The effects of habitual caffeine use on cognitive change: a longitudinal perspective

M.P.J. van Boxtel a,*, J.A.J. Schmitt a, H. Bosma b, J. Jolles a

a Department of Psychiatry and Neuropsychology, European Graduate School of Neuroscience (EURON), Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands
b Department of Care Sciences, Universiteit Maastricht, Maastricht, The Netherlands

Received 25 April 2003; received in revised form 18 June 2003; accepted 18 June 2003

Abstract

The efficiency of higher cortical functions, such as memory and speed of complex information processing, tends to decrease with advancing age in normal healthy individuals. Recently, a high habitual intake of caffeine was found associated with better verbal memory performance and psychomotor speed in several cross-sectional population studies. We tested the hypothesis that habitual caffeine intake can reduce or postpone age-related cognitive decline in healthy adults. For this purpose, the cognitive performance of all participants in the Maastricht Aging Study (MAAS), aged between 24 and 81 years, was reassessed after 6 years. Information on the intake of caffeine-containing beverages was available from the baseline questionnaire. After 6 years, 1376 (75.6%) individuals were available for reassessment. After correction for demographic characteristics, baseline performance and health status, there were small albeit significant associations between the overall estimated caffeine intake at baseline and the 6-year change in complex motor speed (motor choice reaction time). The earlier found association between caffeine intake and verbal memory performance was not apparent in this longitudinal study. These results imply that the longitudinal effect of habitual caffeine intake is limited and will not promote a substantial reduction in age-related cognitive decline at a population level.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Caffeine; Cognitive function; Aging; Longitudinal

1. Introduction

Coffee is a popular beverage in Western societies, and it is estimated that 90% of the Dutch population drinks coffee (Hameleers et al., 2000). The caffeine contained in coffee and other frequently used beverages, such as tea and cola, is considered to have a mild stimulating effect on the central nervous system (CNS), causing increased arousal (for recent reviews, see Fredholm et al., 1999; Nehlig et al., 1992; Smith, 2002). The acute ingestion of caffeine increases extracellular levels of acetylcholine and serotonin by binding to adenosine receptors in the brain. These neurotransmitter systems are involved in many higher cortical circuits implicated in cognitive processes (Buhot et al., 2000; Robbins, 1997). Because of its cognition-enhancing properties, caffeine has been suggested as a potential drug to counteract age-related cognitive decline (Riedel and Jolles, 1996). Indeed, earlier observations have suggested that the daily intake of caffeine-containing beverages is associated with performance on tasks of verbal memory and information processing speed (Hameleers et al., 2000; Jarvis, 1993). Jarvis (1993) demonstrated in a group of 7414 adults (aged 18+ years) that a higher habitual caffeine intake was positively related to better performance on tasks of choice reaction time (incidental verbal learning and visuospatial learning) after controlling for sociodemographic, health and lifestyle variables. Interestingly, older individuals appeared to benefit most from a higher caffeine intake, and performance on some tasks showed a dose–response relationship with caffeine intake. These findings were partially replicated in a more controlled, but methodologically comparable, study of 1875 adults aged between 24 and 81 years (Hameleers et al., 2000). These authors again reported associations between estimated caffeine intake and performance on a choice reaction time task (movement times), the delayed recall of a verbal word learning task and the Stroop...
test (card I: reading speed). However, in contrast to the Jarvis study, no age by caffeine intake interactions were observed. Finally, a recent study in 1538 participants in the Rancho Bernardo Study (mean range 72.6 and 73.3 years, for women and men, respectively) recently produced some evidence that estimated lifetime coffee intake may be beneficial for cognitive performance (Johnson-Kozlow et al., 2002). In this study, the association between estimated coffee intake and better performance on a wide range of neuropsychological tests was most prominent in women aged ≥ 80 years. Thus, these studies gave some support to the notion that habitual caffeine use may boost to some extent the cognitive reserve of the consumer.

Unfortunately, all studies mentioned so far have been cross sectional, which makes a causal interpretation of a relationship between habitual coffee intake and cognitive function still rather speculative. Other unknown variables may have affected the relationship between caffeine intake and cognition. Ideally, a prospective study into the change in (age-related) cognitive performance as a function of caffeine intake is necessary to substantiate the claim that caffeine may postpone or even reduce age-related cognitive decline. We therefore decided to repeat the analysis of Hameleers et al. on the 6-year follow-up data that were derived from the same adult population sample (Maastricht Aging Study or MAAS) (Jolles et al., 1995; Van Boxtel et al., 1998). We specifically investigated whether the decline in cognitive performance over 6 years is associated with a lower habitual intake of caffeine-containing beverages at baseline and whether the effect of caffeine is more pronounced in older than in younger individuals.

