A quick test of cognitive speed (AQT): regression-based norms for cognitively healthy 80 to 94-year olds

Elisabet Classon, Wobbie van den Hurk, Ewa Wressle, Inger Rehn & Maria M. Johansson


To link to this article: https://doi.org/10.1080/13825585.2021.1922585

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 13 Jun 2021.

Submit your article to this journal

Article views: 179

View related articles

View Crossmark data
A quick test of cognitive speed (AQT): regression-based norms for cognitively healthy 80 to 94-year olds

Elisabet Classon\textsuperscript{a,b}, Wobbie van den Hurk\textsuperscript{c,d}, Ewa Wressle\textsuperscript{e}, Inger Rehn\textsuperscript{c} and Maria M. Johansson\textsuperscript{a}

\textsuperscript{a}Department of Acute Internal Medicine and Geriatrics, and Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden; \textsuperscript{b}Research and development, Mindmore AB, Stockholm, Sweden; \textsuperscript{c}Department of Acute Internal Medicine and Geriatrics, Linköping University, Linköping, Sweden

\textbf{ABSTRACT}

Slowed processing speed is part of normal aging but also a symptom of many diseases, including dementia. A Quick Test of Cognitive Speed (AQT) consists of three conditions: color naming (AQT1), form naming (AQT2) and dual color-form naming (AQT3) and offers a user-friendly assessment of processing speed that is used internationally to identify cognitive impairment in elderly patients. Appropriate age-norms have however been lacking. This study provides regression-based norms derived from a Swedish sample of 158 cognitively healthy 80 to 94-year olds. The results show age effects in all three conditions, a non-linear education effect in AQT1, and age by gender interactions in AQT2 and AQT3: men performed worse with increasing age, but women remained on a par. However, irrespective of age and gender, AQT2 and AQT3 mean raw and predicted scores were slower than the hitherto recommended cutoff criteria for suspected cognitive impairment.

\textbf{Introduction}

There is an increasing demand for high-sensitivity cognitive screening instruments that can detect subtle declines associated with mild cognitive impairment (MCI), incipient dementia and other health conditions affecting the elderly (Mitchell & Malladi, 2010). Tests that are highly sensitive to disease-related cognitive decline tend however to be equally sensitive to factors such as education, cultural background, and aging per se (Hebben & Milberg, 2009). This dilemma poses high demands on the availability of adequate normative data for elderly samples, an area in which much work remains to be done. The current study contributes to that work by presenting hitherto lacking age-norms for 80 to 94-year olds on a test of processing speed with proven clinical utility, A Quick Test of Cognitive Speed (AQT; Wiig et al., 2002b).

Processing speed is a broad concept, basically determined by speed of neural transmission within and between brain networks. Like attention, processing speed is typically measured in relation to some timed task, the performance of which requires task-specific combinations of perceptual, motor, and/or cognitive operations that
cannot be easily disentangled. Reduced speed in one or all of the required operations, which also affects the necessary ability to rapidly carry them out simultaneously, sets a ceiling to performance (Salthouse, 1996). Generalized slowing of processing speed has been put forth as a primary mechanism for age-related cognitive change, affecting higher order abilities (Finkel et al., 2007; Salthouse, 1996). Links between white matter damage, reduced processing speed and functional decline are seen in normal aging (Classon et al., 2016; Lu et al., 2013; Makizako et al., 2015; Venkatraman et al., 2011), but accelerated slowing may be an early marker of degenerative disease (Lezak et al., 2012) and often one of the most prominent symptoms in neurocognitive disorders such as vascular dementia, dementia with Lewy bodies and Parkinson’s disease dementia. It is also an indication for treatment of cardiovascular risk factors (Venkatraman et al., 2011), which may in turn reduce risk of dementia (O’Brien & Thomas, 2015). The lack of items assessing processing speed and executive functions in the Mini-Mental State Examination (MMSE; Folstein et al., 1975), still the most used cognitive screen, has been criticized over the years. This is however a disadvantage it shares with the majority of screening tests used in the elderly population (Cullen et al., 2007) and one reason why cognitive decline may remain undetected. Sensitivity in primary care dementia investigations can, for example, be improved by complementing the MMSE with AQT (Kvitting et al., 2013).

AQT is a test that has the advantage of being short and easy to administer and interpret with minimal training and an outcome measure, time in seconds, sensitive enough to capture discreet changes without floor or ceiling effects. It consists of three continuous performance rapid automatized naming (RAN) conditions, where the first two are single dimension (naming of color, naming of form), and the third dual dimension (naming of both color and form). Seemingly simple, efficient rapid automatized naming – mainly examined in children – has been proposed to require swift and precisely timed synchronization and integration of a combination of visual-verbal, attentional and executive processes, along with motor response preparation and activation (Cutting & Denckla, 2001; Norton & Wolf, 2012). The few brain imaging studies that have examined neural correlates of RAN have found activations in networks related to control of eye movements, attention and reading, including areas associated with motor planning, semantic access and articulation (Cummine et al., 2015; Misra et al., 2004). Dual dimension conditions puts additional demands on co-articulation, working memory and set shifting (Norton & Wolf, 2012; Wiig et al., 2005).

