On the paradoxical decrease of self-reported cognitive failures with age

de Winter, JCF; Dodou, D; Hancock, PA

DOI
10.1080/00140139.2015.1019937

Publication date
2015

Document Version
Final published version

Published in
Ergonomics: an international journal of research and practice in human factors and ergonomics

Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: https://www.tandfonline.com/loi/terg20

On the paradoxical decrease of self-reported cognitive failures with age

J.C.F. de Winter, D. Dodou & P.A. Hancock

To cite this article: J.C.F. de Winter, D. Dodou & P.A. Hancock (2015) On the paradoxical decrease of self-reported cognitive failures with age, Ergonomics, 58:9, 1471-1486, DOI: 10.1080/00140139.2015.1019937

To link to this article: https://doi.org/10.1080/00140139.2015.1019937

Ergonomics, 2015
Vol. 58, No. 9, 1471–1486, http://dx.doi.org/10.1080/00140139.2015.1019937

On the paradoxical decrease of self-reported cognitive failures with age

J.C.F. de Winter[a]*, D. Dodou[a] and P.A. Hancock[b]

[a]*Department of BioMechanical Engineering, Delft University of Technology, The Netherlands;* [b]*MIT^2 Laboratories, Department of Psychology and the Institute for Simulation and Training, University of Central Florida, Orlando, FL, USA*

(*Received 9 September 2014; accepted 6 February 2015*)

The science of Human Factors and Ergonomics (HF/E) often relies on self-report. This is a cause for concern because subjective methods are inherently susceptible to bias. Here, we present, examine and discuss a puzzling association between age and self-reported cognitive failures as assessed with Broadbent's Cognitive Failures Questionnaire (CFQ). Despite many well-established age-associated forms of cognitive decline, older persons actually report almost equivalent, or even less, cognitive failures on the CFQ than younger persons. Our present analysis indicates that this paradox may be resolved through the fact that people show age-related learning/adaptation/compensation and by the observation that the CFQ measures peoples' beliefs with respect to an individually idiosyncratic reference. Yet, at the heart of the paradox may be the idea that people cannot remember their own cognitive failures, pointing to even greater concerns with all forms of subjective self-report and its use in HF/E.

Practitioner Summary: Scientists and practitioners often try to understand and improve human performance with the help of self-report questionnaires. Our paper discusses the validity of self-reported errors measured with the Cognitive Failures Questionnaire (CFQ). We look to resolve the curious paradox that older persons tend to report fewer failures than younger persons do.

Keywords: ergonomics tools and methods; memory; human error; cognitive impairment

Introduction: Self-report in ergonomics

What distinguishes Human Factors/Ergonomics [HF/E] (the study of people and technical systems) from pure Engineering is that in the overall study of operations HF/E often elicits nominally subjective responses from the involved human operators (Hancock, Weaver, and Parasuraman 2002). Self-reports can provide valuable insight into an individual's private perceptions and beliefs, and can be used to obtain information about events that happened in the past or otherwise cannot be measured using physical instrumentation. It is incumbent upon those in HF/E to examine the validity of subjective responses which exist alongside the veracity of nominally 'objective' performance measures with which we wish to compare them (Hoffman and Hancock 2014). Much is made of those studies which find strong concordance between subjective apperceptions and objective outcome, noting that these findings are 'strong' evidence of consistent assessment (Hancock, Weaver, and Parasuraman 2002). We tend to rely on such studies accordingly in setting standards and promulgating design advice for both simple and complex technical systems alike. However, subjective methods are not without their flaws, either extraneous to the precise methodology involved or intrinsic to each method itself.

We begin our analysis on cognitive failures from a point of departure featuring the form of anecdotal observations that are common in the everyday world beyond our laboratory thresholds. Some people claim they are good at finding their way through a city; others consider themselves poor in recalling names. Often, such remarks are simply made in a colloquial manner and prove largely inconsequential. However, in some cases self-reported cognitive functioning can have important and even crucial implications (and see Lesch and Hancock 2004). For example, self-reported cognitive complaints are used as a diagnostic criterion in assessing post-concussion syndrome (World Health Organization 1992). They have also been proposed for diagnosing mild cognitive impairment, dementia and post-traumatic stress disorder (Crooks et al. 2005; Pfeifer et al. 2009; Portet et al. 2006). With respect to all such subjective states, objective (ratio) measures of their veracity are hard to establish. But how accurate and trustworthy specifically are self-reported cognitive failures? This question is certainly of relevance to HF/E, a field that has always been preoccupied with understanding and preventing 'human error' (Fitts and Jones 1947; Hale, Stoop, and Hommels 1990; Sharit 2006; Sheridan 2008; Stanton and Stevenage 1998). Indeed, operator error prevention lies at the very heart of the closely associated aspiration for systems safety, which forms part of the on-going mandate for HF/E in general (and see Dul et al. 2012).

