Developmental Psychology
2012, Vol. 48, No. 5, 1262–1276


0012-1649/12/$12.00 DOI: 10.1037/a0027550

# Sex Differences in Reaction Time Mean and Intraindividual Variability Across the Life Span

Dominika Dykiert
University of Edinburgh

Geoff Der
Medical Research Council Social and Public Health Sciences Unit, Glasgow, United Kingdom, and University of Edinburgh

John M. Starr and Ian J. Deary
University of Edinburgh

Men are often found to have faster and less variable reaction times (RTs) than do women. However, it has not been established whether these differences occur in children. One suggestion is that sex differences in RT variability may be due to the effect of sex hormones on the brain and, by implication, may be expected in adults but not in children. The present study investigates sex differences in RT mean and intraindividual variability in a sample that includes both children and adults (age range = 4–75 years). Mean and intraindividual variability of simple RT (SRT) and 4-choice RT (CRT) were measured in 1,994 visitors to science festivals held in Edinburgh, Scotland, in 2008 and 2009 and in Cheltenham and Cambridge, England, in 2008. The commonly reported pattern of decreasing RT mean and variability in childhood and adolescence, followed by an increase in mean and variability through adulthood and into old age, was confirmed. Greater intraindividual variability for females in SRT and CRT was observed in adults but not in children. Males had significantly faster mean SRT than did females across the life span, but there were no sex differences in mean CRT.

*Keywords:* reaction time, intraindividual variability, sex differences, aging

*Supplemental materials:* http://dx.doi.org/10.1037/a0027550.supp

Reaction time (RT) has been used as a test of cognitive functioning for over a century (see Jensen, 1982). Two of the most common RT tests are simple and choice RT (SRT and CRT, respectively). In SRT tasks there is only a single stimulus, which is repeated over trials, and in CRT there are multiple stimuli, with each having its respective response. The appeal of the RT tests is that they are relatively simple and quick to administer, yet they provide a useful measure of cognitive functioning: RTs correlate negatively with general intelligence (Deary, Der, & Ford, 2001) and are slower in patients with neurodegenerative disorders (Burton, Strauss, Hultsch, Moll, & Hunter, 2006).

This article was published Online First March 5, 2012.

Dominika Dykiert, Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom; Geoff Der, Medical Research Council Social and Public Health Sciences Unit, Glasgow, United Kingdom, and Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh; John M. Starr, Centre for Cognitive Ageing and Cognitive Epidemiology, Geriatric Medicine Unit, University of Edinburgh; Ian J. Deary, Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh.

This work was supported by a U.K. Medical Research Council PhD studentship to Dominika Dykiert. Geoff Der was funded by the U.K. Medical Research Council. The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology is funded as part of the joint U.K. Medical Research Council call for Lifelong Health and Wellbeing (G0700704/84698).

Correspondence concerning this article should be addressed to Ian J. Deary, Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, United Kingdom. E-mail: i.deary@ed.ac.uk

## Sex Differences in RT

Sex differences in cognition have been the subject of numerous investigations over the past few decades. Reviews of the literature (e.g., Halpern, 1992; Herlitz & Lovén, 2009) have shown that sex differences are found in specific cognitive abilities but not general intelligence (e.g., Deary, Irwing, Der, & Bates, 2007). For example, women tend to perform better than men on tests of verbal ability and episodic memory, whereas men outperform women on visuospatial and quantitative ability (Halpern, 1992; Herlitz & Lovén, 2009).

Much research interest has been devoted to the study of sex difference in RT, and it is often reported that men have faster RTs than do women. This effect has been found in a number of samples, ranging from university students 18–25 years of age (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 schoolchildren, although the effect was not clear when a two-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 was separated into decision and movement components, females outperformed males on the former, while males outperformed 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 centered not only on mean RT but also on intraindividual variability of RTs; that is, the consistency (or rather inconsistency) of an individual's responses across trials within a test. Intraindividual variability, although highly correlated with mean RT, is a discrete measure of cognitive performance (Jensen, 1992). Only a few studies to date have investigated sex differences in intraindividual variability in RT, and they have suggested that women are less consistent than men (Deary & Der, 2005; Der & Deary, 2006; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 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 average SRT or CRT. One factor that may explain some of the discrepancies in the research findings is the different ages of groups that have been studied. It is now well established that, in general, RTs become shorter with age in childhood (Gilbert, 1894; Goodenough, 1935) and longer with increasing age in adulthood (Deary & Der, 2005; Der & Deary, 2006; Fozard et al., 1994). Studies investigating RT through the life span, which are mostly cross-sectional, usually find a U-shaped relationship between mean RT and age (Koga & Morant, 1923; Wilkinson & Allison, 1989) and between RT intraindividual variability and age (Wilkinson & Allison, 1989; Williams, Hultsch, Strauss, Hunter, & Tannock, 2005; Williams, Strauss, Hultsch, & Hunter, 2007). Few investigators have examined the developmental patterns of RT separately for males and females, and 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 life span.

Many of the studies concerned with sex differences in RT at different ages had small samples ($n = 36$; Fairweather & Hutt, 1972), restricted age ranges (4.5–11.5 and 10–16 years, Eckert & Eichorn, 1977; 6–11 years, Fairweather & Hutt, 1972), or even inadequately matched age groups (Bellis, 1933). It is therefore not surprising that they produced inconsistent findings. Only a few studies have investigated sex and age effects on RT in large samples ($n > 500$) and even fewer in population-representative ones. Noble et al. (1964) studied mean CRTs of 600 people aged 8–87 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, followed by a period of stability and then a 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 males' 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 advantage that males have might arise from their prolonged period of improvement in RT relative to females, which enables them to achieve a higher level of performance before they start declining with age. Observations by Noble and colleagues are interesting, but many were not subjected to statistical testing and are based merely on visual inspection of graphs.

