Mental Health and Cognitive Function in Adults Aged 18 to 92 Years

David Bunce,1,3 Maya Tzur,2 Anusha Ramchurn,3 Felicity Gain,3 and Frank W. Bond3

1Centre for Cognition and Neuroimaging, Brunel University, United Kingdom.
2Department of Psychiatry, University of Cambridge, United Kingdom.
3Department of Psychology, Goldsmiths College, University of London, United Kingdom.

We investigated mental health and cognitive function in 195 community-dwelling adults aged 18 to 92 years (M = 46.64). We assessed several cognitive domains, including psychomotor, executive function, and episodic memory. We found a significant Age × Mental Health interaction in relation to within-person (WP) variability (trial-to-trial variability in reaction time performance) in a four-choice psychomotor task and a Stroop task, but not in relation to mean reaction time measures from those tasks. Poorer mental health was associated with greater WP variability in older adults. We did not find this effect in relation to memory. The findings suggest that measures of WP variability may be sensitive to relatively subtle effects associated with age and poor mental health, and that they provide valuable insights into cognitive function in old age.

Key Words: Intraindividual—Variability—Depression—Anxiety.

Although investigation of reaction time (RT) mean performance for a given cognitive task can provide valuable information, there are good reasons to consider the within-person (WP) variability of RTs across trials of that task. The suggestion that WP variability (also referred to as intraindividual variability) is indicative of neurobiological disturbance or decline (Hendrickson, 1982; Hultsch & MacDonald, 2004; Li & Lindenberger, 1999) is supported by an accumulating body of research. For example, WP variability is greater with older age (e.g., Anstey, 1999; Bunce, MacDonald, & Hultsch, 2004; Hultsch, MacDonald, & Dixon, 2002; West, Murphy, Armilio, Craik, & Stuss, 2002), and also in relation to a range of conditions involving some form of neurobiological disturbance (see MacDonald, Bäckman, & Nyberg, 2006, for review). Relative to matched controls, WP variability is greater in the presence of mild cognitive impairment or mild dementia (Christensen et al., 2005; Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000; Strauss, Bielak, Bunce, Hunter, & Hultsch, 2007), traumatic brain injury (e.g., Stuss, Murphy, Binns, & Alexander, 2003; Stuss, Pogue, Buckle, & Bondar, 1994), Parkinson’s disease (Burton, Strauss, Hultsch, Moll, & Hunter, 2006), and epilepsy (Bruhn & Parsons, 1977). Adding to this behavioral work, functional magnetic resonance imaging (fMRI) work in young adults (Bellgrove, Hester, & Garavan, 2004) suggests that WP variability in a go–no–go response paradigm was systematically related to activity in the frontal cortex. A structural MRI study by Anstey and colleagues (2007) in adults aged 60 to 64 years found that WP variability was associated with corpus callosum size in individuals with mild cognitive disorders but not in individuals in a normative sample. Another investigation by this group in healthy older adults found that the extent of frontal white matter lesioning was associated with measures of WP variability but not with measures of processing speed (mean RT), memory, or global cognition (Bunce et al., 2007). Together, this behavioral and brain imaging work suggests that WP variability (a) reflects neurobiological disturbance or decline and (b) is associated with systematic brain activity or structures in the frontal cortex.

As research suggests that measures of WP variability may reflect disease pathology that involves the central nervous system, in the present study we addressed questions that have received little attention to date. That is, how far are measures of WP variability sensitive to mild psychopathology? Does age have any bearing on associations where they exist? Specifically, we were interested in the possibility that persons experiencing poor mental health in the form of mild depression or anxiety, social dysfunction, and loss of confidence, although functioning apparently normally in their everyday lives, were more variable in their responding than were persons unaffected by such mild psychopathology. Additionally, it was important to assess whether age had any bearing on these associations.

Theoretically, there are both cognitive and neurobiological reasons to expect depression or anxiety, and relatedly, social dysfunction and loss of confidence, to affect cognitive performance. Cognitive perspectives draw largely on resource theory (e.g., Kahneman, 1973), and they emphasize attentional resource reductions and task-irrelevant thoughts. Specifically, the cognitive capacity available to process information may be reduced (but not eliminated) in depressed persons, and this together with the tendency to deploy the remaining resources toward depression-related thoughts results in cognitive deficits, particularly in more effortful processing (Hartlage, Alloy, Vazquez, & Dykman, 1993). Similarly, elevated anxiety may affect cognitive performance by detrimentally influencing executive control mechanisms, and working memory in particular, such that reduced information-processing capacity and efficiency results (e.g., Eysenck & Calvo, 1992).

