Scandinavian Journal of Psychology, 2014, 55, 225–232 DOI: 10.1111/sjop.12114

Age and sex related changes in episodic memory function in middle aged and older adults

ASTRI J. LUNDERVOLD,[1,2,3] DANIEL WOLLSCHLÄGER[1,4] and EIKE WEHLING[1,2,5]

[1]*Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway*
[2]*Center for research on Aging and Dementia, Haraldsplass Deaconal Hospital, Bergen, Norway*
[3]*K. G. Jebsen Centre for Research on Neuropsychiatric Disorders, University of Bergen, Bergen, Norway*
[4]*Institute for Medical Statistics, Epidemiology and Informatics, University Medical Center, Johannes-Gutenberg-University Mainz, Mainz, Germany*
[5]*Department of Physical Medicine and Rehabilitationl Medicine, Haukeland University Hospital, Bergen, Norway*

Lundervold, A. J., Wollschläger, D. & Wehling, E. (2014). Age and sex related changes in episodic memory function in middle aged and older adults. *Scandinavian Journal of Psychology* 55, 225–232.

Age-related change in episodic memory function is commonly reported in older adults. When detected on neuropsychological tests, it may still be difficult to distinguish normal from pathological changes. The present study investigates age-and sex-related changes in a group of healthy middle-aged and older adults, participating in a three-wave study on cognitive aging. The California Verbal Learning test (CVLT-II) was used to assess their episodic memory function. A cross-sectional analysis of results from the first wave showed higher performance in females than males, with a steeper age-related decline in males. This was confirmed in a longitudinal analysis using a mixed effects regression model, but with a lower age-related change and smaller difference between the sexes. Information about learning strategies and errors in the third wave turned out to contribute significantly to explain change in episodic memory function across the three waves. We argue that the results from the longitudinal analyses are generalizable to the population of healthy middle-aged and older individuals, and that they could be useful in guiding clinicians when evaluating individuals with respect to cognitive change.

Key words: CVLT, cognitive aging, longitudinal, mixed effect regression model.

Astri Lundervold, Department of Psychology, University of Oslo, PO Box 1094, Blindern, Oslo, Norway. Tel:+47 22 84 51 31; fax: +47 22 84 50 96; e-mail: Astri.Lundervold@psybp.uib.no

INTRODUCTION

Older adults frequently report changes in cognitive function, very often with a concern about degenerative disorder. However, the precise distinction between normal and pathological changes in cognitive function may be obscure. Knowledge about characteristics of normal cognitive aging and its influence on test results are therefore essential to a clinical neuropsychologist.

This is clearly exemplified when assessing episodic memory function, where changes are known to represent one of the earliest signs of a neurodegenerative disorder (e.g., Bäckman, Small & Fratiglioni, 2001; Petersen, 2004; Twamley, Ropacki & Bondi, 2006). At the same time, episodic memory function is considered to show age-related decline in healthy individuals, with cross-sectional studies reporting a gradual decline with an onset that may start as early as in the twenties (Salthouse, 2009; Zelinski & Burnight, 1997). These findings are, however, nuanced by results from longitudinal studies. Such studies have both reported a markedly later age of onset of cognitive change and large individual differences in developmental pathways (Nyberg, Lövden, Riklund, Lindenberger & Bäckman, 2012; Rønnlund, Nyberg, Bäckman & Nilsson, 2005; Schaie, 2005), leaving some eighty year olds with a function that is not different from what is shown by much younger individuals.

Tests with multiple learning trials are widely used to assess episodic memory function in elderly, and are among the most sensitive measures when changes in memory function indicate subsequent cognitive decline and dementia (Blacker, Lee, Muzikansky *et al*., 2007; Kaltreider, Cicerello, Lacritz, Honig, Rosenberg & Cullum, 2000; Rabin, Par, Saykin *et al*., 2009; Tierney, Yao, Kiss & McDowell, 2005). This is due to characteristics of these tests, involving complex processes of memory function, including working memory, the ability to learn, consolidate and recollect information. Several list-learning tests have therefore been developed, providing information about different aspects of the memory process. Some tests include measures of higher-order cognitive functions, such as the ability to organize information efficiently and correctly, functions that are shown to be important mediators of encoding and recall (Davis, Klebe, Guinther, Schroder, Cornwell & James, 2013). A main strength of the California Verbal Learning Test (CVLT) (Delis, Kramer, Kaplan & Ober, 2000) is its inclusion of measures of learning strategy and errors in addition to more traditional measures of learning, recall and recognition, characteristics that together have contributed to the test's sensitivity to mild cognitive change (Benedict & Zgaljardic, 1998; Greenaway, Lacritz, Binegar, Weiner, Lipton & Munro Cullum, 2006; Lacritz, Cullum, Weiner & Rosenberg, 2001; Rabin *et al*., 2009) and dementia (Delis *et al*., 2000).

It has been shown that females tend to outperform males on verbal tests of episodic memory function (Herlitz, Nilsson & Bäckman, 1997). A recent study by Sunderaraman, Blumen, DeMatteo, Apa and Cosentino (2013) indicated that superior recall in females on CVLT is related to their higher semantic clustering score. This sex difference has also been related to a less lateralized brain function in females than males, but results from studies relating verbal function to brain areas have been conflicting (Wallentin, 2009). Gender effects on cognitive function are obviously modulated by several neurobiological factors (see Reinvang, Winjevoll, Rootwelt & Espeseth, 2010b), emphasizing that sex should be taken into account when evaluating age-related changes in cognitive function.