Several more recent studies have provided evidence based on a more rigorous treatment of data (Deary & Der, 2005; Der & Deary, 2006; Fozard et al., 1994), but none have included children in their samples. Fozard et al. (1994) analyzed 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 20–90, both RTs became slower, but the change was larger for RTs from the go/no-go task than for 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 intraindividual variability in the two RT tasks and found that it, too, increased with age, but a sex difference was found only in the more complex go/no-go task. Again, there was no Sex $\times$ Age interaction.

In a large sample from the U.K. Health and Lifestyle Survey (HALS), including 7,130 adults representing all ages between 18 and 94, Der and Deary (2006) found significant sex differences in RT, with women's RTs being slower and more variable than men's. The sex difference was most marked in CRT intraindividual standard deviation (ISD). Unlike in Fozard et al.'s (1994) data, significant Age $\times$ 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 (2005) 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 significant only 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, and 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 postpubertal exposure to adult levels of sex hormones may lead to differentiation of responding between the genders. If correct, this hypothesis would predict that sex differences in RT intraindividual variability would be found in adults but not in children or young teenagers. The present study aims to test this prediction.

### The Relationship Between Mean RT and Intraindividual Variability

One issue facing researchers interested in RT intraindividual variability is that it is highly correlated with the mean RT; that is, people who have slower RTs tend to have greater intraindividual variability, too (Deary & Der, 2005). Various methods have been used to deal with this issue. Among the simplest ways of controlling for mean RT is using the coefficient of variation (CV), which is the ratio of RT ISD to mean RT for each individual. The simplicity of this measure and ease of computation place it among the most commonly used mean-adjusted measures of intraindividual variability. However, CV is not without its limitations. For example, Hultsch, Strauss, Hunter, and MacDonald (2008) pointed out that it confounds unsystematic variability with systematic effects such as practice and fatigue. Further, they argued that the usefulness of CV is limited because it is a cross-product of the main effect of ISD, the main effect of mean RT, and their interaction. Some authors (e.g., Hultsch, MacDonald, & Dixon, 2002; Williams et al., 2005) have opted for a regression method that allows them to remove group differences in mean RT as well as effects of practice or fatigue. They used residuals from the regression model rather than raw data to calculate RT ISD. Finally, mean RT may be controlled by including it as a covariate in the models of RT intraindividual variability. Although like CV this method does not control for systematic sources of variation, it allows for a more precise modeling of the relationship between mean and ISD. For example, nonlinear associations may be modeled and tested.

Whether mean RT is controlled when studying intraindividual variability, and how it is achieved, can have important implications for the study results. For example, Fozard et al. (1994) found that, when intraindividual 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 modeled 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 (2005) found that, when the regression method is used to control for the mean, the effects of age (but not of sex) on RT intraindividual 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, 2005; 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

In this study, we present cross-sectional SRT and CRT data from a large sample that spans child and adult age ranges. We investigated the shape of relationships between age and mean RT and between age and RT variability across the life span. We expected RT mean and intraindividual variability to decrease with age through childhood and adolescence until early adulthood and increase thereafter. In terms of sex differences, we hypothesized that greater female RT intraindividual 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 intraindividual variability, we used both RT ISD and RT CV as outcomes. In addition, we modeled RT ISD with and without controlling for mean RT.

## Method

### Participants

Participants were visitors to science festivals held in Edinburgh, Scotland, in 2008 and 2009 and in Cambridge and Cheltenham, England, in 2008. Because the festivals were aimed at children, participants under 18 years of age constituted the bulk of our sample. However, the accompanying parents and guardians were also invited to participate, so the overall age range was 3–87 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 eight 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 later in the Data Preparation section. 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 ($M = 15.35$, median = 10.34, $SD = 13.44$).

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. In order to check whether subsamples compared here were of equivalent socioeconomic status, we linked the records to the Index of Multiple Deprivation (IMD) score for their area of residence using their postcode (Office for National Statistics, 2006). The IMD scores are indices of relative deprivation for small geographical areas derived from data on levels of income, employment, education, crime, health, housing, and access to public amenities such as a general practitioner, the post office, or public transportation (Department of Communities and Local Government, 2007; Scottish Government, 2006). IMD scores are derived differently for Scotland and England, and so they were considered separately in our study (the few participants from outside these countries were ignored in these comparisons).

The breakdown of the sample by age, sex, and IMD quintiles is given in Table S2 of the online Supplemental Materials. As expected, the sample was relatively affluent as a whole, but 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: $t(1152) = -0.100$, $p = .920$; age group: $t(1152) = -0.148$, $p = .883$.

## RT Tests and Procedures

RT testing was one of the activities run by the United Kingdom'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 RTs (SRTs) and 4-choice RTs (CRTs) were measured using an RT testing device, which was an upgraded version of an apparatus designed for and used in the HALS study (Cox, Huppert, & Whichelow, 1993; see Figure 1 in Deary et al., 2001, for an illustration of the device). The device had a liquid crystal display screen for the presentation of stimuli and five response buttons labeled 0–4. Buttons 1 and 2 were placed on the left side of the box and were to be operated with the middle and index fingers of the left hand, Buttons 3 and 4 were placed on the right and were to be operated with the index and middle fingers of the 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 eight 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 completed 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 s to 3 s. 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 they were under 16, the consent of their parents or guardians was obtained. Between one and six individuals could participate at any one time, with most tests being carried out in groups of three or four.

