), 129-154.	Wrong task
796	Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. (2007). An EZ-diffusion model for response time and accuracy. Psychonomic Bulletin & Review, 14(1), 3-22.	Wrong task
797	Wainwright, M. A., Wright, M. J., Luciano, M., Geffen, G. M., & Martin, N. G. (2008). Genetic covariation among facets of openness to experience and general cognitive ability. Twin Research and Human Genetics, 11(3), 275-286.	No IIV
798	Wakui, N., Watanabe, M., Ohe, K., Hironaka, M., Hashimoto, Y., Iwasaki, N., et al. (2004). Hypokalemic Myositic Paralysis Caused by Chronic Toluene Inhalation. Journal of the Medical Society of Toho University, 51(1), 26-31.	No RT
799	Wale, J., & Geffen, G. (1981). Dichotic monitoring as a test of hemispheric dominance in cases of epilepsy. Cortex, 17(1), 135-140.	Wrong task
800	Walhovd, K. B., & Fjell, A. M. (2002). One-year test-retest reliability of auditory ERPs in young and old adults. International Journal of Psychophysiology, 46(1), 29-40.	No RT
801	Walhovd, K. B., & Fjell, A. M. (2007). White matter volume predicts reaction time instability. Neuropsychologia, 45(10), 2277-2284.	Wrong task
802	Wall, M., Maw, R. J., Stanek, K. E., & Chauhan, B. C. (1996). The psychometric function and reaction times of automated perimetry in normal and abnormal areas of the visual field in patients with glaucoma. Investigative Ophthalmology & Visual Science, 37(5), 878-885.	No IIV
803	Warburton, D. M., Bersellini, E., & Sweeney, E. (2001). An evaluation of a caffeinated taurine drink on mood, memory and information processing in healthy volunteers without caffeine abstinence. Psychopharmacology, 158(3), 322-328.	Wrong task
804	Washburn, M. F., Keeler, K., New, K. B., & Parshall, F. M. (1929). Experiments on the relation of reaction-time, cube fluctuations, and mirror drawing to temperamental differences. American Journal of Psychology, 41, 112-117.	Matter not relevant
805	Wassenaar, G. M. (1964). The Effect of General Anxiety as an Index of Lability on the Performance of Various Psychomotor Tasks. Journal of General Psychology, 71, 351-357.	No IIV
806	Wassink, T. H., Nopoulos, P., Pietila, J., Crowe, R. R., & Andreasen, N. C. (2003). NOTCH4 and the frontal lobe in schizophrenia. American Journal of Medical Genetics, Part B, Neuropsychiatric Genetics: the Official Publication of the International Society of Psychiatric Genetics. 118(1), 1-7.	No RT
807	Waterloo, K., Omdal, R., Mellgren, S. I., & Husby, G. (1997). Neuropsychological functions in systemic lupus erythematosus: A comparison with chronic whiplash patients. European Journal of Neurology, 4(2), 171-177.	No RT
808	Watson, B. C., Freeman, F. J., Pool, K. D., Finitzo, T., Chapman, S. B., Mendelsohn, D., et al. (1991). Laryngeal reaction time profiles in spasmodic dysphonia: relationship to cortical electrophysiologic abnormality. Journal of Speech & Hearing Research, 34(2), 269-278.	No IIV
809	Wearden, J. H., Wearden, A. J., & Rabbitt, P. M. A. (1997). Age and IQ effects on stimulus and response timing. Journal of Experimental Psychology: Human Perception and Performance, 23(4), 962-979.	No RT

810	Weiler, J. M., Bloomfield, J. R., Woodworth, G. G., Grant, A. R., Layton, T. A., Brown, T. L., et al. (2000). Effects of fexofenadine, diphenhydramine, and alcohol on driving performance - A randomized, placebo-controlled trial in the Iowa Driving Simulator. Annals of Internal Medicine, 132(5), 354-+.	No IIV
811	Weinmann, S. (1979). Psychological assessment in rehabilitation. International Journal of Rehabilitation Research, 2(4 Suppl. 2), 26-28.	No empirical data
812	Weiss, P. H., Jeannerod, M., Paulignan, Y., & Freund, H. J. (2000). Is the organisation of goal-directed action modality specific? A common temporal structure. Neuropsychologia, 38(8), 1136-1147.	No RT
813	West, R. (2001). The transient nature of executive control processes in younger and older adults. European Journal of Cognitive Psychology, 13(1-2), 91-105.	No empirical data
814	Wexler, B. E., Nicholls, S. S., & Bell, M. D. (2004). Instability of cognitive processing systems in schizophrenia. Schizophrenia Research, 71(2), 513-514.	Matter not relevant
815	Whalley, L. J., Deary, I. J., Starr, J. M., Wahle, K. W., Rance, K. A., Bourne, V. J., et al. (2008). n-3 Fatty acid erythrocyte membrane content, APOE epsilon 4, and cognitive variation: an observational follow-up study in late adulthood. American Journal of Clinical Nutrition, 87(2), 449-454.	No RT
816	White, C., Farley, J. D., & Charles, P. (1987). Chronic schizophrenic disorder: II Reaction time, social performance and arousal. British Journal of Psychiatry, 150, 374-379.	No IIV
817	White, H. K., & Levin, E. D. (1999). Four-week nicotine skin patch treatment effects on cognitive performance in Alzheimer's disease. Psychopharmacology, 143(2), 158-165.	No IIV
818	White, P. M., & Yee, C. M. (1997). Effects of attentional and stressor manipulations on the P50 gating response. Psychophysiology, 34(6), 703-711.	Wrong task
819	White, R. F., & Krengel, M. (Eds.). (1995). Toxicant-induced encephalopathy. New York, NY,US: Guilford Press.	No empirical data
820	Whitfield, K. E., Brandon, D. T., Wiggins, S., Vogler, G., & McClearn, G. (2003). Does intact pair status matter in the study of African American twins? The Carolina African American Twin Study of Aging. Experimental Aging Research, 29(4), 407-423.	No RT
821	Whyte, J., Polansky, M., Fleming, M., Coslett, H. B., & Cavallucci, C. (1995). Sustained Arousal and Attention after Traumatic Brain Injury. Neuropsychologia, 33(7), 797-813.	Wrong task
822	Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2007). Developmental changes in error monitoring: An event-related potential study. Neuropsychologia, 45(8), 1649-1657.	Wrong task
823	Wilen, J., Johansson, A., Kalezic, N., Lyskov, E., & Sandstrom, M. (2006). Psychophysiological tests and provocation of subjects with mobile phone related symptoms. Bioelectromagnetics, 27(3), 204-214.	Wrong task
824	Wilkinson, R. T., & Morlock, H. C. (1967). Auditory evoked response and reaction time. Electroencephalography & Clinical Neurophysiology, 23(1), 50-56.	No IIV
825	Williams, S. W., Byrne, E. J., & Stokes, P. (1993). The treatment of diffuse Lewy body disease: A pilot study. International Journal of Geriatric Psychiatry, 8(9), 731-739.	No RT
826	Wilson, K. G. (1984). Electrodermal Lability and Simple Reaction-Time. Psychophysiology, 21(5), 604-604.	No IIV
827	Wilson, R. S., Gilley, D. W., Bennett, D. A., Beckett, L. A., & Evans, D. A. (2000). Person-specific paths of cognitive decline in Alzheimer's disease and their relation to age. Psychology and Aging, 15(1), 18-28.	No RT
828	Witjes-AnÃ©, M.-N. l. W., Vegter-van der Vlis, M., van Vugt, J. P. P., Lanser, J. B. K., Hermans, J., Zwinderman, A., et al. (2003). Cognitive and motor functioning in gene carriers for Huntington's disease: A baseline study. Journal of Neuropsychiatry & Clinical Neurosciences, 15(1), 7-16.	No IIV
829	Wolber, M., & Wascher, E. (2005). The posterior contralateral negativity as a temporal indicator of visuo-spatial processing. Journal of Psychophysiology, 19(3), 182-194.	Wrong task
830	Woods, S. P., Lovejoy, D. W., & Ball, J. D. (2002). Neuropsychological characteristics of adults with ADHD: A comprehensive review of initial studies. Clinical Neuropsychologist, 16(1), 12-34.	No empirical data
831	Wright, M. J., Smith, G. A., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2000). Genetic influence on the variance in coincidence timing and its covariance with IQ: A twin study. Intelligence, 28(4), 239-250.	Sample consisted of teenagers
832	Wrisberg, C. A. (1981). The Effects of Temporal Variability within and between Series of Stimuli on Reaction-Time and Movement Time. Research Quarterly for Exercise and Sport, 52(4), 518-522.	No IIV
833	Xeno Rasmusson, D., & Brandt, J. (1995). Instability of cognitive asymmetry in Alzheimer's disease. Journal of Clinical & Experimental Neuropsychology: Official Journal of the International Neuropsychological Society, 17(3), 449-458.	No RT
834	Yanagihara, M., & Hori, Y. (1995). Differences in Age-Related-Changes in the P300 Elicited by Visual and Auditory-Stimuli. Neurosciences-Japan, 21(2), 59-65.	Wrong task
835	Yang, A. C. C., Tsai, S. J., Hong, C. J., Yang, C. H., Hsieh, C. H., & Liu, M. E. (2008). Association between heart rate variability and cognitive function in elderly community-dwelling men without dementia: A preliminary report. Journal of the American Geriatrics Society, 56(5), 958-960.	No RT
836	Yerkes, R. M. (1904a). Discussion Reaction-time and variability: A criticism. Psychological Bulletin, 1(10), 370-371.	No empirical data
837	Yerkes, R. M. (1904b). Variability of reaction-time. Psychological Bulletin, 1(5), 137-146.	Matter not relevant
838	Ylikoski, R., Ylikoski, A., Keskivaara, P., Tilvis, R., Sulkava, R., & Erkinjuntti, T. (1999). Heterogeneity of cognitive profiles in aging: Successful aging, normal aging, and individuals at risk for cognitive decline. European Journal of Neurology, 6(6), 645-652.	No RT
839	Yogev, G., Giladi, N., Peretz, C., Springer, S., Simon, E. S., & Hausdorff, J. M. (2005). Dual tasking, gait rhythmicity, and Parkinson's disease: which aspects of gait are attention demanding? European Journal of Neuroscience, 22(5), 1248-1256.	Wrong task