Neurobiological accounts include the glucocorticoid hypothesis (e.g., Sapolsky, 1999), which suggests that elevated stress levels trigger the secretion of glucocorticoids. Although this may serve beneficially in adapting to acute stressors, longer term secretion may be harmful to the nervous system, particu-
icular the hippocampus. Empirically, several studies have demonstrated cortisol-related deficits in memory (e.g., Lupien & McEwen, 1997) and also global cognition (Magri et al., 2006). Research has also identified a high density of corticosteroid receptors in the frontal lobes (Murros, Fogelholm, Kettunen, & Vuorela, 1993), the brain structure held to support executive function (Cabeza & Nyberg, 2000). As glucocorticoid exposure increases in old age (Sapolsky, 1999), and anxiety and depression frequently occur together, this raises the possibility that older persons experiencing anxiety or depression may be at greater risk of cognitive deficits, particularly in domains reliant on frontal systems. This possibility was underlined by recent work (Stawski, Sliwinski, & Smyth, 2006) showing stress-related cognitive interference to be associated with poorer performance on working memory and episodic memory tasks in older adults, and also research suggesting measures of executive function to account for depression-related cognitive deficits in older adults (Elderkin-Thompson, Mintz, Haroon, Lavretsly, & Kumar, 2006; Sheline et al., 2006).

Although studies exist that suggest an association between mental health and WP variability, the amount of research in the area is limited. For instance, a recent investigation found a significant positive association between WP variability and negative affect in college students recording higher neuroticism scores (Robinson, Wilkowski, & Meier, 2006). A further study at the WP level that took age into account examined day-to-day variability in stress and WP variability among young adults aged 18 to 24 years, and older adults aged 66 to 96 years (Sliwinski, Smyth, Hofer, & Stawski, 2006). For a two-back working memory task, a significant Stress × RT interaction suggested that higher WP stress was associated with slower, but not faster, RTs. Although there was a trend suggesting this effect was stronger in older adults, it was not significant. Together, even though these investigations point to an association between poor mental health and intraindividual variability, it remains unclear as to whether this effect is stronger in older adults relative to younger adults, and also whether measures of WP variability reveal associations when mean RT measures do not.

In the present study, we recorded self-reports of mental health through the 12-item General Health Questionnaire (GHQ-12; Goldberg, 1978). The original 60-item GHQ was developed to measure general health and included items relating to both psychiatric and somatic complaints. A shorter 12-item version was later developed that specifically assesses minor psychiatric disorders relating to state (rather than trait) feelings. Research investigating the factor structure of this version (e.g., Weneke, Goldberg, Yalcin, & Üstün, 2000; also see Makikangas et al., 2006) suggests the scale to measure constructs relating to anxiety and depression, social dysfunction, and loss of confidence. The measure has been used extensively in various settings, and it effectively identifies persons who would be classified as suffering a minor psychiatric disorder (see Goldberg et al., 1997). Importantly from the present perspective, a study in 5,438 persons (Goldberg et al.) showed that age did not affect the scale’s validity.

In addition, a correlation of 0.69 was reported between the GHQ-12 and the Beck Anxiety Inventory in an outpatient sample with anxiety disorders (Gao et al., 2004), and our own work in a community-based sample (Bond & Bunce, 2000) produced a mean correlation of 0.54 over four time points with the Beck Depression Inventory. On the basis of earlier research, we predicted that poor mental health, as recorded by the GHQ, would be associated with poorer cognitive performance. Because of the well-established evidence of age-related changes in the frontal cortex (Raz & Rodrigue, 2006; West, 1996) and the theoretical rationale presented earlier whereby the mechanisms by which stress and depression affected cognitive function involved frontal regions and, relatedly, executive processes, we expected that associations, where they existed, would be particularly strong in older adults.