2. Materials and methods

2.1. Participants

MAAS started in 1991 as a prospective, observational study of the determinants of normal cognitive aging. Participants in MAAS were recruited from the Registration Network of Family Practices (RNH) (Metsemakers et al., 1992), a sample frame for research in primary care. Individuals were excluded if they had a history of stroke, mental retardation or chronic neurological pathology (e.g., dementia, epilepsy, parkinsonism and CNS malignancy). Participants were stratified for three demographic variables known to be related to cognitive performance: age (12 levels, ranging from 25 ± 1, 30 ± 1, . . . , to 80 ± 1 years), sex and level of general ability (two levels, based on educational level and achievement in professional life) (Van Berkel and Tax, 1990). At baseline measurement, 1821 individuals were tested as part of the MAAS program. Six years later, 1376 (75.6%) individuals returned for follow-up assessment. A group of 271 (14.9%) individuals refused further participation, 118 (6.3%) had died, 37 (2.0%) were medically unfit to participate and 19 (1.0%) did not take part for other reasons. Of this retested group, 10 individuals had incomplete demographic data at baseline, which were interpolated in the regression analyses using a mean substitution procedure. Dropouts who refused to participate were more often women, had lower educational levels and had lower baseline scores on the cognitive tests. Follow-up dropouts who had died were more often men, older and had a poorer performance on cognitive tests than the follow-up participants. Although follow-up participants and dropouts differed in terms of sociodemographic and cognitive characteristics, attrition had only a minor effect on estimates of cognitive change over a 3-year period (Van Beijsterveldt et al., 2002). The MAAS was approved by the Medical Ethics Committee of the University Hospital Maastricht. All participants gave their written informed consent.

2.2. Questionnaire

Information about demographic and lifestyle variables was derived from the baseline questionnaire data set: age at time of first measurement, sex, educational level (range 1–8), smoking status [actual smoking: no (0)/yes (1)], alcohol intake (average number of standard units per week), general health [good to excellent (0)/poor to very poor (1)], housing tenure [owner occupied (0)/rented (1)] and occupation [blue-collar worker (1)/other (0)]. These variables were used in previous cross-sectional studies (Hameleers et al., 2000; Jarvis, 1993) and were again treated as covariates in this analysis.

Caffeine consumption. Participants were asked about their daily caffeine consumption. In answer to the question “Do you drink coffee?” they could tick one of the following answers: “No,” “Yes, 1–3 cups a day,” “Yes, 4–6 cups a day,” “Yes, 7–10 cups a day” and “Yes, more than 10 cups a day.” A similar question was asked about tea. The midpoint of these ranges was used in statistical analyses (i.e., 0, 2, 5, 8.5 and 11 cups). The total daily caffeine consumption was calculated as the number of cups of coffee equivalents consumed per day according to the formula:

\[
\text{Number of daily consumed cups of caffeine (Caf}_{BA}) = \text{round}\left(\frac{85 \times M_{\text{coffee}} + 30 \times M_{\text{tea}}}{85}\right)
\]

in which the weights assigned to mean habitual coffee and tea consumption are the actual average caffeine contents of coffee and tea according to industrial standards (85 and 30 mg, respectively). An extended caffeine intake questionnaire was used at the 6-year follow-up for 1007 individuals. Based on this additional information, which was complete for 928 (92%) individuals, caffeine intake at baseline was calculated as a function of age and beverage type (including cola and “energy drink”). However, due to the availability of these follow-up data in only a subgroup of the participants, they were not used in the regression analyses.
2.3. Cognitive tests

The psychometric test battery used for this study was the same as that used by Hameleers et al. (2000) and covers the cognitive domains of (verbal) memory, reaction time, planning capacity and attention. Before the cognitive tests were administered, participants were offered a cup of coffee or tea according to their own preferences to maintain a preferred internal level of caffeine.