840	Zacny, J. P., Lichtor, J. L., Flemming, D., Coalson, D. W., & Thompson, W. K. (1994). A dose-response analysis of the subjective, psychomotor and physiological effects of intravenous morphine in healthy volunteers. Journal of Pharmacology & Experimental Therapeutics, 268(1), 1-9.	No IIV
841	Zahn, T. P., & Pickar, D. (1993). Autonomic effects of clozapine in schizophrenia: Comparison with placebo and fluphenazine. Biological Psychiatry, 34(1), 3-12.	No IIV
842	Zahn, T. P., Roberts, B. P., Schooler, C., & Cohen, R. (1998). Manual and saccadic reaction time with constant and variable preparatory intervals in schizophrenia. Journal of Abnormal Psychology, 107(2), 328-337.	No IIV
843	Zahn, T. P., Rumsey, J. M., & Van Kammen, D. P. (1987). Autonomic nervous system activity in autistic, schizophrenic, and normal men: Effects of stimulus significance. Journal of Abnormal Psychology, 96(2), 135-144.	No IIV
844	Zanstra, Y. J., Schellekens, J. M. H., Schaap, C., & Kooistra, L. (2006). Vagal and sympathetic activity in burnouts during a mentally demanding workday. Psychosomatic Medicine, 68(4), 583-590.	Wrong task
845	Zappoli, R., Versari, A., Paganini, M., Arnetoli, G., Muscas, G. C., Gangemi, P. F., et al. (1995). Brain Electrical-Activity (Quantitative Eec and Bit-Mapping Neurocognitive Cnv Components), Psychometrics and Clinical Findings in Presenile Subjects with Initial Mild Cognitive Decline or Probable Alzheimer-Type Dementia. Italian Journal of Neurological Sciences, 16(6), 341-376.	No IIV
846	Zec, R. F. (1995). The neuropsychology of aging. Experimental Gerontology, 30(3-4), 431-442.	No empirical data
847	Zeiner, A. R., & Brecher, G. A. (1975). Reaction time performance with and without backscatter from intense pulse light. Aviation Space & Environmental Medicine, 46(2), 125-127.	Matter not relevant
848	Zimprich, D., Hofer, S. M., & Aartsen, M. J. (2004). Short-term versus long-term longitudinal changes in processing speed. Gerontology, 50(1), 17-21.	No RT

Note. RT = reaction time, IIV = intra-individual variability

[a] Detailed description of the inclusion criteria is given in Chapter 2

Appendix E. Retained studies from the general search considered for the healthy ageing review

No.	Study	Reason for exclusion[a]	Details
1	Adam, M., Retey, J. V., Khatami, R., & Landolt, H.-P. (2006). Age-related changes in the time course of vigilant attention during 40 hours without sleep in men. Sleep, 29(1), 55-57.	N/A	INCLUDED IN REVIEW
2	Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. Neuroimage, 8(4), 360-369.	No old	
3	Anderson, L. O. (1921). A Preliminary Report of an Experimental Analysis of Causes of Stuttering. Journal of Applied Psychology, 5(4), 340-349.	No old	Age not reported, but not likely to include older adults
4	Anderson, B., Mennemeier, M., & Chatterjee, A. (2000). Variability not ability: Another basis for performance decrements in neglect. Neuropsychologia, 38(6), 785-796.	No old	No healthy adults aged 60 or over
5	Anstey, K. J. (1999). Sensorimotor variables and forced expiratory volume as correlates of speed, accuracy and variability in reaction time performance in late adulthood. Aging, Neuropsychology, and Cognition, 6(2), 84-95.	No young	Age range: 60-90
6	Anstey, K. J., Mack, H. A., Christensen, H., Li, S.-C., Reglade-Meslin, C., Maller, J., et al. (2007). Corpus callosum size, reaction time speed and variability in mild cognitive disorders and in a normative sample. Neuropsychologia, 45(8), 1911-1920.	No young	Age range: 60-64
7	Anstey, K. J., Dear, K., Christensen, H., & Jorm, A. F. (2005). Biomarkers, health, lifestyle, and demographic variables as correlates of reaction time performance in early, middle, and late adulthood. The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 58(1), 5-21.	N/A	INCLUDED IN REVIEW
8	Arcelin, R., & Brisswalter, J. (1999). Performance stability in simultaneous tasks of pedalling and reaction time. Perceptual and Motor Skills, 88(3), 1193-1199.	No old	
9	Babkoff, H., Kelly, T. L., & Naitoh, P. (2001). Trial-to-trial variance in choice reaction rime as a measure of the effect of stimulants during sleep deprivation. Military Psychology, 13(1), 1-16.	No old	Students in the US Navy; mean age 20.94 ± 2.75
10	Ballard, C. G., Stephens, S., Kenny, R., Kalaria, R., Tovee, M., & O'Brien, J. (2003). Profile of neuropsychological deficits in older stroke survivors without dementia. Dementia and Geriatric Cognitive Disorders, 16(1), 52-56.	No young	Elderly only
11	Ballard, J. C. (2001). Assessing attention: Comparison of response-inhibition and traditional continuous performance tests. Journal of Clinical and Experimental Neuropsychology, 23(3), 331-350.	No young	Elderly only
12	Bankhead, I., & MacKay, D. N. (1982). Fine motor performance in subjects of subnormal, normal, and superior intelligence: I Reaction time and task complexity. Journal of Mental Deficiency Research, 26(2), 73-89.	No old	
13	Barratt, E. S. (1963). Intra-Individual Variability of Performance: Ans and Psychometric Correlates. Texas Reports on Biology & Medicine, 21, 496-504.	No old	Young only
14	Barrett, P. T., Daum, I., & Eysenck, H. J. (1990). Sensory nerve conduction and intelligence: A methodological study. Journal of Psychophysiology, 4(1), 1-13.	No old	
15	Baumeister, A. A., & Kellas, G. (1968). Intrasubject Response Variability in Relation to Intelligence. Journal of Abnormal Psychology, 73(5), 421-423.	No old	No healthy adults aged 60 or over
16	Bekker, E. M., BÃ¶cker, K. B. E., Van Hunsel, F., van den Berg, M. C., & Kenemans, J. L. (2005). Acute effects of nicotine on attention and response inhibition. Pharmacology, Biochemistry and Behavior, 82(3), 539-548.	No old	Only young and middle-aged
17	Benton, A. L., & Blackburn, H. L. (1957). Practice effects in reaction-time tasks in brain-injured patients. The Journal of Abnormal and Social Psychology, 54(1), 109-113.	No healthy adults	Patients from neurological, neurosurgical, and medical services
18	Berkson, G., & Baumeister, A. (1967). Reaction Time Variability of Mental Defectives and Normals. American Journal of Mental Deficiency, 72(2), 262-266.	No old	
19	Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Becic, E. (2006). Testing the limits of cognitive plasticity in older adults: Application to attentional control. Acta Psychologica, 123(3), 261-278.	N/A	INCLUDED IN REVIEW
20	Bielak, A. A. M., Hughes, T. F., Small, B. J., & Dixon, R. A. (2007). It's never too late to engage in lifestyle activities: Significant concurrent but not change relationships between lifestyle activities and cognitive speed. The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 62(6), P331-PP339.	No young	
21	Bleiberg, J., Garmoe, W., Cederquist, J., Reeves, D., & Lux, W. (1993). Effects of dexedrine on performance consistency following brain injury. Neuropsychiatry, Neuropsychology and Behavioral Neurology, 6(4), 245-248.	No old	Age 55 (n=1)
22	Bleiberg, J., Garmoe, W. S., Halpern, E. L., Reeves, D. L., & Nadler, J. D. (1997). Consistency of within-day and across-day performance after mild brain injury. Neuropsychiatry Neuropsychology and Behavioral Neurology, 10(4), 247-253.	No old	Young only

23	Bleiberg, J., Halpern, E. L., Reeves, D., & Daniel, J. C. (1998). Future directions for the neuropsychological assessment of sports concussion. Journal of Head Trauma Rehabilitation, 13(2), 36-44.	Data overlap/ multiple publication	Data used by anotherstudy already included in the review
24	Bookstein, F. L., Streissguth, A. P., Sampson, P. D., Connor, P. D., & Barr, H. M. (2002). Corpus callosum shape and neuropsychological deficits in adult males with heavy fetal alcohol exposure. Neuroimage, 15(1), 233-251.	No old	Age range: 18-36
25	Brewer, N., & Smith, G. A. (1984). How normal and retarded individuals monitor and regulate speed and accuracy of responding in serial choice tasks. Journal of Experimental Psychology: General, 113(1), 71-93.	No old	Age range: 18-36
26	Brewer, N., & Smith, G. A. (1990). Processing speed and mental retardation: deadline procedures indicate fixed and adjustable limitations. Memory & Cognition, 18(5), 443-450.	No old	Age range: 18-30
27	Bruhn, P., & Parsons, O. A. (1971). Continuous reaction time in brain damage. Cortex, 7(3), 278-291.	No healthy adults	Hospitalised patients
28	Bruhn, P., & Parsons, O. A. (1977). Reaction time variability in epileptic and brain-damaged patients. Cortex, 13(4), 373-384.	No healthy adults	All participants had neurological disorders
29	Bunce, D., MacDonald, S. W. S., & Hultsch, D. F. (2004). Inconsistency in serial choice decision and motor reaction times dissociate in younger and older adults. Brain and Cognition, 56(3), 320-327.	N/A	INCLUDED IN REVIEW
30	Bunce, D., Anstey, K. J., Christensen, H., Dear, K., Wen, W., & Sachdev, P. (2007). White matter hyperintensities and within-person variability in community-dwelling adults aged 60-64 years. Neuropsychologia, 45(9), 2009-2015.	No young	
31	Bunce, D., Tzur, M., Rarachurn, A., Gain, F., & Bond, F. W. (2008). Mental health and cognitive function in adults aged 18 to 92 years. Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 63(2), P67-P74.	N/A	INCLUDED IN REVIEW
32	Burley, L. R. (1944). A study of the reaction time of physically trained men. Research Quarterly of the American Association for Health, Physical Education, & Recreation, 15, 232-239.	No old	
33	Burton, C. L., Strauss, E., Hultsch, D. F., Moll, A., & Hunter, M. A. (2006). Intraindividual Variability as a Marker of Neurological Dysfunction: A Comparison of Alzheimer's Disease and Parkinson's Disease. Journal of Clinical and Experimental Neuropsychology, 28(1), 67-83.	No young	
34	Butler, A. J., & Conrad, W. G. (1964). Psychological correlates of abnormal electroencephalographic patterns in familial retardates. Journal of Clinical Psychology, 20(3), 338-343.	No healthy adults	Only mentally retarded patients
35	Carter, J. S., Williams, H. G., Davis, J. M., & Rotter, R. A. (1991). Effects of DDAVP on movement planning and execution processes in healthy young adults. Peptides, 12(1), 1-5.	No old	Age range: 19-32
36	Cation, W. L., Mount, G. E., & Brenner, R. (1951). Variability of reaction time and susceptibility to automobile accidents. Journal of Applied Psychology, 35(2), 101-107.	No old	Age range: 22-55
37	Christensen, H., Dear, K. B. G., Anstey, K. J., Parslow, R. A., Sachdev, P., & Jorm, A. F. (2005). Within-occasion intraindividual variability and preclinical diagnostic status: is intraindividual variability an indicator of mild cognitive impairment? Neuropsychology, 19(3), 309-317.	No young	
38	Cismaru, R., & Chertkow, H. (1999). Is variability a frontal lobe function? Confirmation in dementia subjects. Brain and Cognition, 40(1), 84-85.	No young	
39	Claverie, B., & et al. (1984). Multiparametric outlines with CNV: Application to depressive syndromes. Annals of the New York Academy of Sciences, 425, 556-564.	Inadequate sample size	N=16
40	Collins, L. F., & Long, C. J. (1996). Visual reaction time and its relationship to neuropsychological test performance. Archives of Clinical Neuropsychology, 11(7), 613-623.	No old	Age range: 18-46
41	Cook, D. B., Nagelkirk, P. R., Peckerman, A., Poluri, A., Mores, J., & Natelson, B. H. (2005). Exercise and cognitive performance in chronic fatigue syndrome. Medicine & Science in Sports & Exercise, 37(9), 1460-1467.	No old	Age range: 18-55
42	Cousins, M., & Smyth, M. M. (2003). Developmental coordination impairments in adulthood. Human Movement Science, 22(4), 433-459.	Inadequate sample size	N=17
43	Crawford, T., Goodrich, S., Henderson, L., & Kennard, C. (1989). Predictive responses in Parkinson's disease: Manual keypresses and saccadic eye movements to regular stimulus events. Journal of Neurology, Neurosurgery & Psychiatry, 52(9), 1033-1042.	No young	
44	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), 2499-2507.	No young	
45	Deary, I. J., & Der, G. (2005). Reaction Time, Age, and Cognitive Ability: Longitudinal Findings from Age 16 to 63 Years in Representative Population Samples. Aging, Neuropsychology, and Cognition, 12(2), 187-215.	N/A	INCLUDED IN REVIEW
46	Deary, I. J., Der, G., & Ford, G. (2001). Reaction times and intelligence differences: A population-based cohort study. Intelligence, 29(5), 389-399.	No young	Only 55+