We chose our cognitive variables to measure specific cognitive domains. We were particularly interested in psychomotor performance, because earlier studies (Bunce et al., 2004; MacDonald, Hultsch, & Bunce, 2006) have shown that older adults exhibit a greater response variability in such tasks. Our first aim was to compare standard measures of mean RT with WP variability metrics while taking age and mental health into account. Was it the case that measures of WP variability were sensitive to subtle, but nonetheless important, effects that more commonly used metrics were not? Our second aim, given the empirical research described earlier, was to assess episodic memory and executive function in the expectation that they too would be sensitive to Age × Mental Health associations. Indeed, as theoretical accounts of the mechanisms by which mental health influences cognition commonly suggest frontally modulated or executive control mechanisms, we reasoned that where Age × Mental Health associations were found in relation to memory and psychomotor tasks, they would be accounted for by independent measures of executive function. Our third aim, therefore, was to test this possibility.

### Methods

#### Participants

Through local organizations (e.g., bridge clubs) and researcher contacts, we recruited healthy persons living in the community for the study. During the initial screening, participants provided details of their health (including psychiatric conditions) and medications. Two persons reported neurological conditions (Parkinson’s disease and minor stroke); we did not test them further. The remaining 195 persons (53% women, 47% men) were between 18 and 92 years of age (M = 46.64, SD = 19.10). We screened those individuals older than 65 years for cognitive impairment by using the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). In this study, we included only those individuals recording a score of 26 or higher on that examination. Apart from conditions typically found in older persons (e.g., blood pressure or prostate problems), all participants were free of major health or mental health disorders. None reported using antidepressants. We measured verbal IQ through the National Adult Reading Test (NART; Nelson, 1982), and here we report estimated full-scale IQ (M = 116.46, SD = 8.49). For each decade, descriptive data for biographical, NART, and GHQ variables are presented in Table 1.
Mental Health

GHQ-12.—The GHQ-12 (Goldberg, 1978) is a 12-item scale that assesses mental health through various questions (e.g., Have you recently... “lost much sleep over worry,” “felt that you couldn’t overcome your difficulties,” “been losing confidence in yourself”). We utilized the widely used Likert method of scoring, in which each item was scored 0 (not at all) to 3 (much more than usual). We report the scale mean, in which higher scores indicate poorer mental health. A Cronbach alpha of $\alpha = .89$ suggested that internal consistency was good for this scale.

Cognitive Tasks

Episodic memory.—We created a list of 40 semantically unrelated nouns (frequency, $M = 126.50$). At study, we had 20 target words presented at a rate of 5 s per word on a personal computer (PC) screen. At test following several other tasks, we had the 20 target words randomly intermixed with 20 distracter words and individually presented on the PC screen for 5 s each. Researchers instructed the participants to respond “yes” for target words and “no” for distracter words by using designated keyboard keys. We recorded the hits and false alarms. In order to take response bias into account, here we report on adjusted hits (i.e., hits minus false alarms).

Executive function.—Recent research suggests that executive function is characterized as involving separable but related processes of switching between mental sets, monitoring and updating working memory, and inhibiting prepotent responses (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Therefore, we selected our measures of executive function to reflect these constructs. For switching, we used the Alternate Category Task (Parkin, Walter, & Hunkin, 1995), which required the alternate generation of as many unique animal and country names as possible within 1 minute. We awarded a point for each successful animal–country switch. We summed the points to give a total score that reflected set-switching performance.

For updating, we presented 105 digits from 1 to 9 individually on a PC screen at a rate of 1 s each. Embedded at random intervals were 11 strings of three odd numbers. We instructed participants to press the keyboard spacebar each time they detected a string of three odd numbers. Thus, for each digit presentation, the individual had to update working memory in order to detect three continuous odd-number presentations. We recorded the correct detections (hits) and false alarms. Here, in order to take response bias into account, we report adjusted hits as a summary measure of updating working memory.

For inhibition, we used a version of the Stroop color–word task administered on a PC as a measurement. Participants were required to respond to the ink color (red, blue, yellow, or green) of the printed words red, blue, yellow, and green by using the appropriate keyboard key. We administered 16 practice trials, followed by 96 pseudorandomized test trials, half of which were congruent (word–color matched) and half of which were incongruent (word–color not matched). We computed measures of mean RT and WP variability for this task (see the paragraphs that follow).

Psychomotor.—We administered two-choice RT (2-CRT) and four-choice RT (4-CRT) tasks by means of a PC. In the two-choice version of the task, a black disk (0.5 cm in diameter) was presented pseudorandomly either left or right of a central fixation cross. In the four-choice version of the task, the black disk appeared in one of four locations (lower or upper left or right). Participants were required to press designated keyboard keys that mapped spatially onto the position of the disk on the screen as quickly but as accurately as possible. We counterbalanced the tasks across participants, and in each condition we administered 20 practice trials prior to 100 test trials.