Visual Verbal Learning Test (VVLT) is a computerized, visual version of a test of secondary memory (Brand and Jolles, 1985). In five consecutive trials, a list of 15 monosyllabic nouns, similar in frequency, was presented on a computer screen. The words appeared at a rate of one per second with an interstimulus interval of 2 s. Immediately after the presentation, participants were asked to reproduce these words. Twenty minutes after the final trial, participants were asked again to reproduce the words (delayed recall). The cumulative number of words recalled over the five immediate recall trials was taken as a measure of learning capacity. Delayed recall is regarded as an indicator of memory consolidation.

Motor Choice Reaction Test (MCRT) is a computer test in which measurements are obtained from a push-button panel (Houx and Jolles, 1993). It contains one central red button and five surrounding white target buttons, laid out in an 180° arc, all at 6-cm distance from the red button. The participant was requested to hold down the red button with the index finger of the preferred hand as long as no white button was illuminated. As soon as a white button was illuminated, the participant had to release the red button and shortly press the illuminated button and return to the red central button. Two subsequent task conditions of increasing difficulty were used: simple reaction time (SRT), in which only the upper button was illuminated, and choice reaction time, in which one of a defined set of three buttons was illuminated. Both conditions consisted of 30 button presses. The following dependent variables were recorded: median reaction time (time from stimulus onset until the illuminated button was pressed) and median movement time (time from release of the hold button until the response button was pressed).

Letter-Digit Substitution Test (LDST) is a modified version of the procedurally identical Digit-Symbol Substitution Test (Smith, 1968). The number of correctly completed letters in 90 s was used as the dependent variable.

Fluency is a test for strategy-driven retrieval of information from semantic memory (Luteijn and van der Ploeg, 1983). Participants were asked to produce as many animal names as possible in 1 min. The number of correct responses was taken as the dependent variable.

The Concept Shifting Test (CST) is the Maastricht adaptation of the Trail Making Test (Reitan, 1958). On each test sheet, 16 small circles (Ø 15 mm) were grouped in a larger circle (Ø 16 cm). In the smaller circles, the test items appeared in a fixed random order. The test items were numbers (CST-A), letters (CST-B) or both (CST-C). Participants were asked to cross out the items in the correct ascending order as quickly as possible. The dependent variable (CST-int) was the relative contribution of the need to shift between two concepts, which was computed according to the formula:

\[
\text{CST-int} = \left[ t_{\text{CST-C}} - 0.5(t_{\text{CST-A}} + t_{\text{CST-B}}) \right] / \left[ 0.5(t_{\text{CST-A}} + t_{\text{CST-B}}) \right] \times 100%
\]

in which \( t_{\text{CST-A}} \), \( t_{\text{CST-B}} \) and \( t_{\text{CST-C}} \) refer to the times needed to complete the operations (Jolles et al., 1995).

The Stroop Color-Word Test is a test with three different subtasks, each displaying 10 × 10 items (Houx et al., 1993). In subtask I, color names (red, yellow, green and blue) printed in black ink were to be read as fast as possible. In subtask II, the color of different patches was to be named as quickly as possible. In subtask III, color names were printed in incongruously colored ink. Participants were asked to name the color of the ink in which the words were printed. Time to complete subtask I was taken as a measure of simple response (reading) speed. Interference (Stroop-int), as a measure of attention, was computed with the formula:

\[
\text{Stroop-int} = \left[ t_{\text{STR-II}} - 0.5(t_{\text{STR-I}} + t_{\text{STR-II}}) \right] / \left[ 0.5(t_{\text{STR-I}} + t_{\text{STR-II}}) \right] \times 100%
\]

where \( t_{\text{STR-I}} \), \( t_{\text{STR-II}} \) and \( t_{\text{STR-III}} \) refer to the time to complete the subtasks I, II and III, respectively (Houx et al., 1993).

2.4. Statistical analysis

Demographic and lifestyle characteristics of different caffeine intake groups were compared using a one-way analysis for between-group differences and linear trend in continuous variables and with Pearson’s \( \chi^2 \) tests for group differences and linear trend when variables were dichotomous. Next, multiple hierarchical regression analysis was used to test if the performance at follow-up was related to habitual caffeine intake reported 6 years earlier, adjusting for baseline performance in step 1 and sociodemographic variables (age, sex, educational level, occupation and housing tenure), lifestyle (smoking and alcohol consumption) and subjective health in step 2. Habitual caffeine consumption at baseline was entered in the model in the final step. Residuals were inspected for systematic trends, and collinearity diagnostics were generated to check for unstable regression models. All analyses were performed with the SPSS v10.0 program series for Apple Macintosh. \( P \) values of \( \leq .05 \) were considered statistically significant.