47	Der, G., & Deary, I. J. (2006). Age and Sex Differences in Reaction Time in Adulthood: Results From the United Kingdom Health and Lifestyle Survey. Psychology and Aging, 21(1), 62-73.	N/A	INCLUDED IN REVIEW
48	Diener, H. C., Dichgans, J., Guschlbauer, B., Bacher, M., Rapp, H., & Klockgether, T. (1992). The Coordination of Posture and Voluntary Movement in Patients with Cerebellar Dysfunction. Movement Disorders, 7(1), 14-22.	Inadequate sample size	N=18
49	Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W. S., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. Neuropsychology, 21(3), 381-399.	No young	
50	Doran, S. M., Van Dongen, H. P., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: evidence of state instability. Archives Italiennes de Biologie, 139(3), 253-267.	No old	Age range: 21-47
51	Edinger, J. D., Means, M. K., Carney, C. E., & Krystal, A. D. (2008). Psychomotor performance deficits and their relation to prior nights' sleep amona individuals with primary insomnia. Sleep, 31(5), 599-607.	No sufficient data	Has relevant age groups and IIV, but not presented in the required format in the paper; unable to obtaindata from the author
52	Feldman, R. M., & Reger, S. N. (1967). Relations among Hearing, Reaction Time, and Age. Journal of Speech & Hearing Research, 10(3), 479-495.	No healthy adults	Hospitalised patients
53	Fernandez-Guardiola, A., Ayala, F., & Kornhauser, S. (1968). Eeg, Heart Rate and Reaction Time in Humans: Effect of Variable Vs Fixed Interval Repetitive Stimuli. Physiology & Behavior, 3(2), 231-240.	No old	Age range: 19-32
54	Ferraro, F. R. (1997). Response consistency/inconsistency in Parkinson's disease. Psychology: A Journal of Human Behavior, 34(2), 49-51.	No old	Age range: 18-41
55	Finkel, D., & McGue, M. (2007). Genetic and environmental influences on intraindividual variability in reaction time. Experimental Aging Research, 33(1), 13-35.	N/A	INCLUDED IN REVIEW
56	Fontani, G., Lodi, L., Felici, A., Corradeschi, F., & Lupo, C. (2004). Attentional, emotional and hormonal data in subjects of different ages. European Journal of Applied Physiology, 92(4-5), 452-461.	N/A	INCLUDED IN REVIEW
57	Fontani, G., Lodi, L., Felici, A., Migliorini, S., & Corradeschi, F. (2006). Attention in Athletes of High and Low Experience Engaged in Different Open Skill Sports. Perceptual and Motor Skills, 102(3), 791-805.	No old	
58	Fox, M. D., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2007). Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior.[see comment]. Neuron, 56(1), 171-184.	Inadequate sample size	N=17
59	Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & et al. (1994). Age differences and changes in reaction time: The Baltimore Longitudinal Study of Aging. Journals of Gerontology, 49(4), P179-P189.	N/A	INCLUDED IN REVIEW
60	Fozard, J. L., Thomas, J. C., & Waugh, N. C. (1976). Effects of age and frequency of stimulus repetitions on two-choice reaction time. Journal of Gerontology, 31(5), 556-563.	N/A	INCLUDED IN REVIEW
61	Fuentes, K., Hunter, M. A., Strauss, E., & Hultsch, D. F. (2001). Intraindividual variability in cognitive performance in persons with chronic fatigue syndrome. Clinical Neuropsychologist, 15(2), 210-227.	Inadequate sample size	N=16
62	Galbraith, G. C., Chae, B. C., Cooper, J. R., Gindi, M. M., Ho, T. N., Kim, B. S., et al. (2000). Brainstem frequency-following response and simple motor reaction time. International Journal of Psychophysiology, 36(1), 35-44.	No old	Age range: 18-22
63	Gamberale, F., & Svensson, G. (1974). The effect of anesthetic gases on the psychomotor and perceptual functions of anesthetic nurses. Work-Environment-Health, 11(2), 108-113.	No old	
64	Gardepe, J. P., & Runcie, D. (1983). Heart rate and reaction times for mentally retarded and nonretarded adults in subject- and experimenter-initiated tasks. American Journal of Mental Deficiency, 88(3), 314-320.	No old	
65	Godefroy, O., Lhullier, C., & Rousseaux, M. (1994). Reliability of reaction time measurements in brain-damaged patients. Journal of the Neurological Sciences, 126(2), 168-171.	Inadequate sample size	N=7
66	Goldfarb, W. (1941). An investigation of reaction time in older adults, and its relationship to certain observed mental test patterns. Teachers College Contributions to Education(831), viii + 76.	Strongly suspected insufficient group size (old)	old group age range (55-64) spans middle ages; also given group n=14, 10 adults over 60 not expected
67	Gorus, E., De Raedt, R., & Mets, T. (2006). Diversity, dispersion and inconsistency of reaction time measures: effects of age and task complexity. Aging-Clinical & Experimental Research, 18(5), 407-417.	N/A	INCLUDED IN REVIEW
68	Gorus, E., Lambert, M., De Raedt, R., & Mets, T. (2007). The influence of galantamine on reaction time, attention processes, and performance variability in elderly Alzheimer patients. Journal of Clinical Psychopharmacology, 27(2), 182-187.	No healthy adults	AD patients only
69	Grice, G. R., Spiker, V. A., & Nullmeyer, R. (1979). Variable criterion analysis of individual differences and stimulus similarity in choice reaction time. Perception and Psychophysics, 25(5), 353-370.	No old	Introductory psychology course students