*Corresponding author. Email: J.C.F.deWinter@TUDelft.nl

Illusory superiority: the double curse of the less able

Based on comparisons of self-reported performance with a reference standard or nominally 'true' performance, numerous studies have concluded that people often suffer from 'illusory superiority', also called the 'above average effect' (Dunning, Heath, and Suls 2004; Kruger 1999; Kruger and Dunning 1999; Van Yperen and Buunk 1991). For example, most people believe they are above the level of the average driver, something that is, of course, logically impossible (Sundström 2008; Svenson 1981). Findings show that prisoners judge themselves as more prosocial in terms of kindness and morality than non-prisoners (Sedikides et al. 2014). People also generally consider themselves as having above-average health habits as well as below-average unhealthy habits (Hoorens and Harris 1998). Individuals with below average IQ tend to overestimate that IQ, and under-performers in humour, grammar and logic provide inflated estimates of their actual ability and performance (Kruger and Dunning 1999). This is called 'the double curse' (Dunning et al. 2003, 84), as people of low cognitive ability apparently not only lack the skills required for providing the correct answers but also lack the necessary meta-cognitive skill of recognising when an answer is false. Thus, people are not veridical when it comes to self-assessment (Lesch and Hancock 2004), at least on average. It has also been suggested that people tend to overestimate their ability when it comes to relatively well-practiced tasks, such as driving, but underestimate their ability for more open-ended tasks, such as playing chess (Kruger 1999).

However, much of the research reporting an over-estimation of performance by low achievers (or an under-estimation of performance by high achievers) lacks credibility. This is because results can be, at least to a degree, a manifestation of regression-to-the-mean; a statistical artefact, rather than a true psychological trait (Ackerman, Beier, and Bowen 2002). That is, as a result of random error, any score that is low (or high) on its first measurement tends statistically to be closer to the average score on a subsequent measurement. Thus, the degree to which subjective apperception accords with actual reality remains an important but yet to be fully determined relation.

Correlational validity of self-reports

It has been argued that the corpus of research documenting over-estimation of self-reported as opposed to actual cognitive abilities 'has only documented the experimental side of the scientific divide (which focuses on means and ignores individual differences)' (Ackerman, Beier, and Bowen 2002, 587). Correlational studies, however, have reported moderate-to-strong agreement between self-reported and objective performance. Ackerman, Beier, and Bowen (2002) found a moderately high correlation of 0.45 between self-knowledge and objective knowledge across 18 different domains. This followed upon their earlier work (Ackerman, Kanfer, and Goff 1995) in which they had reported a 0.58 correlation between self-assessed and objective math/spatial ability and a 0.42 correlation between self-assessed and objective verbal ability. A more recent meta-analysis conducted by Beaudoin and Desrichard (2011) reported a smaller but still significant mean correlation of 0.15 between self-rated memory efficacy and objective measures of such episodic memory. We might conclude that some humans can recognise their own foibles and capabilities but not all of the people, and not all of the time.

The Cognitive Failures Questionnaire (CFQ): A reliable measurement, but of dubious validity

In the present critical appraisal, we focus on a now well-known and extensively employed questionnaire specifically developed to measure self-reported cognitive failures, namely the Cognitive Failures Questionnaire (CFQ; Broadbent et al. 1982). The CFQ is certainly not the only questionnaire that has been used to assess cognitive mishaps and complaints (see e.g. the Absent-Mindedness Questionnaire [Reason and Mycielska 1982], the Everyday Memory Questionnaire [Sunderland, Harris, and Baddeley 1983], the Metamemory in Adulthood [Dixon, Hultsch, and Hertzog 1987], the Prospective Memory Questionnaire [Hannon et al. 1995], the Slips of Action Inventory [Reason 1984] and the Subjective Memory Questionnaire [Bennett-Levy and Powell 1980]), but the CFQ is probably the most widely used and so one reason for us to focus on this modal methodology.

Our non-systematic literature searches were conducted using Google Scholar, a service that provides full-text search and a comprehensive coverage of the literature (De Winter, Zadpoor, and Dodou 2014; Gehanno, Rollin, and Darmoni 2013; Shariff et al. 2013). By 31 December 2014, we found 2131 papers in Google Scholar that cited, mentioned or used the 'Cognitive Failures Questionnaire'. As a comparison, searching for 'Cognitive Failures Questionnaire' yielded 176 records in Scopus (search filter: abstracts), 363 records in ScienceDirect (search filter: all fields, that is, full-text) and 178 records in Web of Knowledge (search filter: topic). We believe this suggests that Google Scholar is the most appropriate search engine for our purpose (although we recognise that the period 1982–1995 may be somewhat underrepresented in Google Scholar, see De Winter, Zadpoor, and Dodou 2014). We also searched in the citing and cited papers of the original publication by Broadbent et al. (1982) as well as of the works of Rabbitt (1990), Rabbitt and Abson (1990), Reason (1993) and Rast et al. (2009).