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

The present study investigates changes in episodic memory function associated with normal aging. One obvious aim of studies on age-related changes is to obtain predictions that can be generalized to the population under study. As stated previously, predictions based on cross-sectional and longitudinal studies have come to different conclusions, with the former underestimating cognitive abilities in older age (Nyberg *et al*., 2012). However, longitudinal studies are also challenged by several factors. In addition to measurement errors and learning effects (Levine, Svoboda, Hay, Winocur & Moscovitch, 2002; Salthouse, 2012), longitudinal analyses commonly must take into account varying time-spans between assessments of individual participants, incomplete data sets and different correlation structures over time. This makes it crucial to use a statistical procedure that can handle these factors appropriately. Traditional linear regression models fall short in this regard (Long, 2012), while they are much more appropriately handled by mixed effects regression models (Galecki & Burzykowski, 2013). Another important strength of using mixed effects models is their ability to explicitly incorporate the effect of individual differences, and by this get more reliable estimates of the variability in a population characterized by cognitive heterogeneity.

The present study, including a sample of healthy middle-aged and older adults, investigates age-related changes in episodic memory function both cross-sectionally by using a standard regression model and longitudinally by using a mixed effects regression model. Episodic memory function was assessed by the standard version of the second edition of CVLT (CVLT-II) in three study waves with about three years apart. In the cross-sectional analysis, including measures of learning, recall, recognition, the use of an active learning strategy and accuracy in responses, we expected females to perform better than males, with a steeper age-related decline in males. To minimize measurement errors in the longitudinal analysis, we used a composite memory score calculated from the measures of learning, recall and recognition as a dependent variable. To control for the effect of repeated testing, the effect of a second assessment was included as an independent variable. From earlier studies we expected that the longitudinal analysis would demonstrate a lower age-related change than revealed in the cross-sectional analysis, and that this lower change would be found in both sexes. When a neuropsychologist uses CVLT-II to evaluate if an individual has shown a decline in episodic memory function, performance reflecting learning strategy and errors will commonly contribute to the conclusion. We therefore extended the longitudinal analysis by investigating how much information about these variables from the third wave adds to the longitudinal estimate of change in episodic memory function explained by age and sex.

METHODS

Sample

Healthy individuals were invited through advertisement to take part in the first wave of a longitudinal study on cognitive aging (N = 163). All participants were examined according to an extensive neuropsychological test protocol, including the Norwegian translation of CVLT-II (Delis, Kramer, Kaplan & Ober, 2004) and a test of intellectual function. Subjects with a history of substance abuse, present neurologic or psychiatric disorder, or other significant medical conditions, were excluded from the study. Based on neuropsychological test results, none of the participants showed a mild cognitive impairment (MCI) as defined by Petersen (2004). Five participants were excluded from the present study, three due to pathology detected at the following waves, one due to an IQ < 80 and one due to missing data on the CVLT-II, leaving 158 participants included in the cross-sectional analyses of the data from the first wave. All individuals from the first wave were invited to a follow-up study about three years after the first examination and to a third wave about three years after the second one. In the longitudinal analyses we included all individuals participating in at least two waves (n = 126), of whom 103 participated in all three waves. The analyses of the contribution of performance characteristics in explaining change across the three waves were restricted to these 103 participants. Information about age and intellectual function is presented separately for males and females in Table 1.

Instruments

Intellectual function. First wave performances on two subtests from the Norwegian translation of the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 2007), the Matrix Reasoning and Vocabulary subtests, were used to estimate intellectual function (IQ). Performance was scored and the full-scale IQ (FSIQ) estimated according to norms presented in the test-manual (Wechsler, 1999). The Matrix Reasoning subtest, included in all waves, showed stable age corrected T-scores over time, ($r = 0.65$, $p < 0.001$), and a significant correlation with the FSIQ measure in the second ($r = 0.85$, $p < 0.001$) and third wave ($r = 0.66$, $p < 0.001$). This stability indicates that the IQ measure from the first wave could be used as an estimate across all study waves.

California Verbal Learning Test. In CVLT-II (Delis *et al*., 2000) 16 words (List A), drawn from four semantic categories, are presented five times. After each presentation the participant is asked to repeat as many words as possible. Immediately after the fifth trial, a new list (List B) is read and the participant is asked to repeat as many words as possible. A short delayed test (CVLT-SD) is presented immediately after recall of List B, where the participant is asked to recall the words of List A, first without cues (free) and then with cues. This last procedure is repeated after a delay of 20 minutes (CVLT-LD). Finally, a yes-no recognition test is presented, including the 16 items from List A, the 16 from List B and 16 random distractor items.

In the present study we included three main measures of episodic memory function: the total number of correct words reported across the five learning trials (*Learning*), a sum of words recalled on the free versions of CVLT-SD and CVLT-LD (*Recall*), and the number of hits

Table 1. *Means and standard deviations of demographic and CVLT-II raw-scores for females and males (Wave 1)*

	Females n = 106	Males n = 52
Age	61.1 (7.6)	61.7 (8.3)
Education	13.8 (3.2)	13.4 (3.6)
IQ	114.0 (12.5)	115.5 (12.8)
Matrix Reasoning	58.1 (8.6)	59.9 (8.6)
Vocabulary	57.3 (8.2)	57.3 (8.8)
Learning	54.6 (9.5)	45.2 (10.5)
Recall	24.5 (5.3)	19.5 (6.1)
Recognition	15.0 (1.3)	14.7 (1.4)
cvltSumZ	–0.71 (2.4)	–1.7 (2.7)
Semantic clustering	1.76 (2.1)	0.94 (1.8)
Consistency	85.6 (8.3)	78.6 (9.9)
Primacy effect	29.4 (5.5)	31.1 (7.0)
Intrusions	5.1 (5.7)	5.7 (6.3)
False Positives	1.6 (2.0)	3.6 (3.6)