70	Grunberger, J., & Saletu, B. (1980). Determination of pharmacodynamics of psychotropic drugs by psychometric analysis. Progress in Neuro-Psychopharmacology, 4(4-5), 417-434.	Inadequate sample size	N=10
71	Hall, K. R. L., & Stride, E. (1954). Some factors affecting reaction times to auditory stimuli in mental patients. Journal of Mental Science, 100, 462-477.	No old	No healthy elderly, only <40
72	Halpern, L., & Kugelmass, S. (1956). The variability of reaction time in the sensorimotor induction syndrome with special reference to the effect of colors. Journal of Psychology: Interdisciplinary and Applied, 41, 255-269.	No healthy adults	Patients only
73	Halstead, W. C. (1945). Brain injuries and the higher levels of consciousness. Research Publications of the Association for Research in Nervous & Mental Disease, 24, 480-506.	No old	Age = 49 (n=1)
74	Hamon, J. F., Seri, B., & Camara, P. (1989). Motor skill acquisition influences brain responsiveness in sprinters. Activitas Nervosa Superior, 31(1), 1-6.	No old	
75	Hanninen, H., Matikainen, E., Kovala, T., Valkonen, S., & Riihimaki, V. (1994). Internal load of aluminum and the central nervous system function of aluminum welders. Scandinavian Journal of Work, Environment and Health, 20(4), 279-285.	No old	
76	Hausdorff, J. M., Doniger, G. M., Springer, S., Yogev, G., Giladi, N., & Simon, E. S. (2006). A common cognitive profile in elderly fallers and in patients with Parkinson's disease: The prominence of impaired executive function and attention. Experimental Aging Research, 32(4), 411-429.	No old	
77	Hetherington, C. R., Stuss, D. T., & Finlayson, M. A. J. (1996). Reaction time and variability 5 and 10 years after traumatic brain injury. Brain Injury, 10(7), 473-486.	No old	Under 60 only
78	Hillyard, S. A. (1969). Relationships between the contingent negative variation (CNV) and reaction time. Physiology & Behavior, 4(3), 351-357.	No old	Young only
79	Hogan, M. J. (2003). Average speed or variable speed--what do older adults really need? Irish Journal of Psychology, 24(3), 161-183.	N/A	INCLUDED IN REVIEW
80	Hollingworth, H. L. (1912). The influence of caffein on discrimination and choice reaction times: The Science Press.	No old	Age range: 19-39
81	Hughes, J. R., & Fino, J. J. (1992). Changes in reactivity during the 6/second spike and wave complex. Clinical Electroencephalography, 23(1), 31-36.	No old	
82	Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. Neuropsychology, 14(4), 588-598.	No young	
83	Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 57(2), P101-P115.	N/A	INCLUDED IN REVIEW
84	Iacoboni, M., & Zaidel, E. (Eds.). (2003). Stable and variable aspects of callosal channels: Lessons from partial disconnection. Cambridge, MA,US: MIT Press.	No old	Undergraduate students
85	Jennings, J. R., Wood, C. C., & Lawrence, B. E. (1976). Effects of graded doses of alcohol on speed-accuracy tradeoff in choice reaction time. Perception & Psychophysics, 19(1), 85-91.	Inadequate sample size	N=5
86	Jensen, A. R. (1992). The importance of intraindividual variation in reaction time. Personality and Individual Differences, 13(8), 869-881.	No old	
87	Jensen, A. R. (1993). Why is reaction time correlated with psychometric g? Current Directions in Psychological Science, 2(2), 53-56.	No old	
88	Karlsen, N. R., Reinvang, I., & Froland, S. S. (1992). Slowed reaction time in asymptomatic HIV-positive patients. Acta Neurologica Scandinavica, 86(3), 242-246.	Strongly suspected insufficient group size (old)	Healthy group recruited from armed forces, mean age 43.4; not expecting many individuals aged 60 or over
89	Kellas, G. (1969). Reaction-time and response variability of normal and retarded individuals. American Journal of Mental Deficiency, 74(3), 409-414.	No old	
90	King, H. E. (1962). Reaction-time as a function of stimulus intensity among normal and psychotic subjects. Journal of Psychology: Interdisciplinary and Applied, 54(2), 299-307.	Strongly suspected insufficient group size (old)	Control group age range 16-63, mean age 26.1, n=30; not expecting 10 adults aged 60 or above
91	Kirby, N. H., Nettelbeck, T., & Tiggemann, M. (1977). Reaction time in retarded and nonretarded young adults: Sequential effects and response organization. American Journal of Mental Deficiency, 81(5), 492-498.	No old	
92	Knehr, C. A. (1959). Variability in a schizophrenic. Psychological Reports, 5, 792.	No old	
93	Kofman, O., Meiran, N., Greenberg, E., Balas, M., & Cohen, H. (2006). Enhanced performance on executive functions associated with examination stress: Evidence from task-switching and Stroop paradigms. Cognition & Emotion, 20(5), 577-595.	No old	Age range: 19-29
94	Krabbendam, L., Isusi, P., Galdos, P., Echevarria, E., Bilbao, J. R., Martin-Pagola, A., et al. (2006). Associations between COMTVal158Met polymorphism and cognition: direct or indirect effects? European Psychiatry: the Journal of the Association of European Psychiatrists, 21(5), 338-342.	Strongly suspected insufficient group size (old)	Mean age for control groups = 43.3 and 38.3 and include professionals; strongly suspect not many individuals aged 60 or over

95	Krieg, E. F., Jr., Chrislip, D. W., & Russo, J. M. (1996). A mathematical model of performance on a simple reaction time test. Neurotoxicology and Teratology, 18(5), 587-593.	No old	Mean age = 29.5
96	Landauer, A. A., Jellett, L. B., & Kirk, J. (1976). Propranolol and skilled human performance. Pharmacology, Biochemistry & Behavior, 4(3), 283-287.	No old	Age range: 18-31
97	Landauer, A. A., Pocock, D. A., & Prott, F. W. (1979). Effects of atenolol and propranolol on human performance and subjective feelings. Psychopharmacology, 60(2), 211-215.	No old	Age range: 18-33
98	Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A 'worst performance' analysis of individual differences. Intelligence, 14(3), 309-325.	No old	Navy recruits, mean age=19.8
99	Levav, M., Mirsky, A. F., Herault, J., Xiong, L., Amir, N., & Andermann, E. (2002). Familial association of neuropsychological traits in patients with generalized and partial seizure disorders. Journal of Clinical & Experimental Neuropsychology: Official Journal of the International Neuropsychological Society, 24(3), 311-326.	Strongly suspected insufficient group size (old)	Control group mean age = 25.98; not expecting elderly participants
100	Low, L.-F., Anstey, K. J., Jorm, A. F., Christensen, H., & Rodgers, B. (2006). Hormone replacement therapy and cognition in an Australian representative sample aged 60-64 years. Maturitas, 54(1), 86-94.	No young	Age range: 60-64
101	Lowery, D. P., Wesnes, K., & Ballard, C. G. (2007). Subtle attentional deficits in the absence of dementia are associated with an increased risk of post-operative delirium. Dementia and Geriatric Cognitive Disorders, 23(6), 390-394.	No young	
102	Lykkelund, C., Nielsen, J. B., Lou, H. C., Rasmussen, V., Gerdes, A. M., Christensen, E., et al. (1988). Increased neurotransmitter biosynthesis in phenylketonuria induced by phenylalanine restriction or by supplementation of unrestricted diet with large amounts of tyrosine. European Journal of Pediatrics, 148(3), 238-245.	No old	Age range: 15-24
103	Macchi, M. M., Boulos, Z., Ranney, T., Simmons, L., & Campbell, S. S. (2002). Effects of an afternoon nap on nighttime alertness and performance in long-haul drivers. Accident Analysis & Prevention, 34(6), 825-834.	No old	Middle-aged
104	MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2003). Performance variability is related to change in cognition: Evidence from the Victoria Longitudinal Study. Psychology and Aging, 18(3), 510-523.	No young	
105	Makdissi, M., Collie, A., Maruff, P., Darby, D. G., Bush, A., McCrory, P., et al. (2001). Computerised cognitive assessment of concussed Australian Rules footballers. British Journal of Sports Medicine, 35(5), 354-360.	No old	Age range: 17-26
106	Maruff, P., Falleti, M. G., Collie, A., Darby, D., & McStephen, M. (2005). Fatigue-related impairment in the speed, accuracy and variability of psychomotor performance: Comparison with blood alcohol levels. Journal of Sleep Research, 14(1), 21-27.	No old	Age range: 18-40
107	McAuley, T., Yap, M., Christ, S. E., & White, D. e. A. (2006). Revisiting Inhibitory Control Across the Life Span: Insights From the Ex-Gaussian Distribution. Developmental Neuropsychology, 29(3), 447-458.	N/A	INCLUDED IN REVIEW
108	McFarland, R. A. (1937). Psycho-physiological studies at high altitude in the Andes II Sensory and motor responses during acclimatization. Journal of Comparative Psychology, 23(1), 227-258	No old	Age range: 29-44
109	McKerral, M., Lepore, F., & Lachapelle, P. (2001). Response characteristics of the normal retino-cortical pathways as determined with simultaneous recordings of pattern visual evoked potentials and simple motor reaction times. Vision Research, 41(8), 1085-1090.	No old	Age range: 12-32
110	Morrell, L. K., & Morrell, F. (1966). Evoked potentials and reaction times: a study of intra-individual variability. Electroencephalography & Clinical Neurophysiology, 20(6), 567-575.	No old	Age range: 20-37
111	Murtha, S., Cismaru, R., Waechter, R., & Chertkow, H. (2002). Increased variability accompanies frontal lobe damage in dementia. Journal of the International Neuropsychological Society, 8(3), 360-372.	No young	Elderly only
112	Newell, K. M., & Houk, J. C. (1983). Speed and accuracy of compensatory responses to limb disturbances. Journal of Experimental Psychology: Human Perception and Performance, 9(1), 58-74.	Inadequate sample size	N=8 in each experiment
113	Nicholson, M. E., Wang, M., Airhihenbuwa, C. O., Mahoney, B. S., & et al. (1992). Variability in behavioral impairment involved in the rising and falling BAC curve. Journal of Studies on Alcohol, 53(4), 349-356.	No old	
114	Obrist, W. D. (1953). Simple auditory reaction time in aged adults. Journal of Psychology: Interdisciplinary and Applied, 35, 259-266.	N/A	INCLUDED IN REVIEW
115	Olofsen, E., Dinges, D. F., & Van Dongen, H. P. A. (2004). Nonlinear Mixed-Effects Modeling: Individualization and Prediction. Aviation Space and Environmental Medicine, 75(3), A134-A140.	Inadequate sample size	N=13
116	Onofrj, M., Thomas, A., Iacono, D., Luciano, A. L., & di Iorio, A. (2003). The Effects of a Cholinesterase Inhibitor Are Prominent in Patients With Fluctuating Cognition: A Part 3 Study of the Main Mechanism of Cholinesterase Inhibitors in Dementia. Clinical Neuropharmacology, 26(5), 239-251.	No young	Elderly only