There are virtually no AQT practice effects, test–retest reliability is high (Humes et al., 2013) and cultural bias reportedly low (Dsurvey, 2007; Takahashi et al., 2012; Wiig et al., 2002a) although differences in syllable length across languages, and thereby naming speeds, motivate language-specific norms (Wiig & Al-Halees, 2013). Literacy is not required even if it facilitates performance (Nielsen & Wiig, 2006) and correlations with educational levels beyond 8 years of schooling are typically low, but not in all samples (Nielsen & Wiig, 2006; Petrazzuoli et al., 2014; Takahashi et al., 2012; Wiig & Al-Halees, 2013).

No gender effects have been reported (Nielsen & Wiig, 2006; Petrazzuoli et al., 2014; Subirana-Mirete et al., 2014; Wiig et al., 2007) but age has consistently been associated with decreasing AQT performance, as would be expected in a measure of processing speed. The effect has been described as small, in the range of 1 to 6 s/decade depending
on task and sample (Nielsen & Wiig, 2006; Subirana-Mirete et al., 2014; Wiig et al., 2007). Slowed AQT performance has been associated with cortical thinning in occipital and parieto-temporal cortical areas, reduced parieto-temporal regional cerebral blood flow and reduced connectivity in posterior networks (Jalakas et al., 2019; Park et al., 2018; Warkentin et al., 2008).

With respect to neurodegenerative disorders, AQT has been found to: correlate with other dementia screens but have higher sensitivity; help predict conversion to dementia; discriminate between dementia and normal aging as well as between different types of dementia; and be useful in evaluating treatment effects of cholinesterase inhibitors or Memantine (Aarsland et al., 2009; Andersson et al., 2007; Jalakas et al., 2019; Kvitting et al., 2013; Palmqvist et al., 2010; Park et al., 2018; Takahashi et al., 2012). In studies on cardiovascular disease, AQT has proven sufficiently sensitive to capture subtle differences in processing speed associated with diabetes and cardiovascular disease (Dybjer et al., 2018; Nilsson et al., 2016, 2014; Tufvesson et al., 2013).

AQT performance in normal aging beyond the seventies is however sparsely examined. To our knowledge there are only three normative studies that have included individuals at or above the age of 80 (Petrazzuoli et al., 2014; Subirana-Mirete et al., 2014; Wiig et al., 2007), but in each case as part of a single, wide age-range group (above 59, 70 and 75 years, respectively), suggesting few elderly participants. This is specified in the study by Wiig et al. (2007) which included only three individuals older than 85 years. Collapsing participants across wide age-ranges, typically with considerably fewer old-old (85 years and upwards) than younger participants in the oldest age-group, is common in normative studies but should raise concerns. There is good evidence that age-related declines in cognitive abilities, perhaps particularly processing speed, proceeds and accelerates with advancing age (Bäckman et al., 2004; Cohen-Mansfield et al., 2013; McCarrey et al., 2016; Small et al., 2011). Calculating normative trajectories for the old-old based on findings from younger age groups is therefore associated with a risk of overestimating expected performance. Because AQT is currently used in dementia investigations and other examinations of cognitive health that increasingly involve old-old individuals, this may entail consequences for the individual patient. That said, two of the three norming studies conducted so far (Petrazzuoli et al., 2014; Wiig et al., 2007) generally agree with the recommended cutoffs for risk of cognitive impairment at 35s for the single-dimension tasks, and 70s for the dual-dimension task, irrespective of age (Wiig et al., 2005). But the mean values and standard deviation reported in Subirana-Mirete et al. (2014) rather indicate that these cutoffs are within the range of normal performance in cognitively healthy individuals over 60 years of age.

To summarize, AQT is an interesting tool for evaluation of cognitive impairment in the elderly over broad populations and health conditions, as well as for examining disease progression and treatment outcomes. Age-norms of adequate quality and resolution for individuals above their seventies have however been lacking. The purpose of the current study was to present updated, regression-based, AQT norms for a sample of cognitively healthy Swedish community-dwelling 80 to 94-year olds.
Materials and methods

The data reported below were collected as part of a larger normative study, including two additional cognitive tests: the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) and Cognistat (Kiernan et al., 1987). Results on those will be reported separately for the sake of clarity and readability.

Participants and procedure

A total of 987 mail invitations were sent to a pseudo-randomized selection of Linköping county inhabitants, evenly distributed over gender and age group (80–84 years, 85–89 years, and 90–94 years), acquired from the Swedish population registry. Of these, 218 agreed to a structured telephone interview conducted by experienced staff members at the Linköping university hospital memory clinic. The questions covered living condition, health status, cognitive functioning, sensory and motor functioning, Swedish language proficiency if other native tongue, and consent to supplement the self-reported information with that of medical records (diagnoses). Exclusion criteria were MCI, dementia, or other known cognitive deficit (e.g., following stroke or other neurological or non-neurological disorder), current substance abuse or ongoing moderate-to-severe psychiatric illness. Impaired perceptual or motor ability or lack of knowledge in the Swedish language were also grounds for exclusion if judged to entail non-valid test results.