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.

on the recognition trial (*Recognition*). Three scores were included to represent the learning strategy: *Semantic clustering*, defined as the number of correct list A items consecutively recalled from the same category relative to the expected number based on chance; *Consistency*, defined as the percentage of items from Trial 1 to 4 that were recalled on the next trial; and the *Primacy effect*, defined as the percentage of items recalled from the primacy region for Trials 1-5. Higher scores on each of these variables indicate an active learning strategy. Two scores were used to represent accuracy: number of *Intrusions* across all recall trials and the number of *False positive errors* on the Recognition trial.

Statistics

All analyses were performed using R (version 3.0.1; R Development Core Team, 2013). The impact of age, sex and the interaction between the two (age: sex) were analyzed in a series of separate regression analyses with scores on the selected CVLT-II measures as dependent variables. Significant models were followed by hierarchical analyses to investigate the contribution from the independent variables. The longitudinal analyses were performed using a linear mixed effects regression model (Galecki & Burzykowski, 2013), using the lme4 package (Bates, Maechler & Bolker, 2013) to model the relationship between CVLT-II performance and age-related change for females and males.

The linear mixed effects regression model included a composite score representing episodic memory function as the dependent variable, with retest effect (*cvltRe*), age, sex and the interaction between age and sex as the sequence of independent variables. In addition to these fixed-effects terms, the model included a random intercept term. The composite memory score *cvltSumZ* was defined as the sum of z-transformation of all values of the raw scores from the total learning trials and the raw scores from the short and long free recall trials over all waves from all subjects. The retest effect was accounted for in the longitudinal model by including the time-varying factor *cvltRe* with two levels as a predictor. Following the recommendation from Ferrer, Salthouse, Stewart and Schwartz (2004), *cvltRe* was set to ‘no’ for measurements from the first wave, and to ‘yes’ for measurements from the second and third wave. Thus, parameter estimates for other covariates were adjusted for an additive practice effect, which we assumed to occur mainly at the second testing, with the third testing adding no additional learning benefit. This strategy was successfully employed in a recent longitudinal study on cognitive aging (Finkel, Reynolds, Larsson, Gatz & Pedersen, 2011).

The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used to evaluate the most plausible model in the set of models being tested (Long, 2012). In all subsequent analyses the significance of fixed effect parameters were tested using step-down model comparisons that started with the full model, then sequentially dropped predictor terms. We report two-tailed *p*-values resulting from partial *F*-tests with degrees of freedom estimated according to the Kenward-Roger procedure as implemented in the R package pbkrtest (Halekoh & Højsgaard, 2013). A similar procedure was used to analyze the contribution of learning strategy and error measures in explaining the change in episodic memory function (i.e., *cvltSumZ*). The measures were included in separate analyses in a model controlling for cvltRe, age and sex.

RESULTS

Cross-sectional analyses -wave 1

Demographics, IQ and CVLT-II results. Table 1 presents the means and standard deviations for age, years of education, intellectual function and CVLT-II raw-scores in males (N = 52) and females (N = 106). All sex-differences on demographic variables were non-significant in a series of *t*-tests comparing males and females. The age-corrected IQ score and its subscores showed means in the upper level of normal function, with no significant sex difference. At wave 1, the females obtained statistically significant higher scores on *Learning* ($t = 5.6$, $p < 0.001$), *Recall* ($t = 5.4$, $p < 0.001$), but not *Recognition* ($t = 1.5$, ns). On the measures of learning strategy, the females obtained a significantly higher (better) score on *Semantic clustering* ($t = 2.4$, $p = 0.019$) and *Consistency* ($t = 4.7$, $p < 0.001$), but not on the *Primacy effect* ($t = 1.6$, ns). The sex-difference was also statistically significant for the *False positive errors*, $t = 4.3$, $p > 0.001$, but not for *Intrusions*, $t = 0.61$, ns.

Primary memory measures

Regression models including age, sex and the interaction between age and sex as dependent variables were statistically significant for *Learning*, *Recall*, but not for *Recognition* (Table 2). For *Learning* the model showed a significant effect of age ($t = 4.15$, $p < 0.001$), sex ($t = 2.21$, $p < 0.05$), and the age:sex interaction ($t = 2.98$, $p < 0.01$). For *Recall* there was a significant effect of age ($t = 3.3$, $p < 0.001$). Females showed a lower intercept and less age-related decline than males: with *Learning* as a dependent variable, the average loss was 0.07 units per year for females and 0.7 for males, with *Recall* the average loss per year was 0.1 units for females and 0.3 for males (Table 2). The hierarchical model showed a significant effect of sex when controlling for age both for *Learning*, $F(1,155) = 33.2$, $p < 0.001$, and *Recall*, $F(1,155) = 33.3$, $p < 0.001$.

The results for the composite score of episodic memory function (*cvltSumZ*) confirmed the lower intercept for females than males ($p < 0.05$) at reference age (65), and a steeper age-related change for males (–0.15) than for females (–0.04). The regression line is illustrated in Fig. 1.