3. Results

In Table 1, descriptive information is presented as a function of habitual caffeine intake at baseline. Caffeine
intake was associated with older age [one-way test for trend: \( F(1,1361) = 15.60, P < .001 \)]. Furthermore, there were strong trends toward higher caffeine intake in both men and smokers \( \chi^2 \text{ tests for trend: } \chi^2(1,1361) = 6.66, P = .010 \) and \( \chi^2(1,1361) = 52.60, P < .001 \) for sex and smoking status, respectively.

Fig. 1 displays the mean cumulative intake of caffeine, specified for each beverage type, by age category. Maximum intake was observed in the 34–41-year group, with gradually lower levels in older age categories. In all age groups, coffee accounted for 76% or more of the total daily caffeine intake.

The regression models fitted for each neurocognitive variable separately are presented in Table 2. It should be noted that for VVLT, LDST and Fluency, negative \( \beta \) values for a given predictor should be interpreted as indicative of greater deterioration and vice versa for all other (timed) cognitive parameters. The cognitive performance at baseline explained between 4% (CST-int) and 80% (LDST) of the variance in the cognitive outcome at follow-up. Age contributed significantly to the prediction of change in all parameters, and associations were in the expected direction: the performance of older individuals had decreased more after 6 years than had the performance of younger people. In addition, women showed a smaller change on MCRT movement times and CST-int, but the reverse was true for CST-A. The WLT-total score changed less in women. Multiple effects of educational level on change scores were found, generally indicating that better educated individuals showed less age-related cognitive decline. Effects of lifestyle variables were absent (Smoking) or spurious (Alcohol, Health and Occupation). Interestingly, Housing tenure, which can be regarded as an additional indicator of socio-economic status, was predictive of a change in memory performance: individuals who owned a house showed a smaller decrease in performance on the VVLT than those who rented their house. When total caffeine intake at baseline was entered in the regression model in step 3, additional variance in the change in performance was explained on both MCRT movement times: a higher caffeine intake at baseline was associated with a smaller change in performance after 6 years. However, the associated \( \beta \) weights were small in size ( \( \sim 20\% \) of the overall age effect) and the resulting \( R^2 \) change, albeit significant, was less than 1%. The prediction of other cognitive parameters did not improve on step 3. When the analyses were repeated and total caffeine intake was replaced by caffeine intake through coffee alone in step 3, the pattern of effects did not change substantially (results not shown). In this analysis, the effects of caffeine intake on both MCRT movement times were

![Fig. 1. Total caffeine intake at follow-up as a function of age and beverage type (n=928).](image-url)
Table 2
Results of linear regression analyses of cognitive performance at follow-up on baseline performance, demographical and lifestyle characteristics and overall caffeine intake at baseline (CafBA)

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVLTb</td>
<td>Baseline performance</td>
<td>Age</td>
</tr>
<tr>
<td>VVLTb</td>
<td>Σ Trials 1–5</td>
<td>R²</td>
</tr>
<tr>
<td>Delayed</td>
<td>0.55***</td>
<td>.47***</td>
</tr>
<tr>
<td>MCRTc</td>
<td>SR-ini</td>
<td>0.55***</td>
</tr>
<tr>
<td>SR-mov</td>
<td>0.47***</td>
<td>.39***</td>
</tr>
<tr>
<td>CR-ini</td>
<td>0.60***</td>
<td>.44***</td>
</tr>
<tr>
<td>CR-mov</td>
<td>0.52***</td>
<td>.46***</td>
</tr>
<tr>
<td>CST</td>
<td>CST-A</td>
<td>0.53***</td>
</tr>
<tr>
<td>CST-C</td>
<td>0.44***</td>
<td>.39***</td>
</tr>
<tr>
<td>Interference</td>
<td>0.16***</td>
<td>.04***</td>
</tr>
<tr>
<td>LDST</td>
<td>Number correct</td>
<td>0.78***</td>
</tr>
<tr>
<td>Fluency</td>
<td>Number correct</td>
<td>0.56***</td>
</tr>
<tr>
<td>Stroop</td>
<td>Card I</td>
<td>0.56***</td>
</tr>
<tr>
<td>Card III</td>
<td>0.63***</td>
<td>.58***</td>
</tr>
<tr>
<td>Interference</td>
<td>0.50***</td>
<td>.55***</td>
</tr>
</tbody>
</table>

Displayed are the standardized regression coefficients (β’s) in the final model and the proportion of explained variance (adjusted R²) after each successive step.