The health conditions covered by the interview and medical records search were current or previous substance dependency or neurocognitive, cerebrovascular, neurological, psychiatric, or other (e.g., chronic obstructive pulmonary disease (COPD), thyroid or kidney disorders) disease. Note that, apart from ongoing moderate-to-severe psychiatric illness or substance dependency, presence of one or more of these health conditions was accepted, unless coupled with evidence of subjective or objective cognitive decline. The purpose was to enroll representatives of the general cognitively healthy elderly population but with information about their health condition. Individuals with a verified stroke or intracranial hemorrhage within the last 5 years were however consistently excluded.

Eligible participants (n = 181) completed a 1.5-hour visit at the Linköping university hospital memory clinic, in a quiet and well-lit room. Participants with glasses and/or hearing aids were instructed to bring them and use them under the session, help with settings was provided if needed. The test leader, one of five experienced staff members engaged in the project, collected written informed consent and conducted a structured interview concerning educational and vocative attainment. This was followed by administration of the study tests, always beginning with AQT, and followed by the MMSE, the MoCA and Cognistat in counterbalanced order. After a coffee break, the questionnaires ended the session. Test leaders were not blind to participant age.

To further ensure that only cognitively healthy participants were included, the 23 individuals who scored below 26p on the MMSE (Kvitting et al., 2019) and/or above 9p on the PHQ-9 self-rating scale of depression, corresponding to more than mild depressive symptoms (Kroenke et al., 2001), were exempted from further analyses. These individuals did not differ from the included participants in gender, age, or years of education, but
they performed on average around 20s slower (worse) in each AQT condition ($U_{AQ1} = 2921, z= 4.70, p< 0.001; U_{AQ2} = 2752, z= 4.00, p< 0.001; U_{AQ3} = 2927, z= 4.73, p< 0.001$).

To summarize, cognitive health was assessed by telephone interview, medical records search, MMSE >25p, PHQ-9 < 9p and clinical assessment. Please see Figure 1 for a detailed description of recruitment, drop-out and exclusion.

The final set of 158 participants were evenly distributed over gender, mean age was 85 and they had in average 11 years of education. Three participants were not native Swedish speakers but had lived 57–71 years in this country and spoke Swedish fluently. More detailed demographic information about the participants and their health status are listed in Table 1 and Table 2.

**Measurements**

**AQT**

AQT consists of three rapid automatized naming conditions (hereafter called AQT1-3) that were administered according to the manual (Wiig et al., 2005). Thus, each

![Flow-chart of recruitment, drop-out and exclusion.](image-url)
Table 1. Demographic information and AQT performance divided by age group, by gender, and in total. Presented as mean, SD or number, (%).

<table>
<thead>
<tr>
<th></th>
<th>By age group</th>
<th>By gender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80–84</td>
<td>85–89</td>
</tr>
<tr>
<td>n (women)/n (%)</td>
<td>78 (37)</td>
<td>47 (25)</td>
</tr>
<tr>
<td>AQT performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AQT1</td>
<td>28 47</td>
<td>30 10.6</td>
</tr>
<tr>
<td>AQT2</td>
<td>39 7.4</td>
<td>38 7.6</td>
</tr>
<tr>
<td>AQT3</td>
<td>73 16.1</td>
<td>78 17.8</td>
</tr>
<tr>
<td>Age</td>
<td>82 1.4</td>
<td>87 1.4</td>
</tr>
<tr>
<td>Years of education</td>
<td>12 3.8</td>
<td>10 3.6</td>
</tr>
<tr>
<td>Global cognition (MMSE)</td>
<td>28 1.3</td>
<td>28 1.3</td>
</tr>
<tr>
<td>Self-rated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>depression (PHQ-9)</td>
<td>2 2.0</td>
<td>2 2.2</td>
</tr>
<tr>
<td>health (EQ-VAS)</td>
<td>77 16.2</td>
<td>77 14.7</td>
</tr>
<tr>
<td>functional independence (FAQ)</td>
<td>1 1.2</td>
<td>2 2.5</td>
</tr>
<tr>
<td>Visual function coded as</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fair</td>
<td>49 (58)</td>
<td>23 (27)</td>
</tr>
<tr>
<td>intermediate</td>
<td>27 (44)</td>
<td>20 (32)</td>
</tr>
<tr>
<td>poor</td>
<td>2 (18)</td>
<td>4 (36)</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>primary school</td>
<td>18 (23)</td>
<td>24 (51)</td>
</tr>
<tr>
<td>secondary school</td>
<td>35 (45)</td>
<td>16 (34)</td>
</tr>
<tr>
<td>higher education</td>
<td>25 (32)</td>
<td>7 (15)</td>
</tr>
<tr>
<td>Living situation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>living alone</td>
<td>31 (40)</td>
<td>24 (51)</td>
</tr>
<tr>
<td>living together</td>
<td>47 (60)</td>
<td>23 (49)</td>
</tr>
<tr>
<td>Geography</td>
<td></td>
<td></td>
</tr>
<tr>
<td>city</td>
<td>65 (83)</td>
<td>40 (85)</td>
</tr>
<tr>
<td>smaller community</td>
<td>8 (10)</td>
<td>5 (11)</td>
</tr>
<tr>
<td>countryside</td>
<td>5 (6)</td>
<td>2 (4)</td>
</tr>
</tbody>
</table>