Among the learning strategy measures, the regression model was statistically significant for *Consistency* and *Semantic clustering* (Table 2), with females showing a lower intercept and a lower age-related decline than males. The hierarchical models showed statistically significant effects of sex for both *Semantic clustering*, $F(1,155) = 5.3$, $p = 0.02$ and *Consistency*, $F(1,155) = 21.3$, $p > 0.001$. For the error measures, the regression model

Table 2. *Effects of age, sex and the age:sex interaction on CVLT-II scores in wave 1*

	F-test	*p*-value	Adj.R2	I:F[a]	Age[a]	Age:F[a]
Learning	17.4	<0.001	0.25	–27.2	–0.66	0.59
Recall	14.8	<0.001	0.21	–7.0	–0.30	0.20
Recognition	1.7	0.18	0.01	–1.1	–0.03	0.02
CvltSumZ	16.7	<0.001	0.23	–4.5	–0.15	0.11
Semantic clustering	0.1	0.02	0.05	–2.1	–0.07	0.05
Consistency	8.1	<0.001	0.12	–6.3	–0.24	0.22
Primacy effect	1.0	0.41	0.00	–4.9	–0.01	–0.05
Intrusions	1.2	0.32	0.00	6.9	0.17	–0.12
False positives	8.6	<0.001	0.13	4.2	0.11	–0.10

Note: [a]Estimated unstandardized regression coefficients: I:F = intercept differences females versus males; Age = age-related change with males as a reference; Age:F = the age-related effect for females added to the age-related change with males as a reference.

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.

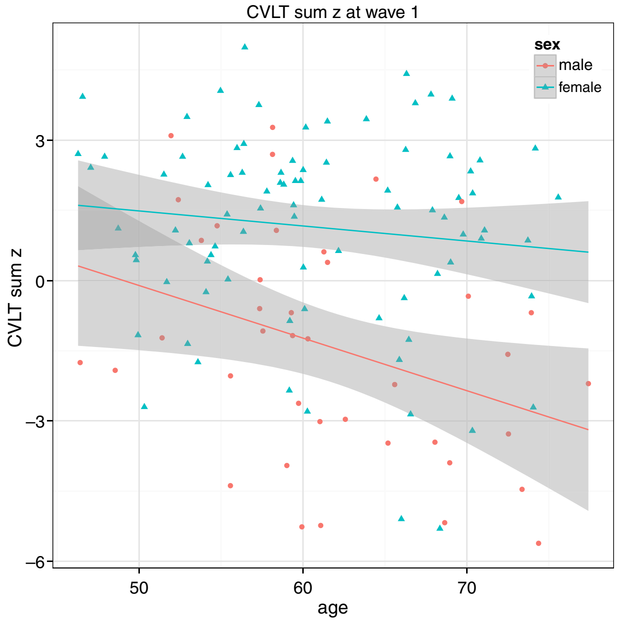

Fig. 1. The composite score *cvltSumZ* in males and females in the cross-sectional analysis of wave 1 data with fixed effects uncertainty.

was only statistically significant for *False positive errors* (Table 2). The intercept was higher in females than males, with a very small age-related change in females (0.01) compared to the change in males (0.11).

Longitudinal analyses

Demographics and CVLT-II variables. There was a mean follow-up time of 6.5 years (range = 5.2 –7.7 years) for participants included in the longitudinal analyzes, with a mean of 3.5 years (range = 2.8 – 4.7 years) between the first two waves, and 3 years (range 2 – 4 years) between Wave 2 and 3. The 23 participants who were lost to follow up for wave 3 were not significantly different on any of the demographic and CVLT-II variables selected for the study. The changes in mean scores on the composite measure of episodic memory function (*cvltSumZ*) and all CVLT-II scores are shown in Table 3. The largest change occurred between Wave 2 and Wave 3, where all except the *Recognition* score were significantly lower in the latter. From the first to the second wave, there was a significant increase on the *Learning*, the composite score *cvltSumZ* and the *Semantic Clustering* scores.

The retest factor cvltRe showed a Kenward-Roger partial F-value of 1.78 ($p = 0.18$) for *cvltSumZ*, with a regression weight of 0.37. This means that on average, the retest was 0.37 *cvltSumZ* units higher than the score in Wave 1.

The effect of age and sex on change in episodic memory function

A linear mixed effects model (Table 4) including cvltRe, age, sex and age:sex interaction as independent variables and the composite score *cvltSumZ* as dependent variable, showed a considerable variation in estimated individual intercepts. The variance was 3.86 units change in *cvltSumZ* score, highlighting the need to take individual differences into account in the model. The age-related change in males was –0.14 units on the *cvltSumZ* score, with females showing a somewhat lower change (–0.12). The AIC and BIC showed the best model fit when cvltRe, age and sex were included, with no improvement of fit by the age:sex interaction. The Kenward-Roger partial F-test showed a significant contribution from both age and sex, but the difference between age-related change in males and females was far lower than in the cross-sectional analyses (Table 4).

The linear mixed effects model computed separately for *Learning* and *Recall* confirmed the high variation in individual intercept, with a variance of 54.4 and 18.4 units change, respectively. Their model fit was not increased by including the age:sex interaction in the model, while the contribution from age and sex was confirmed by the Kenward-Roger partial F-test on both variables. Although the age-related decline was steeper for males than females (–0.52 and –0.30), it was not markedly different from the decline in females (–0.48 and –0.26, see Table 4). The regression line is illustrated in Fig. 2.