117	Orbaek, P., Persson, R., & Osterberg, K. (2005). Impact of trait anxiety and social conformity on responses to experimental chemical challenge. Environmental Toxicology and Pharmacology, 19(3), 659-664.	Strongly suspected insufficient group size (old)	Mean age female: 47.5 ± 6.7, male: 45.4 ±8.0; not expecting enough older adults aged 60+
118	Orton, D. I., & Gruzelier, J. H. (1989). Adverse changes in mood and cognitive performance of house officers after night duty. BMJ, 298(6665), 21-23.	No old	Age range: 24-35
119	Pellizzer, G., & Stephane, M. (2007). Response selection in schizophrenia. Experimental Brain Research, 180(4), 705-714.	Inadequate sample size	N=18
120	Perbal, S., Couillet, J., Azouvi, P., & Pouthas, V. (2003). Relationships between time estimation, memory, attention, and processing speed in patients with severe traumatic brain injury. Neuropsychologia, 41(12), 1599-1610.	No old	Age range: 20-51
121	Persinger, M. A., Lafreniere, G. F., & Mainprize, D. N. (1975). Human Reaction-Time Variability Changes from Low Intensity 3-Hz and 10-Hz Electric-Fields - Interactions with Stimulus Pattern, Sex and Field Intensity. International Journal of Biometeorology, 19(1), 56-64.	No old	Age range: 19-29
122	Porges, S. W. (1972). Heart-Rate Variability and Deceleration as Indexes of Reaction-Time. Journal of Experimental Psychology, 92(1), 103-&.	No old	Volunteers from introductory psychology classes
123	Rabbitt, P. M. A. (Ed.). (2000). Measurement indices, functional characteristics, and psychometric constructs in cognitive aging. New York, NY,US: Oxford University Press.	No young	Age range: 52-79
124	Rakitin, B. C., Scarmeas, N., Li, T., Malapani, C., & Stern, Y. (2006). Single-dose Levodopa Administration and Aging Independently Disrupt Time Production. Journal of Cognitive Neuroscience, 18(3), 376-387.	N/A	INCLUDED IN REVIEW
125	Rammsayer, T. H., & Stahl, J. (2007). Identification of sensorimotor components accounting for individual variability in Zahlen-Verbindungs-Test (ZVT) performance. Intelligence, 35(6), 623-630.	No old	Age range: 19-41
126	Ray, R. L., Piroch, J. F., & Kimmel, H. D. (1977). The effect of task and stimulus variability on habituation of electrodermal and vasomotor reactions. Physiological Psychology, 5(2), 189-196.	No old	Volunteers from an abnormal psychology class
127	Razmjou, S., & Kjellberg, A. (1992). Sustained attention and serial responding in heat: Mental effort in the control of performance. Aviation, Space, and Environmental Medicine, 63(7), 594-601.	No old	
128	Reed, C. L., & Franks, I. M. (1998). Evidence for movement preprogramming and on-line control in differentially impaired patients with Parkinson's disease. Cognitive Neuropsychology, 15(6), 723-745.	Inadequate sample size	N=3
129	Reed, T. E., & Jensen, A. R. (1993). Choice reaction time and visual pathway nerve conduction velocity both correlate with intelligence but appear not to correlate with each other: Implications for information processing. Intelligence, 17(2), 191-203.	No old	Age range: 18-25
130	Reicker, L. I. (2008). The ability of reaction time tests to detect simulation: An investigation of contextual effects and criterion scores. Archives of Clinical Neuropsychology, 23(4), 419-431.	No old	Introductory psychology students
131	Rieger, M., Mayer, G., & Gauggel, S. (2003). Attention deficits in patients with narcolepsy. Sleep, 26(1), 36-43.	No old	Age range: 18-60
132	Rowan, E., McKeith, I. G., Saxby, B. K., O'Brien, J. T., Burn, D., Mosimann, U., et al. (2007). Effects of Donepezil on Central Processing Speed and Attentional Measures in Parkinson's Disease with Dementia and Dementia with Lewy Bodies. Dementia and Geriatric Cognitive Disorders, 23(3), 161-167.	No young	Age range: 71-75
133	Runcie, D., & O'Bannon, R. M. (1975). Relationship of reaction time to deceleration and variability of heart rate in nonretarded and retarded persons. American Journal of Mental Deficiency, 79(5), 553-558.	No old	
134	Saito, Y. (1972). Specification of variation patterns of physiological and performance measurement in sleep loss. Journal of Human Ergology, 1(2), 207-216.	Inadequate sample size	N=3
135	Saletu, B., & Saletu-Zyhlarz, G. M. (2002). Sleep laboratory studies with trazodone in insomnia related to depression and dysthymia. Ceska a Slovenska Psychiatrie, 98(7), 368-376.	Inadequate sample size	N=11
136	Saletu, B., Grunberger, J., Rajna, P., & Karobath, M. (1980). Clovoxamine and fluvoxamine-2 biogenic amine re-uptake inhibiting antidepressants: quantitative EEG, psychometric and pharmacokinetic studies in man. Journal of Neural Transmission - General Section, 49(1-2), 63-86.	No old	Age range: 20-28
137	Saletu, B., Grunberger, J., Anderer, P., Linzmayer, L., Semlitsch, H. V., & Magni, G. (1992). Pharmacodynamics of venlafaxine evaluated by EEG brain mapping, psychometry and psychophysiology. British Journal of Clinical Pharmacology, 33(6), 589-601.	No old	Age range: 21-36
138	Saletu, B., Saletu-Zyhlarz, G., Anderer, P., Brandstatter, N., Frey, R., Gruber, G., et al. (1997). Nonorganic insomnia in generalized anxiety disorder. 2. Comparative studies on sleep, awakening, daytime vigilance and anxiety under lorazepam plus diphenhydramine (Somnium) versus lorazepam alone, utilizing clinical, polysomnographic and EEG mapping methods. Neuropsychobiology, 36(3), 130-152.	No healthy adults	Patients with organic insomnia

139	Saletu, B., Anderer, P., Gruber, G., Mandl, M., Gruber, D., Metka, M., et al. (2001). Insomnia related to postmenopausal syndrome: Sleep laboratory studies on differences between patients and normal controls, and influence of an estrogen-progestogen combination with dienogest Versus estrogen alone and placebo. Drugs of Today, 37(SUPPL. G), 39-62.	Inadequate sample size	N=16
140	Saletu, M., Anderer, P., Saletu, B., Hauer, C., Mandl, M., Semler, B., et al. (2001). Sleep laboratory studies in periodic limb movement disorder (PLMD) patients as compared with normals and acute effects of ropinirole. Human Psychopharmacology, 16(2), 177-187.	Inadequate sample size	N=12
141	Saletu, B., Anderer, P., Saletu, M., Hauer, C., Lindeck-Pozza, L., & Saletu-Zyhlarz, G. (2002). EEG mapping, psychometric, and polysomnographic studies in restless legs syndrome (RLS) and periodic limb movement disorder (PLMD) patients as compared with normal controls. Sleep Medicine, 3(SUPPL. 1), S35-S42.	Strongly suspected insufficient group size (old)	N=33, age range 31-82; n=26 age range 22-78; not expecting 10 participants aged 60 or abouve
142	Saletu-Zyhlarz, G., Saletu, B., Anderer, P., Brandstatter, N., Frey, R., Gruber, G., et al. (1997). Nonorganic insomnia in generalized anxiety disorder. 1. Controlled studies on sleep, awakening and daytime vigilance utilizing polysomnography and EEG mapping. Neuropsychobiology, 36(3), 117-129.	Strongly suspected insufficient group size (old)	Groups ages: 24-66 (mean 41.1, n=34) and 24-65 (mean=43, n=44),
143	Saletu-Zyhlarz, G. M., Hassan Abu-Bakr, M., Anderer, P., Semler, B., Decker, K., Parapatics, S., et al. (2001). Insomnia related to dysthymia: Polysomnographic and psychometric comparison with normal controls and acute therapeutic trials with trazodone. Neuropsychobiology, 44(3), 139-149.	Inadequate sample size	N=11
144	Saletu-Zyhlarz, G., Anderer, P., Brandstatter, N., Dantendorfer, K., Gruber, G., Mandl, M., et al. (2000). Placebo-controlled sleep laboratory studies on the acute effects of zolpidem on objective and subjective sleep and awakening quality in nonorganic insomnia related to neurotic and stress-related disorder. Neuropsychobiology, 41(3), 139-148.	Inadequate sample size	N=15
145	Saunders, E. B., & Isaacs, S. (1929). Tests of reaction time and motor inhibition in the psychoses. American Journal of Psychiatry, 9, 79-112.	No old	Age range: 20-49
146	Schilling, W. (1921). The Effect of Caffein and Acetanilid on Simple Reaction Time. Psychological Review, 28(1), 72-79.	No old	Student volunteers
147	Schwartz, F., Carr, A. C., Munich, R. L., Glauber, S., & et al. (1989). Reaction time impairment in schizophrenia and affective illness: The role of attention. Biological Psychiatry, 25(5), 540-548.	No healthy adults	All had Schizophrenia or an affective disorder
148	Schwartz, F., Munich, R. L., Carr, A., Bartuch, E., & et al. (1991). Negative symptoms and reaction time in schizophrenia. Journal of Psychiatric Research, 25(3), 131-140.	No healthy adults	All had Schizophrenia or an affective disorder
149	Sforza, E., Haba-Rubio, J., De Bilbao, F., Rochat, T., & Ibanez, V. (2004). Performance vigilance task and sleepiness in patients with sleep-disordered breathing. European Respiratory Journal, 24(2), 279-285.	Strongly suspected insufficient group size (old)	Mean age for controls = 42.9; not expecing many participants aged 60 and over
150	Shammi, P., Bosman, E., & Stuss, D. T. (1998). Aging and variability in performance. Aging, Neuropsychology, and Cognition, 5(1), 1-13.	N/A	INCLUDED IN REVIEW
151	Sheridan, M. R., Flowers, K. A., & Hurrell, J. (1987). Programming and execution of movement in Parkinson's disease. Brain, 110(Pt 5), 1247-1271.	Strongly suspected insufficient group size (young)	Three control groups with mean ages of 64.8, 67.0 and 66.5; not expecting a sufficient number of young participants
152	Shipley, B. A., Der, G., Taylor, M. D., & Deary, I. J. (2007). Association between mortality and cognitive change over 7 years in a large representative sample of UK residents. Psychosomatic Medicine, 69(7), 640-650.	Data overlap/ multiple publication	Data used by anotherstudy already included in the review
153	Skuse, N. F., & Burke, D. (1992). Sequence-Dependent Deterioration in the Visual Evoked-Potential in the Absence of Drowsiness. Electroencephalography and Clinical Neurophysiology, 84(1), 20-25.	Inadequate sample size	N=8
154	Smith, C., Carter, M., Sebel, P., & Yate, P. (1991). Mental function after general anaesthesia for transurethral procedures. British Journal of Anaesthesia, 67(3), 262-268.	No young	Elderly only
155	Smulders, F. T. Y., Kok, A., Kenemans, J. L., & Bashore, T. R. (1995). The temporal selectivity of additive factor effects on the reaction process revealed in ERP component latencies. Acta Psychologica, 90(1), 97-109.	No old	Undergraduate students
156	Smulders, F. T. Y., Kenemans, J. L., Jonkman, L. M., & Kok, A. (1997). The effects of sleep loss on task performance and the electroencephalogram in young and elderly subjects. Biological Psychology, 45(1), 217-239.	N/A	INCLUDED IN REVIEW
157	Sparrow, W. A., Begg, R. K., & Parker, S. (2006). Aging Effects on Visual Reaction Time in a Single Task Condition and When Treadmill Walking. Motor Control, 10(3), 201-211.	N/A	INCLUDED IN REVIEW
158	Spirduso, W. W., & Clifford, P. (1978). Replication of age and physical activity effects on reaction and movement time. Journal of Gerontology, 33(1), 26-30.	N/A	INCLUDED IN REVIEW