**Bold** = p < 0.05 as tested by GLMs followed by Bonferroni post hoc, Kruskal–Wallis (FAQ), T-test (gender) or Chi² (educational level, living situation and geography).
† = not statistically analyzed due to less than 5 expected counts in 1 or more cells.
Table 2. Health status in terms of conditions diagnosed in medical records, divided by age group, by gender, and in total.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>80–84</th>
<th>85–89</th>
<th>90–94</th>
<th>p</th>
<th>Male</th>
<th>Female</th>
<th>p</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>50</td>
<td>29</td>
<td>22</td>
<td>67</td>
<td>49</td>
<td>51</td>
<td>61</td>
<td>101</td>
</tr>
<tr>
<td>Heart disease</td>
<td>44</td>
<td>24</td>
<td>20</td>
<td>61</td>
<td>50</td>
<td>37</td>
<td>48</td>
<td>88</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>32</td>
<td>20</td>
<td>12</td>
<td>36</td>
<td>29</td>
<td>33</td>
<td>36</td>
<td>64</td>
</tr>
<tr>
<td>Diabetes</td>
<td>16</td>
<td>6</td>
<td>5</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td>Thyroid disorder</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>12</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>Depression/anxiety</td>
<td>8</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Stroke</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Kidney disorder</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>COPD</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>15</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>TIA</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Confusion</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

TIA = Transient Ischemic Attack, COPD = Chronic Obstructive Pulmonary Disease. Distribution of each diagnosis across age group or gender were examined by Chi², number of diagnoses/participant by GLM (age group) or t-test (gender). Bold = p < 0.05 † = not statistically analyzed due to less than 5 expected counts in 1 or more cells.

Condition was preceded by a short practice trial to ensure ability to accurately identify and name all stimuli (colors and forms) appearing in the task. The actual test stimuli were then presented on a laminated test plate containing 40 stimuli in eight rows of five. In AQT1 the stimuli are blue, yellow, red, and black squares (green is not included to allow for red-green color blindness), in AQT2 they are black lines, squares, triangles and circles, and in AQT they are combinations of the colors and forms of the previous conditions (e.g., blue line, red square, etc.). Before each condition, participants were instructed to name the color, the form, or the color-form combination (in that order) as quickly and accurately as possible, from left to right and top to bottom as in reading. A digital stopwatch was used for timing, starting the moment the participant began naming the first item and stopping the moment the last item was named. The entire test takes approximately 10 minutes to perform and the score is time in seconds to complete each condition. Errors were recorded but not further analyzed in line with the typical procedure for RAN tasks (Norton & Wolf, 2012).

Background variables
The revised Swedish version of the MMSE, MMSE-SR (Palmqvist et al., 2013), was used as a screen of global cognitive functioning. It consists of 20 items assessing orientation to time and place, memory, attention, language, and visuospatial ability. It yields a total score ranging from 0 to 30 with lower scores denoting more impaired cognition. Self-rated depressive symptoms were assessed using PHQ-9 (Kroenke et al., 2001), a short version of the Patient Health Questionnaire (Spitzer et al., 1999). Frequency of nine depressive symptoms over the past two weeks are reported together with an assessment of how symptoms have affected everyday life. Self-reported independence in 10 everyday activities was assessed by the Functional Activities Questionnaire (FAQ; Pfeffer et al., 1982). Summed total score, max 30p with higher scores denoting less independence, was used for analyses. The EQ-5D (EuroQolGroup, 1990) is a self – report inventory
designed to measure health outcome. It includes a visual analog scale (EQ-VAS) that was used to assess self-rated health. The scale runs from 100 (“Best imaginable health state”) to 0 (“Worst imaginable health state”). Vision was not objectively measured but roughly coded into three categories based on unstructured self-reports of function with the correction normally used: fair (generally experiencing that vision is not a problem in everyday life), intermediate (experiencing visual problems not adequately compensated for by glasses), and poor (experiencing vision as a problem in everyday life that glasses cannot compensate for).

**Statistical analyses**

All statistics were conducted using the IBM SPSS Statistics version 25.0 software. Demographic variables and gender differences were analyzed using independent samples T-tests, general linear models (GLMs) followed-up by Bonferroni *post hoc* Chi² tests, Mann–Whitney U-test or Kruskal–Wallis tests, followed-up by Bonferroni *post hoc*, as applicable. The level of significance was set at two-tailed *p* = 0.05.

Separate multiple linear regression analyses were conducted to obtain predicted scores corrected for demographic factors for each AQT condition. AQT performance (time in s), was regressed on candidate lower-order and higher-order predictors: age, years of education and gender; age² and all two-way interactions. Gender was coded as female = 1 and male = 0. To avoid multicollinearity, age and years of education were centered (i.e., individual age – mean age of entire sample rounded to integer) before computing age² and the interaction terms. Additionally, to explore a potential nonlinear effect of education, a second set of regressions were computed replacing linear education by nonlinear education (1/years of education). When this resulted in a new model, the model with nonlinear education was retained if it increased predictive capacity (R²) and visual inspection confirmed the improved fit of the model on the data (Bergman et al., 2016).