Change in episodic memory function explained by learning strategy and errors

In the next series of analyses we investigated the add-on effect of information about learning strategy and errors in the third wave in explaining the change in *cvltSumZ* score. Each measure of learning strategy or errors from the third wave was included as predictor in separate analyses after inclusion of cvltRe, age

Table 3. *Means and standard deviations for CVLT-II raw-scores across the three waves, mean changes across waves and associated p-values from paired t-tests*
Wave 1 Wave 2 Wave 3

CVLT-II variables	$n = 126$	$n = 126$	$n = 103$	t 1–2^a	p 1–2^a	t 2–3^b	p 2–3^b
Learning	52.9 (9.9)	55.6 (11.4)	52.9 (10.9)	–3.5	< 0.001	5.5	< 0.001
Recall	23.8 (5.6)	24.4 (6.0)	21.7 (6.6)	–1.4	0.1	7.8	< 0.001
Recognition	15.0 (1.2)	15.2 (1.1)	15.8 (1.2)	–1.4	0.2	0.9	0.40
cvltSumZ	0.3 (2.5)	0.8 (2.8)	–0.6 (2.9)	–2.5	0.01	9.2	< 0.001
Semantic clustering	1.5 (2.0)	2.4 (2.7)	1.5 (5.7)	–4.2	< 0.001	4.8	< 0.001
Consistency	85.0 (8.3)	85.0 (10.0)	83.0 (5.9)	1.5	0.14	–2.1	0.04
Primacy	29.6 (5.9)	28.7 (2.2)	29.4 (6.3)	–0.2	0.80	2.9	0.01
Intrusions	4.6 (5.2)	4.1 (4.4)	6.3 (5.7)	0.9	0.40	–4.2	< 0.001
False Positives	2.0 (2.7)	1.7 (2.8)	3.3 (4.1)	1.5	0.10	-6.4	< 0.001

Note: ^{a}*t* 1-2/*p* 1-2: *t*-test/*p*-value between score in Wave 1 and 2. ^{b} t 2-3/p 2-3: *t*-test/*p*-value between score in Wave 2 and 3.

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.

Table 4. *Results from mixed effects regression analysis with Kenward-Roger partial F-tests and estimated regression coefficients for the composite episodic memory score (cvltSumZ), and the Learning and Recall scores*

Dependent variables the regression model	AIC[a]	BIC[a]	F-test	df1	df2	P-value	I: F[b]	Age[b]	Age:F[b]
cvltSumZ							1.28	–0.14	0.02
cvltRe	1582	1598	1.8	1	229.2	0.18			
Age	1554	1574	31.4	1	173.2	< 0.001			
Sex	1528	1551	31.8	1	123.7	< 0.001			
Age:Sex	1530	1537	0.1	1	275.8	0.71			
Learning							4.57	–0.52	0.06
cvltRe	2149	2164	0.1	1	229.5	0.76			
Age	2603	2619	27.0	1	162.2	< 0.001			
Sex	2556	2579	27.7	1	123.8	< 0.001			
Age:sex	2558	2585	0.1	1	251.3	0.73			
Recall							2.26	–0.30	0.04
CvltRe	2149	2164	3.1	1	229.2	0.08			
Age	2124	2143	27.2	1	171.1	< 0.001			
Sex	2098	2121	31.6	1	123.7	< 0.001			
Age:sex	2099	2127	0.2	1	272.0	0.64			

Note: [a]AIC = Akaike information criterion; BIC = Bayesian information criterion.
[b]Estimated regression coefficients: I:F = intercept difference females vs. males; Age = age-related change with males as a reference; Age:F = the added age-related effect for females to the age-related change with males as a reference.

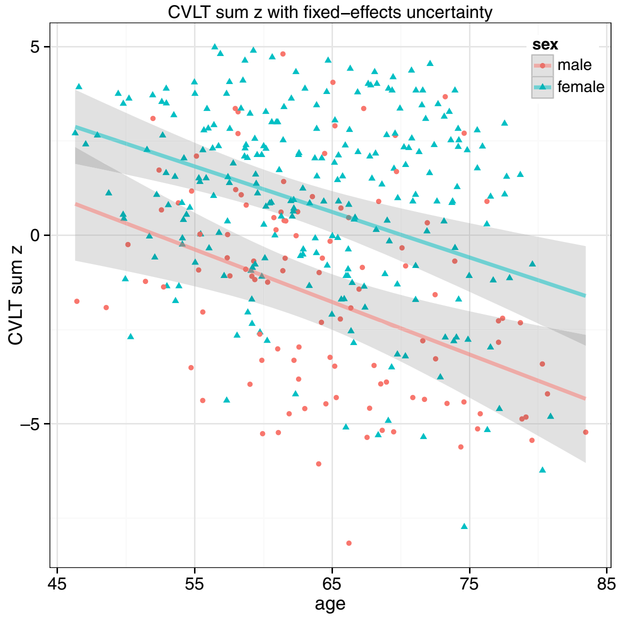

Fig. 2. The composite score *cvltSumZ* in males and females in the longitudinal analysis with fixed effects uncertainty.

and sex. The improved model fit was shown by a lower AIC or BIC score than when only the three basic variables were controlled for. The effects were, however, marginal for the *Primacy effect* and *Intrusions* (Table 5). The Kenward-Roger partial F-tests and the estimated regression weights showed strongest effect for *Semantic clustering*, *Consistency* and *False positive errors*. In the model adding *Semantic clustering* as a predictor, the estimated age-related decline was 0.10 units of the *cvltSumZ* score per year, and *cvltSumZ* was estimated to increase by 0.51 units for each additional unit of the *Semantic clustering* score. The add-on effect of the *Consistency* score was 0.10 units per year, and for each additional *False Positive error*, the *cvltSumZ* was estimated to drop by 0.35 units. (Table 5).