159	Strauss, E., MacDonald, S. W. S., Hunter, M., Moll, A., & Hultsch, D. F. (2002). Intraindividual variability in cognitive performance in three groups of older adults: Cross-domain links to physical status and self-perceived affect and beliefs. Journal of the International Neuropsychological Society, 8(7), 893-906.	No young	Mean age = 75
160	Strauss, E., Bielak, A. A. M., Bunce, D., Hunter, M. A., & Hultsch, D. F. (2007). Within-person variability in response speed as an indicator of cognitive impairment in older adults. Aging, Neuropsychology, and Cognition, 14(6), 608-630.	No young	Age 65+ only
161	Stuss, D. T., Stethem, L. L., Hugenholtz, H., Picton, T., & et al. (1989). Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance. Journal of Neurology, Neurosurgery & Psychiatry, 52(6), 742-748.	Strongly suspected insufficient group size (old)	Control group age range 16-63, mean age 27, n=22; not expecting 10 older adults
162	Stuss, D. T., Murphy, K. J., Binns, M. A., & Alexander, M. P. (2003). Staying on the job: The frontal lobes control individual performance variability. Brain: A Journal of Neurology, 126(11), 2363-2380.	Inadequate sample size	N=12
163	Sullivan, E. V., Desmond, J. E., Lim, K. O., & Pfefferbaum, A. (2002). Speed and efficiency but not accuracy or timing deficits of limb movements in alcoholic men and women. Alcoholism: Clinical & Experimental Research, 26(5), 705-713.	Strongly suspected insufficient group size (old)	Female controls n=12, mean age=50.8, male n=9, mean age = 49.8; given small ns, not expecting 10+ adults aged 60 or over
164	Surwillo, W. W. (1963). The relation of response-time variability to age and the influence of brain wave frequency. Electroencephalography & Clinical Neurophysiology, 15, 1029-1032.	N/A	INCLUDED IN REVIEW
165	Surwillo, W. W. (1975). Reaction-time variability, periodicities in reaction-time distributions, and the EEG gating-signal hypothesis. Biological Psychology, 3(4), 247-261.	No old	Age range: 23-28
166	Sylvester, P. E., & Roy, R. K. (1971). Hemispherical response to fine touch in normal and mentally subnormal epileptic subjects. Journal of Mental Deficiency Research, 15(4), 303-309.	No old	
167	Thoma, R. J., Yeo, R. A., Gangestad, S., Halgren, E., Davis, J., Paulson, K. M., et al. (2006). Developmental instability and the neural dynamics of the speed-intelligence relationship. Neuroimage, 32(3), 1456-1464.	No old	Age range: 18-30
168	Tiffin, J., & Westhafer, F. L. (1940). The relation between reaction time and temporal location of the stimulus on the tremor cycle. Journal of Experimental Psychology, 27(3), 318-324.	No old	Graduate and draduate students
169	Tombaugh, T. N., Rees, L., Stormer, P., Harrison, A. G., & Smith, A. (2007). The effects of mild and severe traumatic brain injury on speed of information processing as measured by the computerized tests of information processing (CTIP). Archives of Clinical Neuropsychology, 22(1), 25-36.	Strongly suspected insufficient group size (old)	Control grop mean age = 33.12; not expecting a sufficient number of adults aged 60 and above
170	Tucha, O., Mecklinger, L., Laufkotter, R., Klein, H. E., Walitza, S., & Lange, K. W. (2006). Methylphenidate-induced improvements of various measures of attention in adults with attention deficit hyperactivity disorder. Journal of Neural Transmission, 113(10), 1575-1592.	Inadequate sample size	N=16
171	van Hilten, J. J., Wagemans, E. A. H., Ghafoerkhan, S. F., & van Laar, T. (1997). Movement characteristics in Parkinson's disease: Determination of dopaminergic responsiveness and threshold. Clinical Neuropharmacology, 20(5), 402-408.	No healthy adults	All had Parkinson's Disease
172	Vernon, P. A. (1981). Reaction time and intelligence in the mentally retarded. Intelligence, 5(4), 345-355.	No old	
173	Walker, M. P., Ayre, G. A., Ashton, C. H., Marsh, V. R., Wesnes, K., Perry, E. K., et al. (1999). A psychophysiological investigation of fluctuating consciousness in neurodegenerative dementias. Human Psychopharmacology-Clinical and Experimental, 14(7), 483-489.	No young	Elderly only
174	Walker, M. P., Ayre, G. A., Cummings, J. L., Wesnes, K., McKeith, I. G., O'Brien, J. T., et al. (2000). Quantifying fluctuation in dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. Neurology, 54(8), 1616-1624.	No young	Elderly only
175	Walker, M. P., Ayre, G. A., Perry, E. K., Wesnes, K., McKeith, I. G., Tovee, M., et al. (2000). Quantification and characterisation of fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease. Dementia and Geriatric Cognitive Disorders, 11(6), 327-335.	No young	Elderly only
176	Wascher, E., Verleger, R., Jaskowski, P., & Wauschkuhn, B. (1996). Preparation for action: An ERP study about two tasks provoking variability in response speed. Psychophysiology, 33(3), 262-272.	No old	Age range: 19-30
177	Waugh, N. C., Fozard, J. L., Talland, G. A., & Erwin, D. E. (1973). Effects of age and stimulus repetition on two-choice reaction time. Journal of Gerontology, 28(4), 466-470.	Data overlap/ multiple publication	Strongly suspected overlap with another studyalready included
178	West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. Brain and Cognition, 49(3), 402-419.	N/A	INCLUDED IN REVIEW
179	Wilkinson, R. T., & Allison, S. (1989). Age and simple reaction time: Decade differences for 5,325 subjects. Journals of Gerontology, 44(2), 29-35.	N/A	INCLUDED IN REVIEW

180	Williams, B. R., Hultsch, D. F., Strauss, E. H., Hunter, M. A., & Tannock, R. (2005). Inconsistency in Reaction Time Across the Life Span. Neuropsychology, 19(1), 88-96.	N/A	INCLUDED IN REVIEW
181	Williams, B. R., Strauss, E. H., Hultsch, D. F., & Hunter, M. A. (2007). Reaction time inconsistency in a spatial Stroop task: Age-related differences through childhood and adulthood. Aging, Neuropsychology, and Cognition, 14(4), 417-439.	N/A	INCLUDED IN REVIEW
182	Willison, J., & Tombaugh, T. N. (2006). Detecting simulation of attention deficits using reaction time tests. Archives of Clinical Neuropsychology, 21(1), 41-52.	Inadequate sample size	N=15
183	Wilson, K. G. (1987). Electrodermal lability and simple reaction time. Biological Psychology, 24(3), 275-289.	No old	Undergraduate students
184	Winterer, G., Ziller, M., Dorn, H., Frick, K., Mulert, C., Dahhan, N., et al. (1999). Cortical activation, signal-to-noise ratio and stochastic resonance during information processing in man. Clinical Neurophysiology, 110(7), 1193-1203.	No sufficient data	Original data no longer available
185	Winterer, G., Musso, F., Vucurevic, G., Stoeter, P., Konrad, A., Seker, B., et al. (2006). COMT genotype predicts BOLD signal and noise characteristics in prefrontal circuits. Neuroimage, 32(4), 1722-1732.	No old	Mean age 22.7 ± 1.7
186	Wood, G. A. (1977). An electrophysiological model of human visual reaction time. Journal of Motor Behavior, 9(4), 267-274.	No old	
187	Yan, J. H., Thomas, J. R., & Stelmach, G. E. (1998). Aging and rapid aiming arm movement control. Experimental Aging Research, 24(2), 155-168.	N/A	INCLUDED IN REVIEW
188	Zahn, T. P., & Mirsky, A. F. (1999). Reaction time indicators of attention deficits in closed head injury. Journal of Clinical and Experimental Neuropsychology, 21(3), 352-367.	No old	Age range: 18-56

[a] Detailed description of the inclusion criteria is given in Chapter 2

Appendix F. Considered studies from the update search

No.	Study	Reason for exclusion[a]	Additional information
1	Adi-Japha, E., Karni, A., Parnes, A., Loewenschuss, I., & Vakil, E. (2008). A shift in task routines during the learning of a motor skill: group-averaged data may mask critical phases in the individuals' acquisition of skilled performance. Journal of Experimental Psychology: Learning, Memory, & Cognition, 34(6), 1544-1551.	No old	Age range: 19-26
2	Basar-Eroglu, C., Schmiedt-Fehr, C., Mathes, B., Zimmermann, J., & Brand, A. (2009). Are oscillatory brain responses generally reduced in schizophrenia during long sustained attentional processing? International Journal of Psychophysiology, 71(1), 75-83.	Inadequate sample size	n=10 (healthy controls)
3	Bowyer, S. M., Hsieh, L., Moran, J. E., Young, R. A., Manoharan, A., Liao, C. C. J., et al. (2009). Conversation effects on neural mechanisms underlying reaction time to visual events while viewing a driving scene using MEG. Brain Research, 1251(C), 151-161.	Inadequate sample size	n=19 with usable data
4	Brenner, E., & Smeets, J. B. J. (2009). Sources of variability in interceptive movements. Experimental Brain Research, 195(1), 117-133.	Inadequate sample size	n=9
5	Bunce, D., Handley, R., & Gaines, S. O., Jr. (2008). Depression, anxiety, and within-person variability in adults aged 18 to 85 years. Psychology and Aging, 23(4), 848-858	N/A	Included in review
6	Burton, C. L., Strauss, E., Hultsch, D. F., & Hunter, M. A. (2009). The relationship between everyday problem solving and inconsistency in reaction time in older adults. Aging Neuropsychology & Cognition, 16(5), 607-632.	No young	
7	Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. Journal of Neuroscience, 29(24), 7869-7876.	Inadequate sample size	n=12
8	Camicioli, R. M., Wieler, M., de Frias, C. M., & Martin, W. R. W. (2008). Early, untreated Parkinson's disease patients show reaction time variability. Neuroscience Letters, 441(1), 77-80.	Inadequate sample size	Control group n=16
9	Champagne, J., & Fortin, C. (2008). Attention sharing during timing: Modulation by processing demands of an expected stimulus. Perception & Psychophysics, 70(4), 630-639.	No old	Age range: 18-40
10	Cho, S. C., Kim, J. W., Kim, B. N., Hwang, J. W., Park, M., Kim, S. A., et al. (2008). Possible association of the alpha-2A-adrenergic receptor gene with response time variability in attention deficit hyperactivity disorder. American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics, 147(6), 957-963.	Strongly suspected insufficient group size (old)	Children with ADHD and parents
11	Crutch, S. J., & Warrington, E. K. (2008). The Influence of refractoriness upon comprehension of non-verbal auditory stimuli. Neurocase, 14(6), 494-507.	Inadequate sample size	n=10 (healthy controls)
12	Depp, C. A., Savla, G. N., Moore, D. J., Palmer, B. W., Stricker, J. L., Lebowitz, B. D., et al. (2008). Short-term course of neuropsychological abilities in middle-aged and older adults with bipolar disorder. Bipolar Disorders, 10(6), 684-690.	No RT	
13	Duchek, J. M., Balota, D. A., Tse, C. S., Holtzman, D. M., Fagan, A. M., & Goate, A. M. (2009). The Utility of Intraindividual Variability in Selective Attention Tasks as an Early Marker for Alzheimer's Disease. Neuropsychology, 23(6), 746-758.	N/A	Included in review
14	Fagot, D., Dirk, J., Ghisletta, P., & de Ribaupierre, A. (2009). Adults' versus children's performance on the Stroop task: Insights from ex-Gaussian analysis. Swiss Journal of Psychology/Schweizerische Zeitschrift fur Psychologie/Revue Suisse de Psychologie, 68(1), 17-24.	No old	Children and young adults only
15	Fjell, A. M., Rosquist, H., & Walhovd, K. B. (2009). Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. Neurobiology of Aging, 30(12), 2065-2079.	Wrong RT	
16	Goel, N., Rao, H., Durmer, J. S., & Dinges, D. F. (2009). Neurocognitive consequences of sleep deprivation. Seminars in Neurology, 29(4), 320-339.	No empirical data	
17	Gooch, C. M., Stern, Y., & Rakitin, B. C. (2009). Evidence for age-related changes to temporal attention and memory from the choice time production task. Aging Neuropsychology & Cognition, 16(3), 285-310.	N/A	Included in review
18	Gorus, E., De Raedt, R., Lambert, M., Lemper, J. C., & Mets, T. (2008). Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. Journal of Geriatric Psychiatry and Neurology, 21(3), 204-218.	No young	"elderly" only
19	Grierson, L. E. M., Gonzalez, C., & Elliott, D. (2009). Kinematic analysis of early online control of goal-directed reaches: a novel movement perturbation study. Motor Control, 13(3), 280-296.	Inadequate sample size	n=10
20	Hamidovic, A., Dlugos, A., Skol, A., Palmer, A. A., & de Wit, H. (2009). Evaluation of genetic variability in the dopamine receptor D2 in relation to behavioral inhibition and impulsivity/sensation seeking: An exploratory study with d-amphetamine in healthy participants. Experimental and Clinical Psychopharmacology, 17(6), 374-383.	No old	Age range: 18-35
21	Hansen, A. L., Johnsen, B. H., & Thayer, J. F. (2009). Relationship between heart rate variability and cognitive function during threat of shock. Anxiety, Stress, & Coping, 22(1), 77-89.	No old	Age range: 18-36