Computation of Mean RT and WP Variability

For correct trials of the Stroop and 2-CRT and 4-CRT tasks we computed mean RT, and the intraindividual standard deviation (ISD) as a measure of WP variability, following the procedures of Hultsch and colleagues (2002). We inspected data distributions for the respective tasks for any extremely fast or slow latencies that may reflect various sources of error (e.g., accidental key presses, task interruption). In order to eliminate such trials, we set a lower boundary of 150 ms, and an upper boundary of the individual mean RT $+ 3 SD$, beyond which trials were excluded. We replaced the eliminated RTs with the age group mean RT for that particular task. For the Stroop task, no trials were replaced at the lower bound, but for the upper bound, 1.56% and 1.25% of trials, respectively, for the congruent and incongruent conditions were replaced. For 2-CRT and 4-CRT tasks, 3.00% and 1.80% of trials, respectively, were replaced for this reason. As removing outliers and replacing missing values in this way reduces variability, the

Table 1. Descriptive Variables According to Age Decade

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Age</th>
<th>Gender</th>
<th>Education</th>
<th>NART</th>
<th>GHQ-12</th>
</tr>
</thead>
<tbody>
<tr>
<td>18–30</td>
<td>22.69</td>
<td>32</td>
<td>14.28</td>
<td>116.61</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>4.02</td>
<td></td>
<td>1.99</td>
<td>6.90</td>
<td>.50</td>
</tr>
<tr>
<td>31–40</td>
<td>35.87</td>
<td>13</td>
<td>14.22</td>
<td>114.00</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>3.08</td>
<td></td>
<td>2.80</td>
<td>10.11</td>
<td>.49</td>
</tr>
<tr>
<td>41–50</td>
<td>45.79</td>
<td>13</td>
<td>13.50</td>
<td>112.32</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>2.81</td>
<td></td>
<td>3.43</td>
<td>9.16</td>
<td>.51</td>
</tr>
<tr>
<td>51–60</td>
<td>55.35</td>
<td>21</td>
<td>12.95</td>
<td>114.86</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>3.07</td>
<td></td>
<td>2.87</td>
<td>9.94</td>
<td>.47</td>
</tr>
<tr>
<td>61–70</td>
<td>64.91</td>
<td>16</td>
<td>13.35</td>
<td>120.56</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>2.59</td>
<td></td>
<td>2.93</td>
<td>5.94</td>
<td>.51</td>
</tr>
<tr>
<td>70+</td>
<td>79.32</td>
<td>10</td>
<td>12.00</td>
<td>120.89</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>6.60</td>
<td></td>
<td>3.16</td>
<td>5.61</td>
<td>.54</td>
</tr>
</tbody>
</table>

Note: For the age groups, given as years, numbers are as follows: 18–30, n = 54; 31–40, n = 23; 41–50, n = 28; 51–60, n = 37; 61–70, n = 34; 70+, n =19. Gender shows the number of women in the Age subgroup; Education refers to the participants’ years of education. NART = National Adult Reading Test (estimated full-scale IQ); GHQ-12 = 12-item General Health Questionnaire (scale average score).
procedure represents a conservative approach to the study of WP variability.

For each task, we computed mean RT from the raw correct response latencies, and for WP variability, we computed the ISD. Because the raw ISD increases in direct relation to age-related increases in RT, it is appropriate to statistically partial out such systematic age differences from ISD measures. Otherwise, larger ISDs are confounded by older adults’ slower responding relative to that of younger adults. Similarly, trial-to-trial variance may reflect learning effects, or other systematic time-on-task influences. Therefore, we used a regression procedure (Hultsch et al., 2002) to partial out the effects of chronological age, and trial-related (i.e., trial number) variation in responding from ISDs, and their higher order interaction. Specifically, we regressed RTs on chronological age (entered as a continuous variable) and trial number to produce residuals that were statistically independent of differences stemming from either of those sources. We then computed the residuals to t scores (M = 50, SD = 10) to facilitate comparisons across tasks, and then we used them to compute ISDs. As nonsignificant correlations between GHQ and mean RT suggested that RTs did not vary according to GHQ (see Table 2), we did not enter GHQ scores into this procedure.