DISCUSSION

Cross-sectional analyses of CVLT-II results in the current sample showed a higher function in females than males across all ages, with a steeper age-related decline in males. This was found on all primary measures of episodic memory function, the learning strategy measure of Consistency and the error measure from the recognition trial (False positive errors). The main pattern of age-related change was confirmed by the longitudinal analyses, but with a somewhat smaller age-related change and a much lower difference between the two sexes. Finally, information about learning strategy and accuracy in the third wave was shown to add to the age-related change across the three study waves. By using a statistical analysis taking both group mean, individual differences and learning effect into account, the findings of the present study should give a valid estimate of change in episodic memory function in a population of healthy middle-aged and older adults. The impact of accuracy and learning strategies on change from previous study waves should be considered by clinical neuropsychologist evaluating individuals reporting that they already have experienced change in episodic memory function.

Overall, the pattern of age-related change was confirmed both by the cross-sectional and longitudinal analyses, with a somewhat lower change in the latter. A similar pattern of change was shown for some of the measures of learning strategy and accuracy in the cross-sectional analyses. However, the main contribution of the present study was the results from the longitudinal analyses using a mixed effects regression model, letting group effects be moderated by individual differences and vice versa. By this, it should give a reasonable estimate of age- and sex-related changes that can be generalized to other groups of middle-aged and older participants, although the relatively few males included in the study may limit the power of the sex-related results. The importance of using a mixed model was emphasized by detecting a substantial variance in individual intercepts, meaning that a strong effect of

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.

Table 5. *Contributions from learning strategy and errors to explain prior change in episodic memory function with Kenward-Roger partial F-test and regression weights.*

Predictor variables the regression model	AIC[a]	BIC[a]	F-test	df2	*p*-value	Age[b]	Add[b]	t-test[b]
Learning strategy and errors								
cvltRe	1377	1392	2.2	205.0	00.14			
Age	1346	1365	34.6	144.8	< 0.001			
Sex	1327	1349	23.5	100.1	< 0.001			
Semantic clustering	1297	1323	38.5	100.0	< 0.001	–0.10	0.51	6.2
Consistency	1291	1318	46.3	99.8	< 0.001	–0.10	0.13	6.8
Primacy effect	1324	1351	4.4	99.7	0.04	–0.14	–0.07	–2.1
Intrusions	1322	1348	7.2	99.5	0.009	–0.14	–0.09	–2.7
False positive errors	1279	1305	68.1	100.3	< 0.001	–0.08	–0.35	–8.3

Note: [a]AIC = Akaike information criterion; BIC = Bayesian information criterion. **Regression weights. [b]Age = the regression weight estimate of Age; **Add = the added regression weight estimate when cvltRe, age and sex were controlled for; **t-test: t-test for the estimate of the learning strategy and error measures.

individual differences may well have influenced results from the cross-sectional analyses. The inclusion of information about the learning-effect in the statistical model was another important aspect of the study, and may explain some differences between our results and results in previous studies. The age-related change in the present main longitudinal analysis was somewhat different from the results in studies documenting a rather stable cognitive function until the individual is more than 60 years old (e.g., (Nilsson, Sternang, Rønnlund & Nyberg, 2009; Rønnlund *et al.*, 2005; Schaie, 2005; Zelinski & Stewart, 1998). Interestingly, some longitudinal studies have shown improvement in cognitive function over time in middle-aged and older adults (e.g., Flicker, Ferris & Reisberg, 1993). This was also found in the present study, but was interpreted as an improvement due to repeated assessment. We argue that the statistical procedure presented by Ferrer *et al*. (2004) gave us a feasible alternative to more resource intensive designs, such as inclusion of new samples when the original longitudinal participants are tested at the second time (e.g., Rønnlund *et al*., 2005) or by using long intervals between test sessions (e.g., Rabbitt, Diggle, Smith, Holland & Mc Innes, 2001).

Finally, the present study showed the importance of information about learning strategy and accuracy in explaining change across the three study waves. The results were convincing by showing a high contribution from almost all included variables. The strongest effect was found for number of false positive errors on the recognition trial. This may be related to what Schacter Koutstaal & Norman (1997) described as a decline in how distinctly older people encode information. It may also explain why errors on the recognition task are more prominent than on the recall measures: answers on the recognition trial rely highly on familiarity (see Martin-Ordas & Call, 2013), giving the participants cues that are on the top of what Stretch and Wixted (1998) referred to as a 'strength-of-evidence axis'. When participants are uncertain about an answer, but find the word familiar, it may be tempting to give a positive rather than a negative answer. Although age-related changes in learning strategy and accuracy may be part of normal cognitive aging, the present study showed an add-on effect: a less active learning strategy and a high number of errors were related to a change in episodic memory function that was more substantial than the change expected from age and sex alone. In the present study we used information about learning strategy and accuracy to explain change from previous study waves, assuming that this may correspond to the information obtained by a neuropsychologist evaluating an individual who reports change in memory function. Future studies of their value in predicting future decline are necessary before concluding about its significance as a symptom of a neurodegenerative disorder.