## Data Preparation

We identified and removed all prepresses, that is, responses made before the appearance of the stimulus. There were 391 trials with prepresses 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 cutoff 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 standard deviations 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. This procedure resulted in the removal of 235 individual trials for SRT and 198 for CRT. The cutoff values used for trimming were less stringent than in many other studies using RT data, in which the lower cutoff of 150 ms was used regardless of the task and the upper cutoff was determined by 3 standard deviations above the mean (e.g., Hultsch et al., 2002; Williams et al., 2005). It was our intention to perform only minimal trimming of the 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, we identified cases in which more than 25% of an individual's trials were removed in either RT task, following the previously mentioned exclusions. These cases were removed completely because a high frequency of prepresses 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 five 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 trials and 0.4% of CRT trials were removed. Following these exclusions, means and ISDs were recalculated for each test and each participant from the trials deemed valid. The final step was to exclude nine participants whose RTs fell beyond 3 interquartile ranges above their age group–specific 75th percentile.

To obtain a simple measure of RT intraindividual variability, which is independent of the mean RT, the RT coefficient of variation (RT CV) was calculated from the trimmed RT data: (RT ISD/mean RT) $\times$ 100.

## Statistical Analyses

Analyses were performed using SAS software (Version 9.2). 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 (age 18 and above). For these analyses, age was centered on the relevant mean (15.35 years for the whole sample, 9.73 for those under 18, and 42.03 for individuals 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 normalize the distribution and to reduce the heteroscedasticity. The optimal transformation parameter for each model was determined using PROC TRANSREG.

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-aged adults, and older adults) are presented in Table 1. The RT means and ISDs are generally largest in the youngest group, intermediate in the oldest group, and smallest in the middle group. Interindividual (between-subjects) variability in mean RT and RT ISD followed a similar pattern, and we tested the observed differences in variances for significance using Levene's test. Children and adolescents were more diverse in their RT means and ISDs than were either of the adult groups (all $p$s $<.05$). 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 1 (see Table S1 in the Supplemental Materials for the number of participants in each band). 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. Intraindividual variability in CRT 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.

### Age and Sex Effects on RT Throughout the Life Span

Modeling was performed on the transformed data (as described in the Statistical Analyses subsection of the Method section), including polynomial terms in age, sex, and their interactions entered as potential predictors. For CRT, the number of errors was also included. Final models were obtained by stepwise elimination of nonsignificant terms.

The analyses were performed in two stages. First, we analyzed the data from the whole sample to determine whether there were any interactions between age and sex; second, we investigated sex differences in more detail by considering the younger (under 18) and older (18 and above) portions of our sample separately.

In the first instance, we fitted exploratory models to the whole sample data. Predicted values obtained from these models for SRT and CRT means, ISDs, CVs, and ISDs adjusted for RT mean were back-transformed to their original units and are shown in Figure S1 of the Supplemental Materials.

All measures showed a steep decrease in childhood until the 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.

Full models with the relevant statistics are given in Table S3 of the Supplemental Materials, and the results are only briefly summarized here. The effects of sex and age, with no significant interaction between them, were found for mean SRT and for SRT ISD adjusted for mean SRT. Males had slightly shorter SRTs than did females across the age range, and the association between age and SRT was best described by a quintic function. SRT ISD adjusted for mean was slightly larger for males and decreased linearly with age throughout the age range. For the remaining models, including SRT ISD, SRT CV, and all four CRT models, there was a significant interaction between age and sex, which, from visual inspection of the graphs in Figure S1 of the Supplemental Materials, appears to suggest that sex effects observed in adults are different from those in children and adolescents.

To directly test our hypothesis that sex differences will be apparent in the adulthood but not in the younger ages, the models were refitted separately on the younger (<18) and older (18 and above) portions of our sample. The cutoff age was chosen to make the older group comparable with that in Der and Deary's (2006) study.

Table 1
*Means and Standard Deviations (in Seconds) of Reaction Time (RT) Measures for Both Genders in Three Broad Age Groups*

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

*Note*. SRT = simple RT; ISD = intraindividual standard deviation; CRT = choice RT.

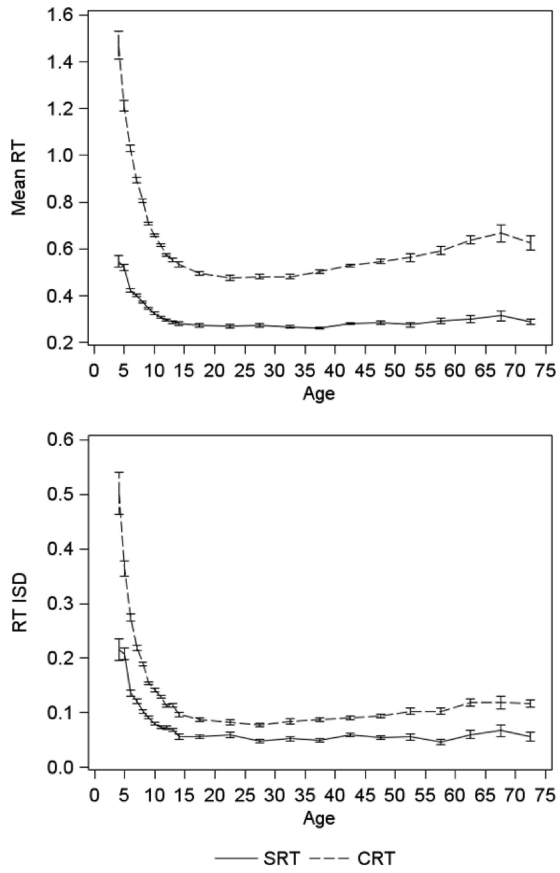


*Figure 1.* Means of simple reaction time (SRT) and choice reaction time (CRT) measures across age (grouped in 1-year bands up to 14, and in 5-year bands thereafter). Error bars represent group standard errors. ISD = intraindividual standard deviation.

## Age and Sex Effects in Childhood and Adolescence

Details of the models fitted to RT data of children and teenagers are presented in Table 2. Figure 2 shows the observed means and standard errors of RT means and ISDs for each age group (left panels) and back-transformed predicted values from the models (right panels). 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. A significant gender effect was found in SRT mean, where males were faster than females.