For the Stroop task, as congruent and incongruent conditions for the respective measures were highly correlated (mean RT = 0.93; WP variability = 0.82), we computed the mean of the two conditions as summary measures of mean RT and WP variability.

### Missing Data

Across the 195 participants, for missing data on a minority of variables (i.e., updating = 2.1%, inhibition = 0.5%), we imputed values with the EM algorithm in SPSS by using all of the variables in the present study (see Shafer & Graham, 2002).

### Results

Bivariate correlations together with means and standard deviations for the respective variables are presented in Table 2. In most cases, age correlated with the cognitive variables in the direction expected. Older age was significantly associated with lower recognition scores, slower RTs, and greater WP variability. Older age was also associated with higher NART scores. There were associations between executive function and WP variability in the direction expected; higher switching scores were significantly correlated with lower variability in both the 2-CRT and 4-CRT tasks, and also the Stroop task. In addition, higher updating scores were significantly associated with lower variability in the 4-CRT task. The correlation between age and GHQ scores was nonsignificant. Correlations between GHQ and cognitive variables were predominantly nonsignificant, the exception being that higher GHQ scores were associated with greater WP variability in the 4-CRT task. Associations between gender and the other variables were all nonsignificant.

We used a series of hierarchical multiple regression models to explore the relationship between age, mental health, and the cognitive variables. As NART scores were associated with both age and several of the cognitive measures, we computed all of the models having adjusted for that variable at Step 1. This procedure controls for the possibility that age differences in IQ underlie age differences in the cognitive variables, and, therefore, act as a confound. At Step 2, we entered chronological age and GHQ scores, followed by the Age × GHQ cross-product interaction term at Step 3. It is important that, if Step 3 significantly added to the variance (R²) explained in the cognitive variable having taken age and GHQ scores into account, then it would suggest that the strength of association between mental health and cognitive function varied according to age. As a result of the number of regression models examined, we set alpha conservatively at p < .01.

The results of the hierarchical multiple regressions are presented in Table 3. At Step 2, the beta weights suggest that the respective associations between age and GHQ scores and the other variables did not differ greatly from the bivariate correlations reported in Table 2. Age independently predicted Stroop and 4-CRT intraindividual variability. Of particular interest, though, was Step 3, where the Age × GHQ cross-product interaction term was entered. For the majority of variables, entry of this term did not add significantly to the variance explained in the cognitive measures. However, for the WP variability in the Stroop and 4-CRT tasks, that interaction

### Table 2. Bivariate Correlations Between Age, Cognitive, and Mental Health Variables

<table>
<thead>
<tr>
<th></th>
<th>M (SD)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>46.64 (19.2)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Gender</td>
<td>—</td>
<td>—</td>
<td>.08</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>NART</td>
<td>116.46 (8.49)</td>
<td>.16*</td>
<td>—.06</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Switching</td>
<td>20.03 (5.43)</td>
<td>—.08</td>
<td>—.07</td>
<td>.47**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Updating</td>
<td>9.20 (3.41)</td>
<td>—.08</td>
<td>.07</td>
<td>—.02</td>
<td>—.04</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4-CRT ISD</td>
<td>787 (216)</td>
<td>.47**</td>
<td>—.03</td>
<td>.18*</td>
<td>.08</td>
<td>—.20**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4-CRT (ms)</td>
<td>7.78 (3.67)</td>
<td>.31**</td>
<td>—.01</td>
<td>—.06</td>
<td>—.18*</td>
<td>—.09</td>
<td>.65**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Recognition</td>
<td>7.59 (5.69)</td>
<td>—.26**</td>
<td>.13</td>
<td>.19**</td>
<td>.34**</td>
<td>.00</td>
<td>—.02</td>
<td>—.23**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2-CRT (ms)</td>
<td>313 (90)</td>
<td>.48**</td>
<td>—.03</td>
<td>.03</td>
<td>—.13</td>
<td>—.17*</td>
<td>.66**</td>
<td>.50**</td>
<td>—.23**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2-CRT ISD</td>
<td>7.32 (5.30)</td>
<td>.37**</td>
<td>.01</td>
<td>—.12</td>
<td>—.29**</td>
<td>—.09</td>
<td>.49**</td>
<td>.54**</td>
<td>—.32**</td>
<td>.87**</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4-CRT (ms)</td>
<td>413 (142)</td>
<td>.51**</td>
<td>—.08</td>
<td>.16*</td>
<td>—.02</td>
<td>—.23**</td>
<td>.73**</td>
<td>.34**</td>
<td>—.10</td>
<td>.81**</td>
<td>.65**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4-CRT ISD</td>
<td>6.68 (4.79)</td>
<td>.47**</td>
<td>—.05</td>
<td>.05</td>
<td>—.16*</td>
<td>—.20**</td>
<td>.62**</td>
<td>.51**</td>
<td>—.27**</td>
<td>.81**</td>
<td>.78**</td>
<td>.84**</td>
<td>—</td>
</tr>
<tr>
<td>GHQ</td>
<td>0.93 (0.49)</td>
<td>—.02</td>
<td>.10</td>
<td>—.01</td>
<td>—.06</td>
<td>.04</td>
<td>.09</td>
<td>.13</td>
<td>.13</td>
<td>.13</td>
<td>.10</td>
<td>.10</td>
<td>.22**</td>
</tr>
</tbody>
</table>