All regressions were conducted using the backwards selection model with p-to-remove set at >0.1. Lower-order predictors that were included in higher-order predictors of the final model were forced back into the model. Regression assumptions were checked by inspection of the distributions and normal probability plots of the residuals, intercorrelations, computation of tolerance and variance inflation factor (VIF). Tolerance should not be lower than 0.2 and VIF should not exceed 10 (Bowerman & O’Connell, 1990; Menard, 1995). Influential outliers were identified by visual inspection, computation of Cook’s distances (Di) and SDs from mean results. Homoscedasticity was assessed by visual inspection of plots of standardized predicted values and unstandardized residuals and by applying Levene’s tests to the residuals divided into tertiles. The unstandardized residuals measure the vertical distance between each raw score and the regression equation line, i.e., the positive or negative deviation between each individual raw score and it’s corresponding predicted score. The SDs of the residuals are thus measures of the present sample’s normal distributions of variances around the predicted scores, within which the relative deviations of individual raw scores can be defined. Homoscedasticity is violated when residual variance differs across the regression equation line (aging is for example, associated with worse cognitive performance and increasing individual variance, hence residual variance may be larger around higher than lower predicted values). This needs to be considered when computing normative values by using the residual SD for each tertile. If homoscedasticity is not violated, the same
SD can be used for all cases (please see a detailed description of how to compute individual AQT norms below, section Application of the normative model to individual raw scores).

**Ethics**

All participants gave their written informed consent. The study procedure was approved by the Research Ethics Committee of Linköping University, Sweden (D.nr 2016/394-31), and conducted according to the Helsinki Declaration.

**Results**

**AQT performance and demographics**

AQT raw score performance and demographics are detailed in Table 1 and health conditions diagnosed in medical records are presented in Table 2 (see also Supplementary materials where scores fully stratified by age and gender are presented in Table S1 and traditional norms in Table S2). It should also be noted that there were no test leader effects in either AQT condition. A few comments are that AQT2 and AQT3 mean raw performances were slower than the hitherto recommended cutoff criteria for suspected cognitive impairment at 35s and 70s: 66% and 57%, respectively, performed below cutoff in these conditions. Functional independence declined with age but depressive symptoms, self-rated health and frequency, as well as number, of diagnosed health conditions where generally similar across age groups and genders. As expected at these high ages, the majority had at some time received a cardiovascular diagnosis, most commonly hypertension. There were some substantial overlaps where, for example, 50 (32%) of those with hypertension also had hyperlipidemia. Analyses examining whether either specific diagnosis impacted AQT performance showed that presence of hypertension or heart disease had significant effects but only on AQT1 (hypertension: \( U= 3466, z= 2.04, p= 0.041 \); heart disease: \( U= 3805, z= 2.51, p= 0.012 \), slowing down processing speed with 2–3s. As noted, sensory function was not objectively measured but some degree of visual impairment was ubiquitous. Compared to those with “fair” vision, those with “intermediate” vision were in average 0–6s slower, and those with “poor” vision 4–8s slower, depending on AQT condition. Due to the small number of participants with “poor” vision, two categories: “fair” and “intermediate-poor” were created for statistical analyses. These found no significant differences in AQT1 or AQT2 performance but those with poorer vision were slower in AQT3 (\( M_{fair} = 73.88, M_{intermediate-poor} = 80.23; t(156) = -2.10, p= 0.038 \).

**Regression-based normative data**

The final regression models are presented in Table 3. Note that for AQT2, the backwards selection resulted in a model including only age and an age by gender interaction, but Table 3 shows the results when gender was forced back into the model. Multicollinearity was well within recommended levels, tolerance was ≥0.52 and VIF ≤1.95. There were however two influential outlier scores in AQT1, representing the only raw scores >5 SD above the mean and with \( DIs \) of 0.596 and 0.278. These two scores were excluded which improved model fit as assessed by visual inspection and \( R^2 \) (please see Figures S1 and S2).
Table 3. Final regression models for AQT performance.

<table>
<thead>
<tr>
<th>AQT condition</th>
<th>Predictor</th>
<th>B</th>
<th>SE B</th>
<th>p</th>
<th>( \beta )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQT 1</td>
<td>Constant</td>
<td>-0.185</td>
<td>10.225</td>
<td>0.010</td>
<td>0.204</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>0.318</td>
<td>0.122</td>
<td>0.010</td>
<td>0.180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonlinear education</td>
<td>31.932</td>
<td>14.175</td>
<td>0.026</td>
<td>0.180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-2.341</td>
<td>0.977</td>
<td>0.018</td>
<td>-0.186</td>
<td>0.113</td>
</tr>
<tr>
<td>AQT 2</td>
<td>Constant</td>
<td>-4.415</td>
<td>18.388</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>0.511</td>
<td>0.215</td>
<td>0.019</td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.213</td>
<td>1.269</td>
<td>0.867</td>
<td>-0.013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age*gender</td>
<td>-0.519</td>
<td>0.313</td>
<td>0.181</td>
<td>-0.181</td>
<td>0.035</td>
</tr>
<tr>
<td>AQT 3</td>
<td>Constant</td>
<td>-101.016</td>
<td>41.922</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>2.114</td>
<td>0.490</td>
<td>0.000</td>
<td>0.447</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-4.991</td>
<td>2.893</td>
<td>0.093</td>
<td>-0.128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age*gender</td>
<td>-2.154</td>
<td>0.713</td>
<td>0.003</td>
<td>-0.313</td>
<td>0.129</td>
</tr>
</tbody>
</table>

for plots and regression diagnostics. Homoscedasticity, as assessed by Levene’s Test, could be assumed for the residuals of the AQT1 and AQT2 models (\( p \geq 0.215 \)), but not for the AQT3 model (\( p = 0.008 \)). This violated homoscedasticity was however not confirmed by visual inspection of residuals and predicted scores. The small difference could be inflated by the unequal group sizes of the tertiles of predicted scores: one very large group with 111 individuals and two very small groups with 19 to 21 individuals. Instead of splitting the participants in arbitrary groups and calculate separate SD values, it was decided to continue with one SD value for each of the models. The SDs of the residuals are 5.939 for AQT1, 7.850 for AQT2, and 17.898 for AQT3.