Models of intraindividual variability adjusted for the mean RT are also summarized in Table 2, and predicted values from these models are shown in Figure 3. With the effects of SRT mean controlled, the effect of age on SRT variability was cubic. There was also an effect of gender on RT intraindividual variability, with females slightly less variable than males. These age and sex effects were present in the model of SRT CV and of SRT ISD controlling for mean SRT. As far as CRT intraindividual variability is concerned, females were again found to be less variable using both methods of controlling for RT mean, and there was a quadratic effect of the number of errors. One notable difference 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 summarized in Table 3. Figure 4 shows the means and standard errors by age group in 5-year bands (left panels) and predicted values transformed back to their original units (right panels). 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 for 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 middle-aged people; however, the Age × Sex interaction was not significant.

Predicted values of intraindividual variability adjusted for mean RT are presented in Figure 5 (see also Table 3 for the models statistics). SRT CV did not vary with age or sex, whereas CRT CV was a linear function of age, with women slightly more variable than men. 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.

# Discussion

The discussion of our findings are presented in three sections: Sex Differences in RT, Age Effects, and RT Intraindividual Variability Controlling for RT Mean.

## Sex Differences in RT

The key finding of the present study is that, as hypothesized, greater intraindividual variability 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, 2005; 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 those under 18 and more variable RTs of adult females, is in line with Deary and Der's (2005) 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 cognitive performance of the genders from early adulthood onward.

Table 2
*Summary Statistics of Reaction Time (RT) Models for Individuals Below Age 18 Years*

| RT measure and parameter | Estimate | *SE* | *F* | *p* |
|---|---|---|---|---|
| Mean SRT | | | | |
| Intercept | −1.703888702 | 0.01568759 | | |
| Age | −0.133129629 | 0.00360536 | 1,363.49 | <.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 | 4,744.49 | <.001 |
| Age$^2$ | 0.006216372 | 0.00047116 | 174.07 | <.001 |
| No. errors | −0.040732910 | 0.00630167 | 41.78 | <.001 |
| No. 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 |
| No. 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 |
| No. errors | 0.033123963 | 0.00958356 | 11.95 | .001 |
| No. 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 | 3,212.06 | <.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 |
| No. errors | 0.043475335 | 0.00733225 | 35.16 | <.001 |
| No. errors$^2$ | −0.003542471 | 0.00111079 | 10.17 | .002 |

*Note.* SRT = simple RT; ISD = intraindividual standard deviation; CV = coefficient of variation; CRT = choice RT.

Considering cognition more broadly, sex hormones have been proposed as a potential source of sex differences in performance (for review, see Erlanger, Kutner, & Jacobs, 1999). One potential mechanism by which sex hormones could exert their effect is through acting directly on human brain functioning (see McEwen, 2001). Interestingly, Hampson (1990) demonstrated that during phases of the menstrual cycle, when estrogen levels are high, performance of females on tasks on which males tend to perform better (e.g., spatial ability) is poorer, and performance on tasks that females tend to score better on (e.g., articulatory skills) is better. In terms of the potential variability-producing mechanisms of estrogen, it could affect attentional systems, as estrogen receptors can be found in the cingulate cortex (McEwen, 2001). It could also act on the cholinergic system, as the dopaminergic system in particular