**Notes:** Gender 1 = Men, 2 = Women; NART = National Adult Reading Test; 2-CRT or 4-CRT = two- or four-choice reaction time; ISD = intraindividual standard deviation; GHQ = General Health Questionnaire-12.

*p < .05; **p < .01.
Mental Health and Cognitive Function in Adults

Table 3. Hierarchical Regression Analyses: Cognitive Variable Regressed on Age, GHQ-12, and the Age × GHQ Cross-Product Interaction Term

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Switching</th>
<th>Updating</th>
<th>Stroop (ms)</th>
<th>Stroop ISD</th>
<th>Recognition</th>
<th>2-CRT (ms)</th>
<th>2-CRT ISD</th>
<th>4-CRT (ms)</th>
<th>4-CRT ISD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NART</td>
<td>$b = .47^*$</td>
<td>.22*</td>
<td>$R^2 = -.02$</td>
<td>.00</td>
<td>.18</td>
<td>.03</td>
<td>-.06</td>
<td>.00</td>
<td>.19*</td>
</tr>
<tr>
<td>Age</td>
<td>$b = -.16$</td>
<td>-.08</td>
<td>$.46^*$</td>
<td>.33*</td>
<td>-.29*</td>
<td>.49*</td>
<td>.40*</td>
<td>.50*</td>
<td>.48*</td>
</tr>
<tr>
<td>GHQ-12</td>
<td>$b = -.06$</td>
<td>.03</td>
<td>.04</td>
<td>.01</td>
<td>.10</td>
<td>.22*</td>
<td>.14</td>
<td>.12*</td>
<td>.07</td>
</tr>
<tr>
<td>Age × GHQ</td>
<td>$b = -.28$</td>
<td>.01</td>
<td>.15</td>
<td>.00</td>
<td>-.02</td>
<td>.00</td>
<td>.57*</td>
<td>.03*</td>
<td>-.12</td>
</tr>
</tbody>
</table>

Notes: NART = National Adult Reading Test; GHQ-12 = 12-item General Health Questionnaire; CRT = Choice Reaction Time; ISD = Intraindividual Standard Deviation. Step 1, $df = 1, 193$; Step 2, $df = 2, 191$; Step 3, $df = 1, 190$.

*p < .01.

term was significant. In both instances, the equivalent interactions for mean RT were nonsignificant (ps > .90). Regression lines for the significant interaction terms are presented in Figure 1. Although the Stroop task exhibited greater variability overall, in both tasks, for persons with low GHQ scores, greater age is not associated with a marked increase in WP variability. Importantly, however, for individuals recording higher GHQ scores, not only is WP variability greater, but the degree of variability increases with older age. In other words, poorer mental health is having a greater impact on the response variability of older persons.

It is of note that the significant Age × GHQ interaction was found for ISDs in the more demanding 4-CRT task and not for the two-choice version of the task. As this may reflect a selective effect related to task demands (i.e., the increase from two to four choices), we repeated the regression model, but this time we controlled for 2-CRT ISD at Step 1. This made no difference to our original finding. As this procedure effectively controls for the WP variability attributable to the two-choice task and leaves residuals reflecting the increase from two to four response choices, the finding supports the view that it is WP variability in the more effortful psychomotor condition that is particularly sensitive to Age × Mental Health relations.