There are several limitations to the present study. First of all, the study has not included information about biological factors of importance to cognitive change in samples of presumably healthy individuals, for example, hormonal sex-differences (Sherwin, 1998; Shaywitz, Shaywitz, Pugh *et al*., 1999), small age-dependent vascular lesions (Papp, Kaplan, Springate *et al*., 2013), genetic factors (Bender & Raz, 2012, Reinvang, Espeseth & Westlye, 2013) and sensory problems (Rønnberg *et al*., 2011). In addition, there are several other essential factors, which are well described in cognitive reserve- (Stern, 2002, 2009), scaffolding (Park & Reuter-Lorenz, 2009) and brain maintenance (Nyberg *et al*., 2012; Pudas, Persson, Josefsson, de Luna, Nilsson & Nyberg, 2013) models of cognitive aging. The composite score of episodic memory function was calculated from subtests from CVLT-II. Although this may restrict our conclusions to results on CVLT-II, we will argue that it may also represent a strength: the scores were expected to tap the same cognitive domain and to have the same sensitivity and level of difficulty, which may be difficult to obtain when using tests from different test-batteries and traditions. Although we showed that CVLT-II was sensitive to age-related cognitive changes, information about other aspects of memory function (see Wang, Li, Li & Zhang, 2013), attention and executive function (Reinvang, Deary, Fjell, Steen, Espeseth & Parasuraman, 2010a, Reinvang *et al*. 2010b) are of importance before concluding about cognitive change in an individual. Finally, we used a linear regression model, which gave us a good approximation of cognitive change. More complicated regression models could be employed, for example by using quadratic terms and splines. This would, however, require a larger sample and should also include more waves.

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.

CONCLUSIONS

The CVLT-II identified cognitive change in a sample of healthy middle-aged and older individuals. The lower age-related change and sex difference in the longitudinal than cross-sectional analyses were probably related to using a linear mixed effects model taking both individual and group factors into account. The add-on value of information about learning strategy and accuracy in explaining age-related change in episodic memory function emphasizes the importance of taking such qualitative characteristics into account when evaluating patients reporting cognitive change. Assuming that these measures reflect abilities to associate and organize the words in the learning list, it is obvious that aspects of attention and executive function are involved in the observed cognitive change. Further longitudinal studies of larger cohorts including a wide range of individual factors in a proper statistical framework are thus warranted.

We thank Professor Ivar Reinvang for initiating the first wave of the study on cognitive aging and for providing fruitful collaboration with the Department of Psychology, University of Oslo. The study was financially supported by Western Norway Health Authority (grant 911397 and 911687 to AJL, and grant 911461 to EW).

REFERENCES

Bäckman, L., Small, B. J. & Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*, *124*, 96–102.

Bates, D., Maechler, M. & Bolker, B. (2013). *lme4: Linear mixed-effects models using S4 classes* [Computer software]. Retrieved 1 September 2013 from https://github.com/lme4/lme4/ (R package version 0.999999-2).

Bender, A. R. & Raz, N. (2012). Age-related differences in episodic memory: A synergistic contribution of genetic and physiological vascular risk factors. *Neuropsychology*, *26*, 442–450.

Benedict, R. H. & Zgaljardic, D. J. (1998). Practice effects during repeated administrations of memory tests with and without alternate forms. *Journal of Clinical and Experimental Neuropsychology*, *20*, 339–352.

Blacker, D., Lee, H., Muzikansky, A., Martin, E. C., Tanzi, R., McArdle, J. J., *et al.* (2007). Neuropsychological measures in normal individuals that predict subsequent cognitive decline. *Archives of Neurology*, *64*, 862–871.

Davis, H. P., Klebe, K. J., Guinther, P. M., Schroder, K. B., Cornwell, R. E. & James, L. E. (2013). Subjective organization, verbal learning, and forgetting across the life span: From 5 to 89. *Experimental Aging Research*, *39*, 1–26.

Delis, D. C., Kramer, J. H., Kaplan, D. & Ober, B. A. (2000). *California Verbal Learning Test* (2nd edn). San Antonio, TX: Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, D. & Ober, B. A. (2004). *California Verbal Learning Test (CVLT-II)*. Pearson Assessment: Norwegian manual supplement. Stockholm.

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

Finkel, D., Reynolds, C. A., Larsson, M., Gatz, M. & Pedersen, N. L. (2011). Both odor identification and ApoE-ε4 contribute to normative cognitive aging. *Psychology and Aging*, *26*, 872–883.

Flicker, C., Ferris, S. H. & Reisberg, B. (1993). A longitudinal study of cognitive function in elderly persons with subjective memory complaints. *Journal of the American Geriatrics Society*, *41*, 1029–1032.

Galecki, A. & Burzykowski, N. (2013). *Linear mixed-effects models using R: A step-by-step approach*. New York: Springer.

Greenaway, M. C., Lacritz, L. H., Binegar, D., Weiner, M. F., Lipton, A. & Munro Cullum, C. (2006). Patterns of verbal memory performance in mild cognitive impairment, Alzheimer's disease, and normal aging. *Cognitive and Behavioral Neurology*, *19*, 79–84.

Halekoh, U. & Højsgaard, S. (2013). *pbkrtest: Parametric bootstrap and Kenward Roger based methods for mixed model comparison* [Computer software]. Retrieved 1 September 2013 from http://CRAN.R-project.org/package = pbkrtest (R package version 0.3-5).

Herlitz, A., Nilsson, L. G. & Bäckman, L. (1997). Gender differences in episodic memory. *Memory and Cognition*, *25*, 801–811.

Kaltreider, L. B., Cicerello, A. R., Lacritz, L. H., Honig, L. S., Rosenberg, R. N. & Cullum, M. C. (2000). Comparison of the Cerad and CVLT list-learning tasks in Alzheimer's disease. *Clinical Neuropsychology*, *14*, 269–274.

Lacritz, L. H., Cullum, C. M., Weiner, M. F. & Rosenberg, R. N. (2001). Comparison of the Hopkins Verbal Learning Test-revised to the California Verbal Learning Test in Alzheimer's disease. *Applied Neuropsychology*, *8*, 180–184.