22	Haraldsson, H. M., Ettinger, U., Magnusdottir, B. B., Sigmundsson, T., Sigurdsson, E., & Petursson, H. (2008). Eye movement deficits in schizophrenia: investigation of a genetically homogenous Icelandic sample. European Archives of Psychiatry & Clinical Neuroscience, 258(6), 373-383.	No old	Age range: 18-55
23	Hayes, S. J., Timmis, M. A., & Bennett, S. J. (2009). Eye movements are not a prerequisite for learning movement sequence timing through observation. Acta Psychologica, 131(3), 202-208.	No old	Age range: 18-21
24	Heerey, E. A., Bell-Warren, K. R., & Gold, J. M. (2008). Decision-Making Impairments in the Context of Intact Reward Sensitivity in Schizophrenia. Biological Psychiatry, 64(1), 62-69.	Wrong RT	
25	Hogan, M. J., Kelly, C. A. M., Verrier, D., Newell, J., Hasher, L., & Robertson, I. H. (2009). Optimal time-of-day and consolidation of learning in younger and older adults. Experimental Aging Research, 35(1), 107-128.	Wrong RT	
26	Hultsch, D. F., Strauss, E., Hunter, M. A., & MacDonald, S. W. (2008). Intraindividual variability, cognition, and aging. In M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (3rd ed., pp. 491-556). Mahwah, NJ: Erlbaum.	No empirical data	
27	Iyer, D., & Zouridakis, G. (2008). Single-Trial Analysis of the Auditory N100 Improves Separation of Normal and Schizophrenia Subjects. In 2008 30th Annual International Conference of the Ieee Engineering in Medicine and Biology Society, 1-8, 3840-3843.	No old	
28	Johns, M., Crowley, K., Chapman, R., Tucker, A., & Hocking, C. (2009). The effect of blinks and saccadic eye movements on visual reaction times. Attention, Perception, & Psychophysics, 71(4), 783-788.	No old	Age range: 17-32
29	Jones, C. R. G., Malone, T. J. L., Dirnberger, G., Edwards, M., & Jahanshahi, M. (2008). Basal ganglia, dopamine and temporal processing: performance on three timing tasks on and off medication in Parkinson's disease. Brain & Cognition, 68(1), 30-41.	Strongly suspected insufficient group size (young)	PD patients and controls; mean age 62.83 and 67.65
30	Kamienkowski, J. E., & Sigman, M. (2008). Delays without mistakes: response time and error distributions in dual-task. PLoS ONE, 3(9), e3196.	Inadequate sample size	n=16
31	Kato, A., Watanabe, Y., Sawara, K., & Suzuki, K. (2008). Diagnosis of sub-clinical hepatic encephalopathy by Neuropsychological Tests (NP-tests). Hepatology Research, 38, S122-S127.	No IIV	
32	Kim, C. Y., Lee, G., Choi, H., & Goh, J. (2009). Repeated measures of reaction times among patients with schizophrenia. Clinical Psychopharmacology and Neuroscience, 7(1), 20-22.	No old	
33	Kollins, S. H., Anastopoulos, A. D., Lachiewicz, A. M., Fitzgerald, D., Morrissey-Kane, E., Garrett, M. E., et al. (2008). SNPs in dopamine D2 receptor gene (DRD2) and norepinephrine transporter gene (NET) are associated with continuous performance task (CPT) phenotypes in ADHD children and their families. American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics, 147(8), 1580-1588.	Strongly suspected insufficient group size (old)	Parents of children with ADHD; mean age = 41.1
34	Kollins, S. H., McClernon, F. J., & Epstein, J. N. (2009). Effects of smoking abstinence on reaction time variability in smokers with and without ADHD: an ex-Gaussian analysis. Drug & Alcohol Dependence, 100(1-2), 169-172.	Inadequate sample size	Smokers with no ADHD; n=12
35	Kuningas, M., Mooijaart, S. P., Jolles, J., Slagboom, P. E., Westendorp, R. G. J., & van Heemst, D. (2009). VDR gene variants associate with cognitive function and depressive symptoms in old age. Neurobiology of Aging, 30(3), 466-473.	No young	85+ only
36	Laessoe, U., Hoeck, H. C., Simonsen, O., & Voigt, M. (2008). Residual attentional capacity amongst young and elderly during dual and triple task walking. Human Movement Science, 27(3), 496-512.	No RT	
37	Lafortune, L., Beland, F., Bergman, H., & Ankri, J. (2009). Health status transitions in community-living elderly with complex care needs: a latent class approach. BMC Geriatrics, 9, 6.	No young	64+ only
38	Leue, A., & Beauducel, A. (2008). A meta-analysis of Reinforcement Sensitivity Theory: On performance parameters in reinforcement tasks. Personality and Social Psychology Review, 12(4), 353-369.	No empirical data	
39	Levinthal, B. R., Morrow, D. G., Tu, W. Z., Wu, J. W., & Murray, M. D. (2008). Cognition and health literacy in patients with hypertension. Journal of General Internal Medicine, 23(8), 1172-1176.	No healthy adults	All had hypertension
40	Li, S.-C., Hammerer, D., Muller, V., Hommel, B., & Lindenberger, U. (2009). Lifespan development of stimulus-response conflict cost: Similarities and differences between maturation and senescence. Psychological Research/Psychologische Forschung, 73(6), 777-785.	N/A	Included in review
41	Lipnicki, D. M., Gunga, H. C., Belavy, D. L., & Felsenberg, D. (2009). Bed Rest and Cognition: Effects on Executive Functioning and Reaction Time. Aviation Space and Environmental Medicine, 80(12), 1018-1024.	No old	Age range: 21-45
42	Lovden, M., Schaefer, S., Pohlmeyer, A. E., & Lindenberger, U. (2008). Walking variability and working-memory load in aging: A dual-process account relating cognitive control to motor control performance. Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 63(3), P121-P128.	No RT	