The regressions showed that years of education reached significance as a predictor for AQT1 performance only (although there was also a significant correlation between years of education and AQT3, \( r = -0.17, p < 0.05 \)). Nonlinear education provided a better fit than linear, reflecting a positive effect of the first 8–10 years of attained education but with little additional benefit thereafter (please see Figure 2: b). For AQT1 there were also effects of linear age and gender, with \( B \)-values indicating a consistent 2s advantage for women together with a general slowing with increasing age for both genders, corresponding to around 3s per decade (see Figure 2: a). Significant effects of linear age were also found for AQT2 and AQT3, with \( B \)-values indicating around 5s and 20s slower performance per decade and condition, but together with an age by gender interactions of similar size: at 80, men were faster than women, but their performance declined with age, while that of women remained on a par. After the age of around 85 for AQT2 and 83 for AQT3 males therefore performed worse than females (Figure 2: c and d).

As shown by Tables 1 and 2, there were no obvious demographic explanation for these different trajectories. Women had fewer years of education (t(156) = 3.1, \( p = 0.036 \)) and were more inflicted by thyroid disease (\( \chi^2(1) = 16.99, p < 0.001 \)), factors that would rather have put them at a disadvantage. Diseases could become disproportionately more common in men than in women with increasing age, and demographic factors could vary as a function of gender and age group, potentially explaining the age by gender interactions. This was however not supported by the data. There was no age group by gender interactions in years of education, global cognition, number of diagnoses, self-rated depressive symptoms, health, or functional ability. Examination of gender distribution of hypertension and heart disease (i.e., the two conditions found to influence AQT
Figure 2. Normative models displaying effects of age, education, and gender. a) AQT1, effect of age. Line solidity indicates years of education: dashed = 8 years, solid = 16 years. b) AQT1, non-linear effect of education with worse performance particularly in individuals with few years of education. Line solidity indicates age: dashed = 82 years, solid = 92 years. c) and d) show effects of age, gender, and age by gender interactions in AQT2 and AQT3. Male performance declines with age while female performance remains stable. Line color indicates sex: red = female; blue = male. Shaded area = CI.
results) or level of visual function (“fair” vs “intermediate-poor”) across age groups did not reveal emerging gender differences with age. As a final check, the same regression analyses as used for the AQT norms were repeated with MoCA and Cognistat, the other two cognitive tests performed by the same sample, as outcome variables. Neither of these regressions showed significant age by gender interactions (results not shown here), suggesting they are specific for AQT.

**Application of the normative model to individual raw scores**

To facilitate for clinicians, predicted AQT scores by age and gender are displayed in Table 4 together with cutoffs for slower-than-typical performance (≥1 SD), suspected impairment (≥1.5 SD), and atypical/pathological performance (≥2 SD).

Note that years of education is fixed at 11 years in the Table 4 normative data. That doesn’t affect AQT2 or AQT3, for which no education effects were found, but needs to be considered for AQT1: if the test taker has less than 11 years of education (and particularly when 8 or fewer), computation of individual normative data using the two-step procedure presented below is recommended.

As explained in detail in Van Der Elst et al. (2006), normative data for individual test takers can be computed from the regression parameters and, in combination with the SDs

| Table 4. Regression norms stratified by age and gender. As noted, there was a nonlinear effect of years of education in AQT1 and the predicted AQT1 scores below are computed for 11 years of education. |
|---------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                                | AQT1        | AQT2        | AQT3        | Age         | 0 SD        | +1 SD       | +1.5 SD     | +2 SD       | 0 SD        | +1 SD       | +1.5 SD     | +2 SD       | 0 SD        | +1 SD       | +1.5 SD     |
| Men                            |             |             |             | 80          | 28          | 34          | 37          | 40          | 36          | 44          | 48          | 52          | 68          | 86          | 95          | 104         |
|                                |             |             |             | 81          | 29          | 35          | 38          | 41          | 38          | 46          | 50          | 54          | 74          | 92          | 101         | 110         |

![Table image with data](image-url)
of the residuals, indicate this individual’s standardized deviation in performance compared to the norms. This requires a two-step procedure where the first step is to calculate the predicted score for the test taker (TT in the equations below) which is computed as:

\[
TT \text{ predicted score} = \text{Constant} + (B_{\text{age} \times TT \text{ age}}) + [(B_{\text{nonlinear education} \times x1/TT \text{ education}})] + (B_{\text{gender} \times TT \text{ gender}}) + (B_{\text{age} \times \text{gender} \times (TT \text{ age} - 85)} \times TT \text{ gender})
\]