We also wished to confirm that our regression procedure for removing age and time-on-task influences on RT from ISD metrics had been successful and had no bearing on the current findings. Therefore, we repeated the regression models for Stroop and 4-CRT ISDs, but this time we controlled for the respective mean RT measures at Step 1. Although in both cases the first-order effects for age and GHQ were attenuated, the Age × GHQ interactions were unaffected (Stroop $\Delta R^2 = .03$, $p < .01$; 4-CRT $\Delta R^2 = .03$, $p < .01$); this suggests that the effects for WP variability were independent of mean RT.

Finally, our further aim in this study was to assess if the independent measures of executive function accounted for significant Age × GHQ interactions where they were found. This was of particular interest, given (a) the finding of a significant Age × GHQ interaction in respect to Stroop and 4-CRT ISDs, and (b) the theoretical role of executive control in mental health–cognition relations, and also in explanations of WP variability. Therefore, we ran a further series of regression models for Stroop and 4-CRT ISDs, but, respectively, we controlled at Step 1 for switching, updating, and, in the case of 4-CRT, inhibition. If the respective executive function measures accounted for the associations found in the original models, then the effect size of the Age × GHQ interaction would be attenuated in these repeated analyses. When adjusting for switching and updating, we found that only marginal attenuation occurred (for 4-CRT controlling for switching, $\Delta R^2 = .03$, $p < .01$; for updating, $\Delta R^2 = .03$, $p < .01$; for Stroop controlling for switching, $\Delta R^2 = .03$, $p = .012$; for updating, $\Delta R^2 = .03$, $p < .01$). However, when examining 4-CRT ISDs having controlled for Stroop ISDs, we found that $R^2$ was markedly attenuated ($\Delta R^2 = .01$, $p > .06$). This latter finding suggests that variability in inhibition is accounting for some of the Age × GHQ-related variability in 4-CRT.

Discussion

To our knowledge, this is the first study to investigate age, mental health, and WP variability in a continuous age range spanning young adulthood to persons in their early 90s. The work extends research describing an association between intraindividual variability and neurological disorders, and it suggests that having mild mental health problems (i.e., depression and anxiety, social dysfunction, loss of confidence) in old age is associated with variability in cognitive performance. Several important findings were produced. First, although the association between mental health and cognitive function did not vary according to age for the majority of variables, significant Age × GHQ interactions were produced in relation to WP variability in the Stroop and 4-CRT tasks. Poorer mental health and older age were associated with greater variability. Importantly, a dissociation was evident, as measures of mean RT for those variables were not similarly related to age and mental health. Second, the association for 4-CRT remained even when we controlled for the two-choice equivalent of this measure, suggesting the effect was selective to the more effortful task condition. Finally, the significant Age × Mental Health interaction in relation to 4-CRT WP variability was accounted for by executive function measures of inhibition, but not by switching or updating. Overall, the findings were unrelated to gender, and they were independent of verbal IQ.

The present findings build on earlier research at the WP level that found a weak but nonsignificant association between age, day-to-day stress, and variability (Sliwinski et al., 2006). It is
important that the dissociations found here suggest that variability measures may be sensitive to relatively subtle effects of age and mild mental health problems, whereas measures of mean RT are not. This sensitivity to mild effects is underlined by the observation that persons aged 18 to 29, 30 to 39, 40 to 49, and 50 to 65 years recorded mean GHQ scores in line with functioning community-dwelling persons in those age bands elsewhere (Mullarkey, Wall, Warr, Clegg, & Stride, 1999). This suggests that participants were representative of the normal population and were not experiencing major mental health problems.

The reason for the dissociation with mean RT measures may be related to the finding that increased WP variability is largely due to a greater proportion of slower responses falling into the tail of the individual’s RT distribution (e.g., Spieler, Balota, & Faust, 1996; West et al., 2002). It has been suggested that this may represent fluctuations in executive control (Bunce et al., 2004; West et al.) or, relatedly, attentional lapses (Bunce, Warr, & Cochrane, 1993), although this latter view has been questioned (Saltzhouse, 1993). It was noted earlier that attentional resources may be reduced in depression, and that task-irrelevant depressive thoughts may result in cognitive deficits, particularly in more effortful processing (Hartlage et al., 1993). As the cognitive effects of depression are modulated by frontally supported executive processes, and as fluctuations in these processes are thought to underlie WP variability, it is plausible that measures of intraindividual variability are more sensitive to the deleterious influence of poor mental health at a lower threshold than mean RT measures. Had our participants been experiencing greater mental health problems, we would expect measures of mean RT to capture Age × GHQ associations too.