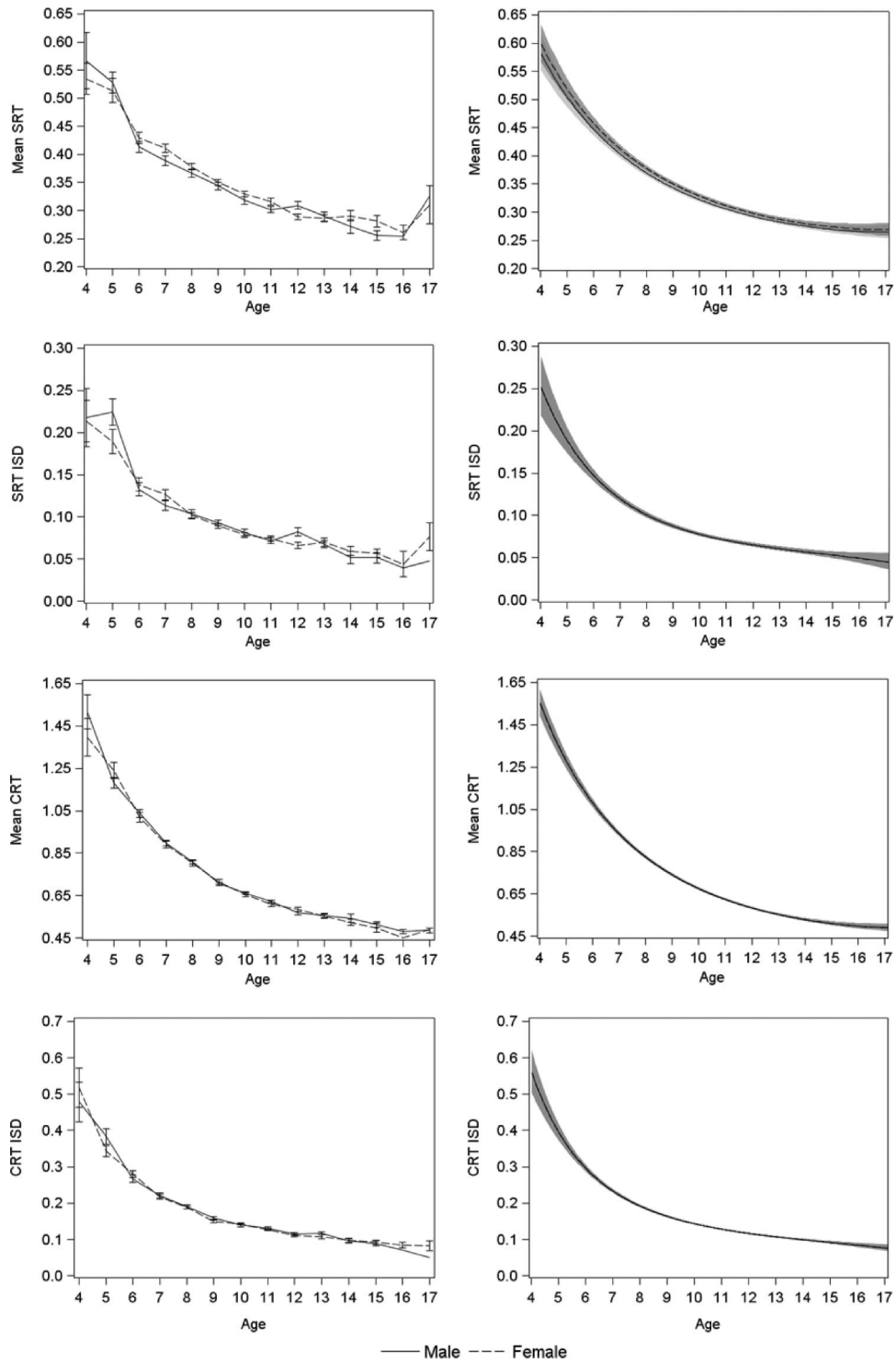


*Figure 2.* Observed and predicted values of reaction time measures for males and females under 18 years of age. Left panels: Means and standard errors (in 1-year age bands). Right panels: Predicted values and their corresponding 95% confidence interval bands. SRT = simple reaction time; ISD = intraindividual standard deviation; CRT = choice reaction time.

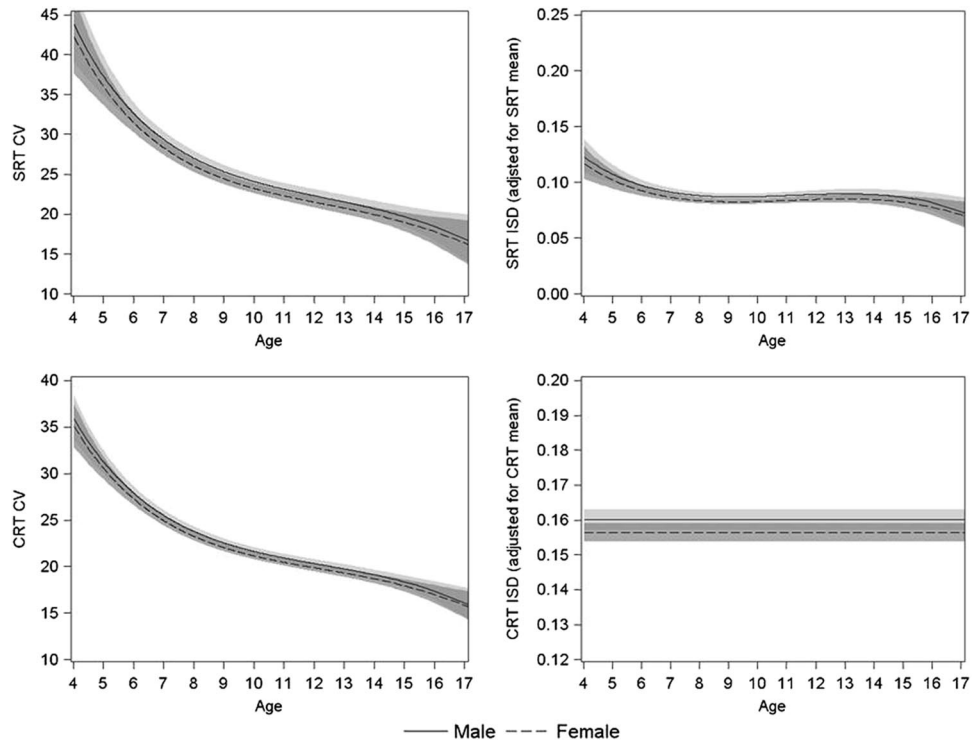


*Figure 3.* Predicted values and their corresponding 95% confidence interval bands of simple reaction time (SRT) and choice reaction time (CRT) coefficients of variation (CV) and intraindividual standard deviation (ISD) from models controlling for mean reaction time for participants under age 18.

is thought to affect age-related variability in cognition (Li, Lindenberger, & Sikstrom, 2001). Although there are plausible theoretical mechanisms linking estrogens to variability, empirical studies to date do not support the association. For example, there is little evidence of the effect of estrogen supplementation (through hormone replacement therapy [HRT]) on RT intraindividual variability (Low, Anstey, Jorm, Christensen, & Rodgers, 2006; Wegesin & Stern, 2004). The level of testosterone has been linked with RT performance in men (Müller, 1994). Men with lower salivary testosterone concentration had slower and more variable RTs. However, in a sample including men and women (Fontani, Lodi, Felici, Corradeschi, & Lupo, 2004), testosterone was associated with only mean RT and not variability. According to Fontani et al. (2004), testosterone could have a direct effect on the brain, as testosterone receptors can be found in the hippocampus, or it could aromatize to estradiol. Taken together, empirical findings regarding sex hormone levels and RT performance do not give a clear indication of hormonal effects on intraindividual variability. Both estrogen and testosterone studies have their limitations. The former focused on HRT and so may not be applicable to endogenous estrogen effects. The latter were limited by small sample sizes and require replication.

As well as influencing the functioning of the brain, hormones might affect the brain structure. 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 gray 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 that may underlie sex differences in RT variability.