The constants and B-values and their signs (±) to use are found in Table 3. The second step is to standardize the difference between the actual test result (raw score time in s) and the predicted score of the test taker:

\[
z - \text{score} = (\text{raw score} - TT \text{ predicted score})/SD \text{ of residuals}
\]

As already noted, the SDs of residuals are 5.939 for AQT1, 7.850 for AQT2, and 17.898 for AQT3. For example, computation of normative AQT1 data for a male 91-year-old test taker with 6 years of education and a raw score of 45s:

\[
TT \text{ predicted score} = -0.185 + (0.318x91) + (31.932x(1/6)) - (2.341x0) = 34.075
\]

\[
z - \text{score} = (45 - 34.075)/5.939 = 1.84
\]

A z-score of 1.84 lies between 1.5 and 2 SD above (slower than) average and thus in the range of suspected impairment. Had the same test taker had 11 years of education, Table 4 shows that his score would have been more than 2 SD above average and thus in the range of pathological performance.

**Discussion**

Slowed processing speed is a natural part of normal aging, but also a symptom of many somatic or psychiatric disorders and, if accelerated, dementias. AQT offers a quick and user-friendly assessment of processing speed that is used internationally to differentiate between normal aging and cognitive impairment, also in patients above their seventies. Normative data for this age-group has however been lacking. The purpose of the present study was therefore to provide regression-based AQT norms for cognitively healthy 80 to 94-year olds. Results showed interacting effects of age and gender, with age-related slowing of processing speed in men but not in women in two of the three AQT conditions. The majority, irrespective of age and gender, also performed slower than the recommended cutoffs for suspected cognitive impairment in these conditions. Easy-to-use regression-based norms, stratified by age and gender, are provided in Table 4 and the results are discussed in detail below.

The nonlinear effect of years of education on AQT1 performance only (although AQT3 performance also correlated with years of education) is in line with previous findings that AQT is relatively independent of education once a certain level of attainment, around 8–10 years of schooling, has been reached (Nielsen & Wiig, 2006; Petrazzuoli et al., 2014; Wiig & Al-Halees, 2013). Processing speed tends to be less influenced by education than other cognitive domains, supporting a closer relation to neurobiological function than to acquired skills (Lu et al., 2013; Tucker-Drob et al., 2009). White matter tract myelin content has for example been put forth as a predictor of healthy adults individual variability in processing speed (Chopra et al., 2018).
Age has consistently been linked to AQT performance and linear age effects were found for all three AQT conditions even in the restricted age range of old-old individuals presented here. Previous studies have suggested a 1–6s increase (worsening) in AQT performance per decade (Nielsen & Wiig, 2006; Subirana-Mirete et al., 2014; Wiig et al., 2007) but the present results indicate steeper performance declines, with 3–20 s increases per decade, in the old-old. In AQT2 and AQT3, this was however only true for men: while their performance decreased with increasing age, women continued to perform on a par.

AQT gender effects have not been previously reported but the literature is not very extensive. As noted, we’ve only found three previous publications of normative data that include individuals above their seventies. These cover Spanish (Subirana-Mirete et al., 2014), Italian (Petrazzuoli et al., 2014), and multi-ethnic American (Wiig et al., 2007) samples and present traditional norms for discrete age-groups, with the elderly collapsed across age-spans of around 20 years or more (e.g., 60 to 85-year olds) and gender differences only examined by comparison of means over the entire samples. We have found no studies examining similar RAN tasks during aging; they are predominantly studied in children and dyslexics due to their capacity to predict reading ability and typically use other stimuli, such as letters. There is however a well-documented female advantage in verbal ability and processing speed as measured by other instruments (Daseking et al., 2017; Gerstorf et al., 2011; Irving, 2012; Siedlecki et al., 2019). McCarrey et al. (2016) found gender-dependent aging trajectories with steeper rates of decline in men as compared to women in visuospatial abilities and perceptuomotor speed and integration. Because women did not show steeper rates of decline than men in any domain, the results were interpreted to indicate a greater female resilience to cognitive aging. Other studies have found similar results, including in processing speed and executive function (Ritchie et al., 2016; Schaie, 1994; Zaninotto et al., 2018; Zelinski & Stewart, 1998) although results are mixed (e.g., Reas et al., 2017). If gender-dependent trajectories may also be indexed by AQT remains to be confirmed in longitudinal studies.

Women were faster than men at all ages in AQT1, while men were faster than women up to the age of around 83 to 85 years of age in AQT2 and AQT3, but not thereafter. These differences between AQT conditions are interesting. Speculatively, the consistent female advantage in color naming could reflect the female advantage in verbal performance and processing speed discussed above. AQT2 and AQT3 include naming of geometrical forms, arguably more related to visuospatial ability and presumably facilitated by life-long experience with naming of geometric forms. Men typically perform better than women in tasks loading on visuospatial ability and are more likely to have had technical jobs than women in these age cohorts, potentially explaining their initial advantage in AQT2 and AQT3 and subsequent decline (cf McCarrey et al., 2016). But it should be remembered that the regression models could only account for 3.5% to 12.9% of the total variation in performance and gender differences in z-scores did not exceed 1.4 SD. Indeed, the relatively low proportion of variance explained by demographic factors suggest AQT primarily captures differences in cognitive functioning.