Levine, B., Svoboda, E., Hay, J. F., Winocur, G. & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*, 677–689.

Long, J. (2012). *Longitudinal data analysis for the behavioral sciences using R*. Iowa City, IO: Sage.

Martin-Ordas, G. & Call, J. (2013). Episodic memory: A comparative approach. Frontiers in Behavioral Neuroscience, 7, 63. http://dx.doi.org/10.3389/fnbeh.2013.00063.

Nilsson, L.-G., Sternang, O., Rønnlund, M. & Nyberg, L. (2009). Challenging the notion of an early-onset of cognitive decline. *Neurobiology of Aging, 30*, 521–524; discussion 530–533.

Nyberg, L., Lövden, M., Riklund, K., Lindenberger, U. & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Science*, *16*, 292–305.

Papp, K. V., Kaplan, R. F., Springate, B., Moscufo, N., Wakefield, D. B., Guttmann, C. R. G., *et al.* (2013). Processing speed in normal aging: Effects of white matter hyperintensities and hippocampal volume loss. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *21*, 197–213.

Park, D. C. & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*, 173–196.

Petersen, R. C. (2004). Mild Cognitive Impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*, 183–194.

Pudas, S., Persson, J., Josefsson, M., de Luna, X., Nilsson, L.-G. & Nyberg, L. (2013). Brain characteristics of individuals resisting age-related cognitive decline over two decades. *Journal of Neuroscience*, *33*, 8668–8677.

R Development Core Team (2013). R: *A Language and Environment for Statistical Computing* [Computer software manual]. Vienna, Austria. Retrieved xxxxx form http:// www.r-project.org/

Rabbitt, P., Diggle, P., Smith, D., Holland, F. & Mc Innes, L. (2001). Identifying and separating the effects of practice and of cognitive aging during a large longitudinal study of elderly community residents. *Neuropsychologia*, *39*, 532–543.

Rabin, L. A., Par, N., Saykin, A. J., Brown, M. J., Wishart, H. A., Flashman, L. A., *et al.* (2009). Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *16*, 357–376.

Reinvang, I., Deary, I. J., Fjell, A. M., Steen, V. M., Espeseth, T. & Parasuraman, R. (2010a). Neurogenetic effects on cognition in aging brains: A window of opportunity for intervention? *Frontiers Aging Neuroscience*, *2*, doi:10.3389/fnagi.2010.00143.

Reinvang, I., Espeseth, T. & Westlye, L. T. (2013). APOE-related biomarker profiles in non-pathological aging and early phases of Alzheimer's disease. *Neuroscience and Biobehavioral Reviews*, *37*, 1322–1335.

Reinvang, I., Winjevoll, I. L., Rootwelt, H. & Espeseth, T. (2010b). Working memory deficits in healthy APOE epsilon 4 carriers. *Neuropsychologia*, *48*, 566–573.

Rønnberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternang, O., Wahlin, A., *et al.* (2011). Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory. *Journal of Speech, Language, and Hearing Research*, *54*, 705–726.

Rønnlund, M., Nyberg, L., Bäckman, L. & Nilsson, L.-G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging*, *20*, 3–18.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*, 507–514.

Salthouse, T. A. (2012). Robust cognitive change. *Journal of the International Neuropsychological Society*, *18*, 749–756.

Schacter, D. L., Koutstaal, W. & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Science*, *1*, 229–236.

Schaie, K. W. (2005). What can we learn from longitudinal studies of adult development? *Research in human development*, *2*, 133–158.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., Mencl, W. E., *et al.* (1999). Effect of estrogen on brain activation patterns in postmenopausal women during working memory tasks. *JAMA*, *281*, 1197–1202.

Sherwin, B. B. (1998). Cognitive assessment for postmenopausal women and general assessment of their mental health. *Psychopharmacology Bulletin*, *34*, 323–326.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448–460.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*, 2015–2028.

Stretch, V. & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 1397–1410.

Sunderaraman, P., Blumen, H. M., DeMatteo, D., Apa, Z. L. & Cosentino, S. (2013). Task demand influences relationships among sex, clustering strategy, and recall: 16-word versus 9-word list learning tests. *Cognitive and Behavioral Neurology*, *26*, 78–84.

Tierney, M. C., Yao, C., Kiss, A. & McDowell, I. (2005). Neuropsychological tests accurately predict incident Alzheimer's disease after 5 and 10 years. *Neurology*, *64*, 1853–1859.

Twamley, E. W., Ropacki, S. A. L. & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society*, *12*, 707–735.

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language*, *108*, 175–183.

Wang, P., Li, J., Li, H. & Zhang, S. (2013). Differences in learning rates for item and associative memories between amnestic mild cognitive impairment and healthy controls. *Behavior and Brain Function, 9* 29, http://dx.doi.org/10.1186/ 1744-9081-9-29.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence. WASI manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2007). *Wechsler Abbreviated Scale of Intelligence (WASI).* Pearson Assessment: Norwegian Manual Supplement. Stockholm.

Zelinski, E. M. & Burnight, K. P. (1997). Sixteen-year longitudinal and time lag changes in memory and cognition in older adults. *Psychology and Aging*, *12*, 503–513.

Zelinski, E. M. & Stewart, S. T. (1998). Individual differences in 16-year memory changes. *Psychology and Aging*, *13*, 622–630.

Received 27 September 2013, accepted 21 January 2014

© 2014 The Authors Scandinavian Journal of Psychology published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.