The present dissociations underline the possibility that WP variability is a particularly sensitive behavioral marker of cognitive or neurobiological disturbance. Indeed, regarding the latter point, recent research (Bunce et al., 2007) has shown that, in a relatively young (60–64 years) group of healthy older adults, that white matter lesioning of the frontal cortex is related to WP variability, but not mean RT, for the same psychomotor task. This work, together with the findings of the present study, clearly suggests that measures of WP variability have the potential to detect associations or conditions in circumstances where more commonly used measures of accuracy and mean RT do not. Therefore, the measures may provide a valuable addition to screening and assessment batteries in health care settings.

Two further aims of the study were to assess how far executive function and memory varied as a function of age and mental health, and also to establish if findings for WP variability were accounted for by measures of executive control. Although there was no evidence that memory varied according to age and mental health, our research did show that, consistent with research elsewhere (e.g., Elderkin-Thompson et al., 2006; Sheline et al., 2006), executive function was associated with mental-health-related cognitive deficits in older adults. Specifically, we found a significant Age × Mental Health interaction in respect to inhibition, but not in relation to measures of switching and updating. This finding is consistent with the view that a failure to inhibit intrusive thoughts related to depression and anxiety, social dysfunction, and loss of confidence underlies mental-health-related cognitive deficits, and that reduced attentional resources in old age exacerbates this effect.

This interpretation was underlined by the additional finding that the Age × GHQ interaction in respect to 4-CRT variability was substantially attenuated when we controlled for variability in the Stroop task, our measure of inhibition. One should exercise caution when interpreting this latter finding, however, because of statistical concerns relating to variance-partitioning procedures, and population mean confounds, in cross-sectional samples (e.g., Hofer & Sliwinski, 2001; Lindenberger & Pöpper, 1998; Sliwinski & Hofer, 1999). Furthermore, consideration of Table 2 suggests that the respective measures of executive function were weakly correlated, and Table 3 indicates there was little shared variance between age and both switching and updating. As these findings were unexpected, it raises the possibility that the present measures did not capture those constructs with
sufficient rigor and that alternative measures should be considered for future research. Despite these limitations, though, we believe the findings relating to the measure of inhibition have potentially important theoretical implications that should be investigated further through longitudinal research.

There are several further limitations to the present study that we should acknowledge. First, as noted, cross-sectional data were used, and therefore we are unable to draw any conclusions concerning causality. Second, although the finding of a significant Age × GHQ interaction in relation to WP variability in the 4-CRT but not the 2-CRT task suggests that it is during more effortful processing, which places greater demands on executive control, that mental health–cognition effects are strongest, we did not independently measure the perceived workload associated with the respective tasks. Therefore, we cannot confirm that an increase in perceived demands occurred, and accordingly, that interpretation should be treated with appropriate caution. Finally, although the GHQ possesses good psychometric properties and is widely used as a measure of mental health (Goldberg et al., 1997), the short version we used does not specifically delineate between anxiety and depression. Therefore, although we can assume that participants recording higher GHQ scores were experiencing elevated anxiety or depression, it is not possible to say which. Current work in our laboratories involving an independent sample of adults distinguishes between those mental health constructs and will shed light on this issue.

To conclude, the finding of significant Age × Mental Health interactions in respect to WP variability, but not measures of mean RT, raises the possibility that moment-to-moment fluctuations in information processing may characterize the influence of poor mental health on cognition in older persons. This finding not only is of theoretical interest but also suggests that measures of WP variability are sensitive to relatively subtle, but nonetheless important, mental-health-related effects that may be masked in other metrics and cognitive measures. As the present study builds on research suggesting that measures of WP variability and mean RT dissociate in the presence of possible disease pathology, it is plausible that measures of WP variability may not only provide valuable theoretical insights but also have considerable potential for assessment and diagnostic purposes in health care contexts. It is important for future research to explore this potential further.

Acknowledgments

This research was partly funded by Grant RES-000-22-1399 from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce. Anusha Ramchurn was supported by a studentship from the Economic and Social Research Council, UK, to David Bunce.

References