Because the hormone levels were not tested in our study, the explanations just given are only speculative. It is possible that the differences found have different roots, for example, different strategies used by males and females. One study suggested that females have a slower response on the first trial in an RT test but not the subsequent ones and that this creates an overall sex difference in intraindividual variability (Reimers & Maylor, 2006). Another explanation could be in terms of cohort effects. Because our study was cross-sectional, it may be that the differences we observed between younger and older participants are a reflection of genuine change in the 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 have 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 the bulk of previous literature, which found men to

Table 3
*Summary Statistics of Reaction Time (RT) Models for Individuals 18 Years of Age 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 | 5,3971.00 | <.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$^2$ | 0.0000697 | 0.0000337 | 4.27 | .040 |
| No. 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 |
| No. 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 |
| No. errors | 0.0208817 | 0.0075432 | 7.66 | .006 |

*Note*. SRT = simple RT; ISD = intraindividual standard deviation; CV = coefficient of variation; CRT = choice RT.

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, 2005; 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 male and female adults. The two RT tasks used in the present study (SRT and CRT) were 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, 2005; Der & Deary, 2006), all of which used a 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 intraindividual 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 underreported. 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, 2005), HALS (Der & Deary, 2006), and the science festivals sample in the present study. These are not accompanied by reports of the opposite effect, which should not be subject to publication bias. Therefore, although we would encourage attempts to replicate our findings, we think that the effect, alas small and problematic to explain, is genuine.

Noble et al. (1964) 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.

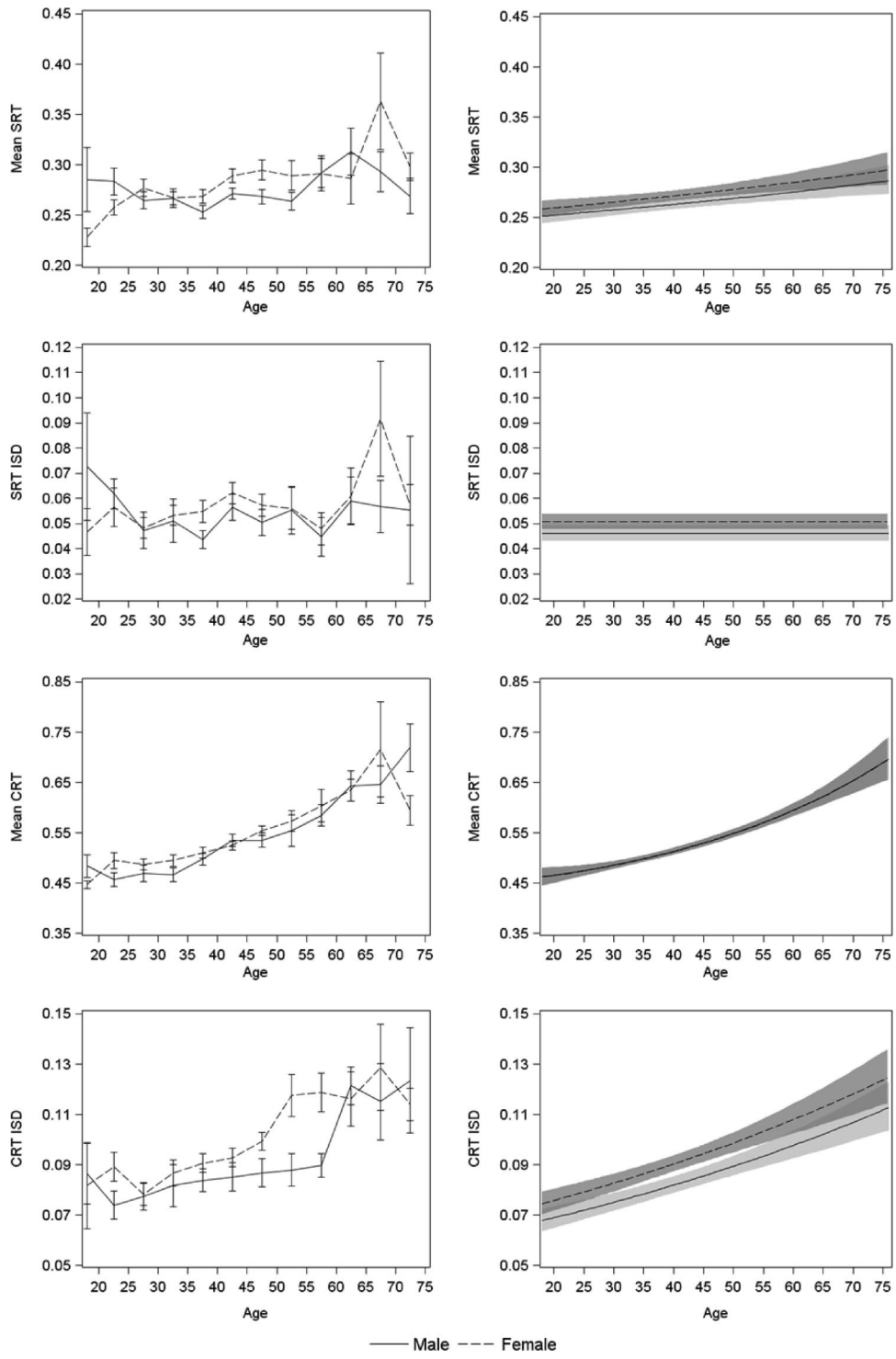


*Figure 4.* Observed and predicted values of reaction time measures for males and females 18 years of age or over. Left panels: Means and standard errors (the first age band includes ages 18 and 19; the remaining ones are 5-year bands). Right panels: Predicted values and their corresponding 95% confidence interval bands. SRT = simple reaction time; ISD = intraindividual standard deviation; CRT = choice reaction time.