Our searches were not limited to the CFQ per se, but also extended towards various objective and subjective memory measures and cognitive performance tests.

The CFQ itself consists of 25 items divided between perceptual (e.g. 'Do you have trouble making up your mind?'), memorial (e.g. 'Do you forget where you put something like a newspaper or a book?') and motoric functions (e.g. Do you drop things?). The CFQ instructs respondents in the following manner: 'We want to know how often these things have happened to you in the last six months' and responses are given via five-point Likert scales with anchors ranging from 'very often' to 'never'. The underlying idea here is that the CFQ assesses everyday events that do not proceed as intended. According to Broadbent et al. (1982) the 25 failures 'were assembled from events which had happened to ourselves or acquaintances, the aim being to find episodes which the majority of people would accept as occurring to them at least occasionally' (3).

The CFQ exhibits reasonable test-retest reliability, between $r = 0.49$ and 0.84 for periods ranging from 6 weeks to 2 years (Broadbent et al. 1982; Merckelbach et al. 1996; Rabbitt et al. 1995; Vom Hofe, Mainemarre, and Vannier 1998). It is meaningfully correlated with other self-reported measures. For example, in a sample of 94 students, Broadbent et al. (1982) reported moderate to strong correlations between their CFQ total score and a number of other questionnaires, including the Slips of Action Forms A & B ($r = 0.57$ and 0.58, respectively), the Absent-Mindedness Questionnaire ($r = 0.62$) and the Short Inventory of Memory Experiences ($r = 0.59$). Martin (1983) found that CFQ total score correlated 0.74 ($N = 20$) with the Short Inventory of Memory Experiences, and -0.64 ($N = 20$) with the Everyday Memory Questionnaire. Further, Reason and Lucas (1984) recorded a correlation of 0.46 ($N = 78$) between the CFQ and the Absentmindedness in Shops Questionnaire. Smilek, Carriere, and Cheyne (2010) found a 0.82 ($N = 363$) correlation of the CFQ total score with the Attention-Related Cognitive Errors questionnaire and a correlation of 0.68 ($N = 363$) with the Mindful Attention Awareness Scale.

The CFQ also correlates with other observations such as self-reported traffic accidents ($r = 0.25$, $N = 152$, Larson and Merritt 1991; $r = 0.06$, $N = 2379$, Larson et al. 1997; $r = 0.39$, $N = 240$, Wallace and Vodanovich 2003), driving aggression ($r = 0.20$, $N = 112$, Gulian et al. 1989; $r = 0.43$, $N = 33$ [men], and $r = 0.60$, $N = 33$ [women], Simon and Corbett 1996), the Boredom Proneness Scale (correlations between 0.09 and 0.44 for $N = 137$ and $N = 126$, Wallace, Vodanovich, and Restino 2003), the Epworth Sleepiness Scale ($r = 0.32$, $N = 137$ and $r = 0.28$, $N = 126$, Wallace, Vodanovich, and Restino 2003), behavioural inhibition of the Behavioural Inhibition/Activation System Scale ($r = 0.20$, $N = 222$, Flehmig et al. 2007) and the neuroticism and psychoticism scales of the Eysenck Personality Questionnaire ($r = 0.26$ and 0.27, respectively), but not with extraversion: $r = 0.03$, $N = 222$ (Flehmig et al. 2007). However, correlations between the CFQ and other self-reports are not necessarily trustworthy. This is because correlations between two self-report questionnaires can be influenced by common method variance (Podsakoff et al. 2003). That is, specific types of overall reporting tendency could either inflate or deflate correlations independent of the actual scales themselves. This, as opposed to any underlying 'ground truth' relationship. In an attempt to circumvent common method variance, Broadbent et al. (1982) investigated the congruence between self-reported CFQ ratings and ratings of the respondents' marital partners. Results revealed that these two types of CFQ ratings significantly inter-correlated ($r \approx 0.34$), indicating that 'the respondent's own view of his or her liability to cognitive failure does tend to be shared by somebody who has a good opportunity to judge' (9).