* P ≤ 0.05. ** P ≤ 0.01. *** P ≤ .001.

Sex: male = 0, female = 1; Smoking (actual): no = 0, yes = 1; Health: good = 0, poor = 1; Housing tenure: owner occupied = 0, rented = 1; Occupation: white collar/other = 0, blue collar = 1.

Delayed = delayed recall.

SR = simple task condition, CR = complex task condition, ini = initiation time, mov = movement time.
attenuated but remained significant. Finally, no additional variance was explained when an Age × Habitual caffeine intake interaction term was added to the regression models.

4. Discussion

This study was undertaken to test earlier found associations between habitual caffeine intake and cognitive function in the MAAS in a longitudinal perspective. We used the same battery of tests and covariate measures that were used in cross-sectional studies (Hameleers et al., 2000; Jarvis, 1993) to investigate whether habitual caffeine intake at baseline could predict performance change over a 6-year period. The measure of caffeine intake was identical to that used in the study by Hameleers et al. in which cross-sectional associations were found between habitual caffeine intake and a verbal memory measure (delayed recall) and two reaction time measures (MCRT movement times). Results indicated that performance (movement times) of the motor choice reaction time tasks (MCRT) was preserved whereas verbal memory performance was not. The effects on MCRT were, however, small and would not survive a Bonferroni correction to adjust the significance levels for multiple testing. Earlier reports suggested that older individuals may be more sensitive to the CNS effects of caffeine (Rees et al., 1999; Swift and Tiplady, 1988). To test this longitudinally, we looked for Age × Caffeine intake interaction effects on performance at follow-up but found none. Thus, based on our results, habitual caffeine intake does not modulate or attenuate the age-related decline in cognitive performance.

There are several methodological issues that may affect the results of this study. The predictive value of caffeine intake for cognitive performance may have been affected by selective attrition in this sample. Attrition in MAAS has been related to sociodemographic variables, such as age, sex and educational level, and to lower levels of cognitive functioning at baseline, although estimates of cognitive change remain relatively unaffected over a 3-year interval (e.g., Van Beijsterveldt et al., 2002). We cannot exclude, however, that attrition promoted a healthy survivor effect on the remaining sample, in which longstanding associations between a health-related habit and cognitive functioning may have been underestimated due to a “regression to the mean” phenomenon.

It could be argued that the reliability of questionnaire information is limited and that there may have been a report bias with respect to quantity of the caffeine-containing drinks. However, the questions about habitual caffeine consumption were based on actual intake and not on past intake. Current intake was taken as being representative for previous intake as coffee/tea drinking habits are relatively stable over time. A related point is that we did not distinguish between caffeine-containing and decaffeinated coffee or between black tea and other types of tea with a potentially lower caffeine content, which may limit the generalizability of our findings. Furthermore, not all sources of caffeine in the daily diet were included in the habitual caffeine intake measure. For instance, caffeine in cola or different varieties of energy drinks was not included; however, studies have shown that these sources of caffeine contribute much less to the total daily caffeine intake (Massey, 1998). This can be seen in Fig. 1, which shows the mean daily caffeine intake at follow-up as a function of age and caffeine source in a subgroup of this population sample. Caffeine intake via cola added only marginally to the overall caffeine intake, particularly in younger age groups. The contribution of caffeine from modern beverages such as “energy drinks” appears to be almost negligible.

In theory, compounds other than caffeine may be involved in the combined behavioral effects of coffee and black tea (Hindmarch et al., 1998). Using our method for calculating the caffeine content in the consumed beverages, we also may have indexed the presence of other substances in these drinks. Although the psychoactive properties of caffeine are by now well documented, it cannot be ruled out completely that the observed associations were related to the action of pharmacologically active compounds other than caffeine.

The present observational, prospective study did not produce substantial evidence that habitual and prolonged caffeine intake reduces cognitive decline in aging individuals. From a public health standpoint, it seems therefore unfounded to promote the use of caffeine-containing beverages to prevent cognitive deterioration in later life.

Acknowledgements

This study was supported by a grant from the Institute for the Scientific Information on Coffee/Physiological Effects of Coffee Committee (ISIC/PEC), which is gratefully acknowledged.

References