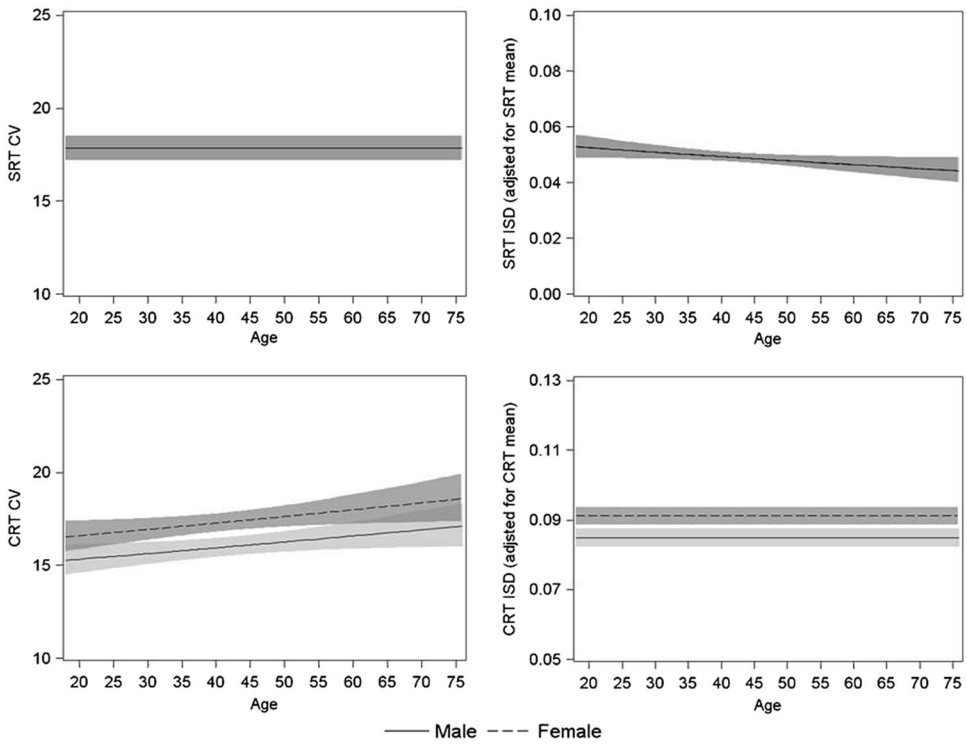


*Figure 5.* Predicted values and their corresponding 95% confidence interval bands of simple reaction time (SRT) and choice reaction time (CRT) coefficients of variation (CV) and intraindividual standard deviation (ISD) from models controlling for mean reaction time for participants 18 years of age and over.

## Age Effects

The findings in terms of age effects on RT mean and intraindividual variability are in line with the majority of evidence from the life span studies. Like Koga and Morant (1923), Wilkinson and Allison (1989), Williams et al. (2007), and others, we found that RT trajectories are characterized by a rapid improvement in childhood and adolescence and a much slower deterioration through the adult age range. Some of these effects may be due to a difference in ability to sustain attention at different ages. For example, West et al. (2002) suggested attentional lapses as a possible explanation for the greater variability in the older ages. 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). Our work added to the existing knowledge by comparing and contrasting, within a single study, different measures of RT intraindividual 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 our finding is consistent with that of Fozard et al. (1994), who reported larger age effects in a more complex task.

The increase in RT intraindividual variability in old age found in our 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. In Williams et al.'s study, participants had to ignore the location of the stimulus (left, right, or center) 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 tasks (e.g., West et al., 2002). Therefore, one 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 center of the screen). It is possible that the mere presence of incongruent trials interspersed throughout a test affects responses even on neutral trials.

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. First, 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. Second, 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, and 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 had 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.

## RT Intraindividual Variability Controlling for RT Mean

Sex effects were apparently affected by adjustment of intraindividual standard deviations for RT mean. 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 variability 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. The findings regarding sex effects on RT variability adjusted for RT mean among the adult portion of our sample agree with those reported by Der and Deary (2006). However, they are in opposition to the findings reported by Hultsch et al. (2002), where females had greater mean-adjusted variability in SRT but not CRT.

Although the pattern of sex differences in intraindividual variability adjusted for mean RT in our study may appear 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 favoring 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 modeled with mean RT will remain greater among women.

In terms of the relationship between age and RT intraindividual variability, 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 modeled 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 (2005) 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 age effects on intraindividual variability are found depends on the method of controlling for RT mean that 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 on the comparability of the two approaches. Since there has been no consensus on which method is best, a variety of methods are currently employed by the investigators in the field. Some authors have opted for a simple CV (Gorus, De Raedt, & Mets, 2006), others have used mean RT as a covariate in their models (Shammi, Bosman, & Stuss, 1998), and still others have regressed 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 those under 18. This finding supports claims by 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 nonlinearity among the older participants in our 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.

## 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 is found in 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 mean deprivation scores 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 that 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, our under 18 sample is among the largest out of the RT studies we reviewed and is the largest one on which sex differences in RT variability were investigated. The smaller number of adults, particularly of those over age 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 we had more power to detect any differences.

The cross-sectional design of the present study is a limitation: The data do not provide information about the developmental change in RT but only about differences between participants of different ages. Moreover, any patterns found in the present study 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 participants 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 have suggested that sex difference in RT have been 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 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 among young children might have been overestimated. In order to be able to meaningfully compare the amount of intraindividual variability observed in children and older adults, stimuli easily identifiable by both groups should be used (e.g., colored 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 an unusually large number of children. The key finding of the present study is that greater intraindividual variability of females in SRT and CRT was observed in adults but not in children. 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 did 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.

## References


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

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. doi:10.1111/j.1469-7610.2006.01611.x

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

Bunce, D., Tzur, M., Ramchurn, 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, 63,* P67–P74. doi:10.1093/geronb/63.2.P67

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. doi:10.1080/13803390490918318

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. doi:10.1037/0882-7974.16.4.588

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi:10.1037/0033-2909.112.1.155

Cox, B. D., Huppert, F. A., & Whichelow, M. J. (1993). *The Health and Lifestyle Survey: Seven years on.* Aldershot, United Kingdom: Dartmouth.