43	Luciano, M., Gow, A. J., Harris, S. E., Hayward, C., Allerhand, M., Starr, J. M., et al. (2009). Cognitive ability at age 11 and 70 years, information processing speed, and APOE variation: The Lothian Birth Cohort 1936 study. Psychology and Aging, 24(1), 129-138.	No young	Data from age 11 and 70 (RT from 70)
44	Lutz, A., Slagter, H. A., Rawlings, N. B., Francis, A. D., Greischar, L. L., & Davidson, R. J. (2009). Mental training enhances attentional stability: Neural and behavioral evidence. The Journal of Neuroscience, 29(42), 13418-13427.	Strongly suspected insufficient group size (old)	n=17, ages 22-64; n=23, ages 20-62
45	Macdonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2008). Predicting impending death: inconsistency in speed is a selective and early marker. Psychology & Aging, 23(3), 595-607.	No young	VLS participants initially aged 55-85
46	MacDonald, S. W., Li, S.-C., & Backman, L. (2009). Neural underpinnings of within-person variability in cognitive functioning. Psychology and Aging, 24(4), 792-808	No empirical data	
47	MacDonald, S. W. S., Cervenka, S., Farde, L., Nyberg, L., & Backman, L. (2009). Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. Neuropsychologia, 47(11), 2299-2304.	Inadequate sample size	n=16
48	Madison, G., Forsman, L., Blom, O., Karabanov, A., & Ullen, F. (2009). Correlations between intelligence and components of serial timing variability. Intelligence, 37(1), 68-75.	Strongly suspected insufficient group size (old)	Age rage: 19-62, n=30
49	Martin, D. M., Burns, N. R., & Wittert, G. (2009). Free Testosterone Levels, Attentional Control, and Processing Speed Performance in Aging Men. Neuropsychology, 23(2), 158-167.	N/A	Included in review
50	Martino, D. J., Strejilevich, S. A., Scapola, M., Igoa, A., Marengo, E., Ais, E. D., et al. (2008). Heterogeneity in cognitive functioning among patients with bipolar disorder. Journal of Affective Disorders, 109(1-2), 149-156.	No old	Controls matched on age to 18-55
51	McCann, J. C., & Ames, B. N. (2008). Is there convincing biological or behavioral evidence linking vitamin D deficiency to brain dysfunction? FASEB Journal, 22(4), 982-1001.	No empirical data	
52	Medina, J. M. (2009). 1/falpha noise in reaction times: a proposed model based on Pieron's law and information processing. Physical Review E Statistical, Nonlinear, & Soft Matter Physics, 79(1 Pt 1), 011902.	No human participants	Simulation data
53	Miller, J., Sproesser, G., & Ulrich, R. (2008). Constant versus variable response signal delays in speed-accuracy trade-offs: Effects of advance preparation for processing time. Perception & Psychophysics, 70(5), 878-886.	No old	n=2 students; n=12 students
54	Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. Psychological Science, 19(11), 1071-7.	No old	Age range: 18-25
55	Nebes, R. D., Buysse, D. J., Halligan, E. M., Houck, P. R., & Monk, T. H. (2009). Self-reported sleep quality predicts poor cognitive performance in healthy older adults. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 64(2), 180-187.	No young	Age range: 65-80
56	Negahban, H., Hadian, M. R., Salavati, M., Mazaheri, M., Talebian, S., Jafari, A. H., et al. (2009). The effects of dual-tasking on postural control in people with unilateral anterior cruciate ligament injury. Gait & Posture, 30(4), 477-481.	Strongly suspected insufficient group size (old)	n=27; mean 26.29
57	Negash, S., Greenwood, P. M., Sunderland, T., Parasuraman, R., Geda, Y. E., Knopman, D. S., et al. (2009). The Influence of Apolipoprotein E Genotype on Visuospatial Attention Dissipates After Age 80. Neuropsychology, 23(1), 81-89.	Wrong RT	
58	Payton, A. (2009). The impact of genetic research on our understanding of normal cognitive ageing: 1995 to 2009. Neuropsychology Review, 19(4), 451-477.	No empirical data	
59	Personnier, P., Paizis, C., Ballay, Y., & Papaxanthis, C. (2008). Mentally represented motor actions in normal aging II. The influence of the gravito-inertial context on the duration of overt and covert arm movements. Behavioural Brain Research, 186(2), 273-283.	No RT	
60	Pietrzak, R. H., Snyder, P. J., Jackson, C. E., Olver, J., Norman, T., Piskulic, D., et al. (2009). Stability of cognitive impairment in chronic schizophrenia over brief and intermediate re-test intervals. Human Psychopharmacology, 24(2), 113-121.	Strongly suspected insufficient group size (old)	mean 39.2; n=20
61	Ratcliff, R. (2008). Modeling aging effects on two-choice tasks: Response signal and response time data. Psychology and Aging, 23(4), 900-916.	Wrong RT	
62	Reicker, L. I. (2008). The ability of reaction time tests to detect simulation: An investigation of contextual effects and criterion scores. Archives of Clinical Neuropsychology, 23(4), 419-431.	No old	University students
63	Reimers, S., & Stewart, N. (2008). Using Adobe Flash Lite on mobile phones for psychological research: Reaction time measurement reliability and interdevice variability. Behavior Research Methods, 40(4), 1170-1176	No old	Undergraduate psychology students
64	Rodriguez-Sanchez, J. M., Perez-Iglesias, R., Gonzalez-Blanch, C., Pelayo-Teran, J. M., Mata, I., Martinez, O., et al. (2008). 1-year follow-up study of cognitive function in first-episode non-affective psychosis. Schizophrenia Research, 104(1-3), 165-174.	No old	Age range: 15-60

65	Schmiedek, F., Lovden, M., & Lindenberger, U. (2009). On the relation of mean reaction time and intraindividual reaction time variability. Psychology and Aging, 24(4), 841-857.	Wrong RT	WM (3 back) component
66	Sculthorpe, L. D., Stelmack, R. M., & Campbell, K. B. (2009). Mental ability and the effect of pattern violation discrimination on P300 and mismatch negativity. Intelligence, 37(4), 405-411.	No old	n=30; M=19.8(2.3)
67	Sedlackova, S., Rektorova, I., Srovnalova, H., & Rektor, I. (2009). Effect of high frequency repetitive transcranial magnetic stimulation on reaction time, clinical features and cognitive functions in patients with Parkinson's disease. Journal of Neural Transmission, 116(9), 1093-1101.	No healthy adults	Patients with PD
68	Sestieri, C., Pizzella, V., Cianflone, F., Romani, G. L., & Corbetta, M. (2008). Sequential activation of human oculomotor centers during planning of visually-guided eye movements: a combined fMRI-MEG study. Frontiers in Human Neuroscience, 1.	No old	Age range: 20-31
69	Shapkova, E. Y., Shapkova, A. L., Goodman, S. R., Zatsiorsky, V. M., & Latash, M. L. (2008). Do synergies decrease force variability? A study of single-finger and multi-finger force production. Experimental Brain Research, 188(3), 411-425.	Inadequate sample size	n=7
70	Smyrnis, N., Karantinos, T., Malogiannis, I., Theleritis, C., Mantas, A., Stefanis, N. C., et al. (2009). Larger variability of saccadic reaction times in schizophrenia patients. Psychiatry Research, 168(2), 129-136.	No old	Age range:18-25
71	Sosnoff, J. J., & Newell, K. M. (2008). Age-Related Loss of Adaptability to Fast Time Scales in Motor Variability. Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 63(6), P344-P352.	No RT	
72	Sparrow, W. A., Begg, R. K., & Parker, S. (2008). Variability in the foot-ground clearance and step timing of young and older men during single-task and dual-task treadmill walking. Gait & Posture, 28(4), 563-567.	Data overlap/ multiple publication	Same as Sparrow et al. 2006
73	Staines, D. R. (2008). Are multiple sclerosis and amyotrophic lateral sclerosis autoimmune disorders of endogenous vasoactive neuropeptides? Medical Hypotheses, 70(2), 413-418.	No empirical data	
74	Steinborn, M., Rolke, B., Bratzke, D., & Ulrich, R. (2008). Intraindividual reaction time variability as a measure of circadian rhythms in cognitive control. International Journal of Psychology, 43(3-4), 656-656.	Strongly suspected insufficient group size (old)	n=24, mean age=26.2; n=30, mean age=26; n=30. mean=23.7
75	Stoet, G., Ruge, H., & Snyder, L. H. (2008). Modification of response time variability in a decision-making task. NeuroReport: For Rapid Communication of Neuroscience Research, 19(13), 1321-1324.	No old	Median age = 22
76	Stopford, C. L., Snowden, J. S., Thompson, J. C., & Neary, D. (2008). Variability in cognitive presentation of Alzheimer's disease. Cortex, 44(2), 185-195.	No RT	
77	Sullivan, E. V., Rose, J., Rohlfing, T., & Pfefferbaum, A. (2009). Postural sway reduction in aging men and women: relation to brain structure, cognitive status, and stabilizing factors. Neurobiology of Aging, 30(5), 793-807.	No RT	
78	Terry, P., Doumas, M., Desai, R. I., & Wing, A. M. (2009). Dissociations between motor timing, motor coordination, and time perception after the administration of alcohol or caffeine. Psychopharmacology, 202(4), 719-729.	Strongly suspected insufficient group size (old)	Mean age = 24.3 (university employees or students)
79	Towgood, K. J., Meuwese, J. D. I., Gilbert, S. J., Turner, M. S., & Burgess, P. W. (2009). Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. Neuropsychologia, 47(13), 2981-2988.	No old	Age range: 20-43
80	Ullen, F., Forsman, L., Blom, O., Karabanov, A., & Madison, G. (2008). Intelligence and variability in a simple timing task share neural substrates in the prefrontal white matter. The Journal of Neuroscience, 28(16), 4238-43.	No old	Age range: 19-49
81	Valdez, A. B., & Amazeen, E. L. (2009). Target dimension affects 1/f noise in aiming. Nonlinear Dynamics, Psychology, & Life Sciences, 13(4), 369-392.	No old	Age range: 25-30
82	Werner, S., Bock, O., & Timmann, D. (2009). The effect of cerebellar cortical degeneration on adaptive plasticity and movement control. Experimental Brain Research, 193(2), 189-96.	Inadequate sample size	n=17 controls healthy
83	Wester, A. E., Bockner, K. B. E., Volkerts, E. R., Verster, J. C., & Kenemans, J. L. (2008). Event-related potentials and secondary task. performance during simulated driving. Accident Analysis and Prevention, 40(1), 1-7.	No old	Age range: 21-30
84	Wicks, P., Abrahams, S., Papps, B., Al-Chalabi, A., Shaw, C. E., Leigh, P. N., et al. (2009). SOD1 and cognitive dysfunction in familial amyotrophic lateral sclerosis. Journal of Neurology, 256(2), 234-241.	No RT	
85	Wijnants, M. L., Bosman, A. M. T., Hasselman, F., Cox, R. F. A., & Van Orden, G. C. (2009). 1/f scaling in movement time changes with practice in precision aiming. Nonlinear Dynamics, Psychology, & Life Sciences, 13(1), 79-98.	No old	Undergraduate students
86	, P., Biebl, R., & Ansorge, U. (2008). The impact of stimulus and response variability on S-R correspondence effects. Journal of Experimental Psychology: Learning, Memory, & Cognition, 34(3), 533-545.	No old	Age range:18-33
87	Wylie, S. A., van den Wildenberg, W. P. M., Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., et al. (2009). The effect of Parkinson's disease on interference control during action selection. Neuropsychologia, 47(1), 145-157.	Strongly suspected insufficient group size (young)	Mean age 67.7

88	Xu-Wilson, M., Chen-Harris, H., Zee, D. S., & Shadmehr, R. (2009). Cerebellar contributions to adaptive control of saccades in humans. Journal of Neuroscience, 29(41), 12930-12939.	Inadequate sample size	2 control groups; total n=17
89	Yang, A. C. C., Tsai, S.-J., Hong, C.-J., Yang, C.-H., Hsieh, C.-H., & Liu, M.-E. (2008). Association between heart rate variability and cognitive function in elderly community-dwelling men without dementia: a preliminary report. Journal of the American Geriatrics Society, 56(5), 958-960.	No young	Elderly subjects
90	Yarkoni, T., Barch, D. M., Gray, J. R., Conturo, T. E., & Braver, T. S. (2009). BOLD correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fMRI analysis. PLoS ONE, 4(1), e4257.	No old	Young adults

Note. RT = reaction time, IIV = intra-individual variability

[a] Detailed description of the inclusion criteria is given in Chapter 2