Average performances of the present sample (collapsed across age groups for comparability) are 1–10s slower on AQT1, 3–13s slower on AQT2, and 18–29s slower on AQT3 than those of the oldest age groups of earlier normative studies (Petrazzuoli et al., 2014; Subirana-Mirete et al., 2014; Wiig et al., 2007). However, mean age is 6–12 years higher than that of the oldest age groups in the Mediterranean studies (not reported in the
American) and sample size is almost double that of the Spanish study, which included 90 individuals at or above 60 years of age and is the largest of the three in terms of elderly participants. The normative values for individuals older than 60 years of age presented by Subirana-Mirete et al. (2014) show that the recommended cutoffs for suspected impairment are within their normal performance ranges (i.e., within +1SD from the mean). Average performance of both men and women in the present sample were also consistently slower (worse) than recommended cutoffs for AQT2 and AQT3, with more than 55% qualifying for suspected impairment. The 35s cutoff for AQT1 is close to the +1SD predicted score for 80-year olds in the current sample and is exceeded by in total 14% of the participants. Similar cutoffs for AQT2 and AQT3 at around +1SD from predicted scores, 45s and 85s, would be exceeded by 15% and 22% of the participants.

While reliable detection of MCI requires high sensitivity, adequate specificity is equally important for an instrument to be clinically useful. Demographically adjusted norms increase specificity by reducing the influence of factors that are not disease specific, but at the same time reduce sensitivity when the demographic factors adjusted for are also risk-factors for the disease in question (Strauss et al., 2006). Age is indeed the most important risk factor for MCI and dementia, supporting arguments against using demographically adjusted norms, particularly for screening purposes where the intention is to identify all individuals at risk. On the other hand, if non-adjusted criteria, such as the recommended AQT cutoffs, are used as part of dementia diagnostics or as a basis for referral for in-depth investigation in specialized care, then there is an evident risk for misdiagnoses and over-investigation. This is especially true for a test like AQT which has also proved sensitive to non-dementia diseases that increase with aging (Dybjer et al., 2018; Nilsson et al., 2016, 2014; Tufvesson et al., 2013). The present data suggest that adjustment of AQT cutoffs for the old-old could increase specificity while retaining sensitivity as compared to the MMSE. While diagnostic studies are required to find optimum criteria with respect to specific diseases, availability of appropriate age-norms for the old-old is essential for a balanced clinical interpretation of individual test scores.

RAN tasks such as AQT index speeded integration of multiple neural processes: visual-verbal, attentional and executive together with motor response preparation and activation (Norton & Wolf, 2012). Processing speed has been found to directly contribute to RAN in children, with rate of articulation playing a minor role (Cutting & Denckla, 2001). Although similar neural networks seem to be engaged in older as in younger individuals, aging could differentially affect subprocesses such as speed of articulation, verbal retrieval, attentional or executive processes. Differences in processing demands between the AQT conditions during aging are however yet to be fully investigated. For example, AQT3 naming times are not simply an additive of AQT1 and AQT2 naming times. A separate measure, AQT3-(AQT1+AQT2), has been suggested to evaluate cognitive efficiency or shift cost (Nielsen & Wiig, 2011, 2013) but age-related slowing of single dimension naming times apparently distort the ratio (Nielsen & Wiig, 2011). These are interesting question for future studies, together with exploration of associations between AQT and other neuropsychological measures as well as brain activations in normal aging.

Strengths of the present study include the pseudorandomized recruitment from the aging population together with the careful screen of health status including a medical records search. The sample size is large for the age-range and the use of regression-based, as compared to traditional, norms avoids problems with arbitrary age-bands and varying
subsample sizes (Knight et al., 2006). There are also limitations: number of participants decrease with increasing age, reflecting the challenges of recruiting participants at or above their eighties. Further, exploration of biomarkers or follow-up testing to rule out incipient cognitive disease was beyond our scope and sample selectivity differences can never be ruled out in cross-sectional designs – the educational level of the present sample is, for example, relatively high. Combined with exclusion criteria that may be considered stringent (e.g., MMSE <26p), the present norms may also be considered stringent, especially for less advantaged populations.

Conclusion

AQT is a short and sensitive test of processing speed that is well tolerated by the old-old but performance declines with increasing age, particularly in males, and motivate the complementary use of age- and gender adjusted norms for individuals above their seventies. The normative values presented in the present paper will aid clinical interpretation of AQT performance in elderly patients.

Acknowledgments

We would like to thank the participants for their collaboration, sharing personal information and devoting their time and interest to this project. We would also like to thank our colleagues at the Linköping university hospital memory clinic for their enthusiastic help with recruitment and data collection.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Alzheimerfonden under Grant number AF-648031; and by Region Östergötland (RÖ) under Grant numbers RÖ-430341, RÖ-751381, RÖ-829391, RÖ-631391, RÖ-825821, FORSS-482934, and RÖ ALF Grant numbers RÖ-790811, RÖ-696631, RÖ-602761

ORCID

Elisabet Classon  http://orcid.org/0000-0003-1403-6357
Wobbie van den Hurk  http://orcid.org/0000-0003-0849-0158
Maria M. Johansson  http://orcid.org/0000-0003-4166-7269

References