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, Neuroscience, and Cognition, 12,* 187–215. doi:10.1080/13825580590969235

Deary, I. J., Der, G., & Ford, G. (2001). Reaction times and intelligence differences: A population-based cohort study. *Intelligence, 29,* 389–399. doi:10.1016/S0160-2896(01)00062-9

Deary, I. J., Irwing, P., Der, G., & Bates, T. C. (2007). Brother-sister differences in the g factor in intelligence: Analysis of full, opposite-sex siblings from the NLSY1979. *Intelligence, 35,* 451–456. doi:10.1016/j.intell.2006.09.003

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,* 552–557. doi:10.1093/cercor/11.6.552

Department of Communities and Local Government. (2007). *Indices of Deprivation 2007* [Data file]. Retrieved from http://www.communities.gov.uk/communities/research/indicesdeprivation/

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. doi:10.1037/0882-7974.21.1.62

Eckert, H. M., & Eichorn, D. H. (1977). Developmental variability in reaction time. *Child Development, 48,* 452–458. doi:10.2307/1128638

Erlanger, D. M., Kutner, K. C., & Jacobs, A. R. (1999). Hormones and cognition: Current concepts and issues in neuropsychology. *Neuropsychology Review,* 9, 175–207. doi:10.1023/A:1021634622577

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* (pp. 159–175). London, England: Churchill Livingstone.

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. doi:10.1007/s00421-004-1108-3

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: Series B: Psychological Sciences, 49,* P179–P189. doi:10.1093/geronj/49.4.P179

Gilbert, J. A. (1894). Researches on the mental and physical development of school-children. In Yale University, Psychological Laboratory (Ed.), *Studies From the Yale Psychological Laboratory* (Vol. 2, pp. 40–100). Retrieved from http://www.archive.org

Goodenough, F. L. (1935). The development of the reactive process from early childhood to maturity. *Journal of Experimental Psychology, 18,* 431–450. doi:10.1037/h0062460

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

Halpern, D. F. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hampson, E. (1990). Variations in sex-related cognitive abilities across the menstrual cycle. *Brain and Cognition, 14,* 26–43. doi:10.1016/0278-2626(90)90058-V

Herlitz, A., & Lovén, J. (2009). Sex differences in cognitive functions. *Acta Psychologica Sinica, 41,* 1081–1090. doi:10.3724/SP.J.1041.2009.01081

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, 57,* P101–P115. doi:10.1093/geronb/57.2.P101

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, NY: Psychology Press.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). Berlin, Germany: Springer-Verlag. doi:10.1007/978-3-642-68664-1_4

Jensen, A. R. (1992). The importance of intraindividual variation in reaction time. *Personality and Individual Differences, 13,* 869–881. doi:10.1016/0191-8869(92)90004-9

Kalb, R., Jansen, S., Reulbach, U., & Kalb, S. (2004). Sex differences in simple reaction tasks. *Perceptual and Motor Skills, 98,* 793–802. doi:10.2466/pms.98.3.793-802

Koga, Y., & Morant, G. M. (1923). On the degree of association between reaction times in the case of different senses. *Biometrika, 15,* 346–372. doi:10.1093/biomet/15.3-4.346

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. doi:10.1111/j.1467-9450.1985.tb01175.x

Landauer, A. A., Armstrong, S., & Digwood, J. (1980). Sex differences in choice reaction time. *British Journal of Psychology, 71,* 551–555. doi:10.1111/j.2044-8295.1980.tb01766.x

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. doi:10.1016/j.neuroimage.2007.03.053

Li, S.-C., Lindenberger, U., & Sikstrom, S. (2001). Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences, 5,* 479–486. doi:10.1016/S1364-6613(00)01769-1

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. doi:10.1016/j.maturitas.2005.09.001

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. Retrieved from http://jap.physiology.org/content/91/6/2785.full.pdf

Müller, M. J. (1994). Salivary testosterone and simple reaction time parameters. *Neuropsychobiology, 30,* 173–177. doi:10.1159/000119157

Noble, C. E., Baker, B. L., & Jones, T. A. (1964). Age and sex parameters in psychomotor learning. *Perceptual and Motor Skills, 19,* 935–945. doi:10.2466/pms.1964.19.3.935

Office for National Statistics. (2006). *National statistics postcode directory* [Data file]. Census Geography Data Unit (UKBORDERS), EDINA (University of Edinburgh). Retrieved from http://edina.ac.uk/ukborders/

Reed, T. E., Vernon, P. A., & Johnson, A. M. (2004). Sex differences in brain nerve conduction velocity in normal humans. *Neuropsychologia, 42,* 1709–1714. doi:10.1016/j.neuropsychologia.2004.02.016

Reimers, S., & Maylor, E. A. (2006). Gender effects on response time variability and trial-to-trial performance: Reply to Deary and Der (2005). *Aging, Neuropsychology, and Cognition, 13,* 479–489. doi:10.1080/138255890969375

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. doi:10.1037/a0017799

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. doi:10.1076/anec.5.1.1.23

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. doi:10.1097/01.psy.0000195867.66643.0f

Silverman, I. W. (2006). Sex differences in simple visual reaction time: A historical meta-analysis. *Sex Roles, 54,* 57–68. doi:10.1007/s11199-006-8869-6

Wegesin, D. J., & Stern, Y. (2004). Inter- and intraindividual variability in recognition memory: Effects of age and estrogen use. *Neuropsychology, 18,* 646–657. doi:10.1037/0894-4105.18.4.646

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. doi:10.1006/brcg.2001.1507

Wilkinson, R. T., & Allison, S. (1989). Age and simple reaction time: Decade differences for 5,325 subjects. *Journal of Gerontology, 44,* P29–P35. doi:10.1093/geronj/44.2.P29

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. doi:10.1037/0894-4105.19.1.88

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. doi:10.1080/13825580600584590

Received July 20, 2010
Revision received January 12, 2012
Accepted January 19, 2012 ■