CFQ scores are, however, known to positively correlate with some facets of objective performance such as sustained attention ($r = 0.21$, weighted mean across 6 studies, Smilek, Carriere, and Cheyne 2010) and measured brain activity as assessed via fMRI (correlations ranging from $r = 0.26$ to $r = 0.36$, $N = 71$ for various brain regions, Garavan et al. 2006), and negatively with the inhibitory neuro-transmitter GABA in the visual cortex as measured with MRI ($r = -0.42$, $N = 36$, Sandberg et al. 2014). These linkages suggest that higher capacity for cognitive inhibition and suppression of distractions are associated with lower CFQ scores. Hohman et al. (2011) found that higher CFQ scores were associated with increased brain activity during memory tasks as well as longer verbal reaction times ($r = 0.22$, $N = 101$). Hester, Fassbender, and Garavan (2004) reported distinctively different brain activation patterns as measured with fMRI in high versus low CFQ participants, with the former exhibiting greater activation in brain regions related to error processing. It should be noted here, however, that small-sample neuroscience research should encourage a cautious interpretation at the present time, as the field is known for generating spurious significant effects (e.g. Button et al. 2013).

Curiously, the CFQ does not correlate strongly with laboratory memory tests. Rabbitt (1990) found no significant correlations between the CFQ and five laboratory memory tests (effect sizes not reported, $N = 3900$ (– to the best of our knowledge, the largest CFQ sample published to date). The CFQ does not correlate strongly either with IQ tests ($r = -0.04$, $N = 442$, Rabbitt and Abson 1990), vocabulary tests ($r = -0.10$, $N = 442$, Rabbitt and Abson 1990), objective cognitive tests of attention ($r = 0.07$, $N = 96$), psychomotor speed ($r = 0.12$, $N = 108$), information processing speed ($r = 0.10$, $N = 108$), attentional switching ($r = 0.00$, $N = 108$), executive functioning ($r = -0.04$, $N = 108$), visuospatial capacities ($r = -0.13$, $N = 96$), visual memory ($r = 0.05$, $N = 108$), verbal memory ($r = 0.00$, $N = 108$;

$r = -0.11$, $N = 96$), abstract verbal reasoning ($r = 0.01$, $N = 96$) or practical judgement ($r = 0.05$, $N = 96$, Horner, Harvey, and Denier 1999; Van der Werf-Eldering et al. 2011). This, of course, should be, and is, a source of rather significant concern.

The age paradox

Although the CFQ exhibits validity in the sense that it correlates interpretably with other self-reported and some selected, objective criteria, there remains a curious and embedded paradox. It is well known that older people in general perform more poorly on fluid reasoning and perceptual speed tests than younger cohorts, with age-performance correlations ranging between -0.6 and -0.4 (e.g. Salthouse 2015; see also Figure 1). Vigilance, that is, the ability to sustain attention to a task, also decreases sharply with age (e.g. Deaton and Parasuraman 1993). These declines go hand in hand with the shrinking brain volume that occurs as people grow older (Figure 1). Given such physiological premises as those established above, one would anticipate that CFQ scores should increase with age and significantly so.

A substantial number of CFQ studies, however, have actually found a ‘paradoxical negative correlation’ (Reason 1993, 410) between age and the CFQ. This indicates that older people report *fewer* errors via the CFQ than younger people (e.g. $r = -0.18$, $N = 1826$, age range: 16–85, Mecacci and Righi 2006; $r = -0.09$ and $r = -0.13$ for $N = 2100$ and $N = 900$, respectively, age range: 50–96, and $r = -0.12$, $N = 442$, age range: 50–85, Rabbitt and Abson 1990; $r = -0.19$, $N = 312$, age range: 50–96, Rabbitt and Abson 1991; age-CFQ item correlations ranging from 0 to -0.17, $N = 3497$, Rabbitt et al. 1995). Rabbitt et al. (1995) reported CFQ results per item for 475 undergraduate students and 3509 participants between 50 and 86 years old (data from the two samples collected in different studies) and found that in all but four items, the group of older participants reported fewer lapses than the group of younger participants. Also within the group of older participants, the number of self-reported failures decreased with age.

In a study assessing the discrepancy between self-reported CFQ scores and the corresponding scores of an external observer (e.g. a person who knew the participant for at least 2 years and had spent more than 6 h together during the last 2 months), Harty et al. (2013) found that while younger participants (age range = 18–34) over-estimated their cognitive failures relative to the informant’s report (mean CFQ self = 40.8 [SD = 12.5] vs. informant = 36.6 [SD = 16.0], $N = 45$ participant-informant pairs), older participants (age range = 66–90) under-estimated their cognitive failures relative to the informant’s report (mean CFQ self = 41.0 [SD = 11.0] vs. informant = 47.9 [SD = 15.5], $N = 45$ participant-informant pairs). Some other studies have reported no statistically significant correlations between age and total CFQ score

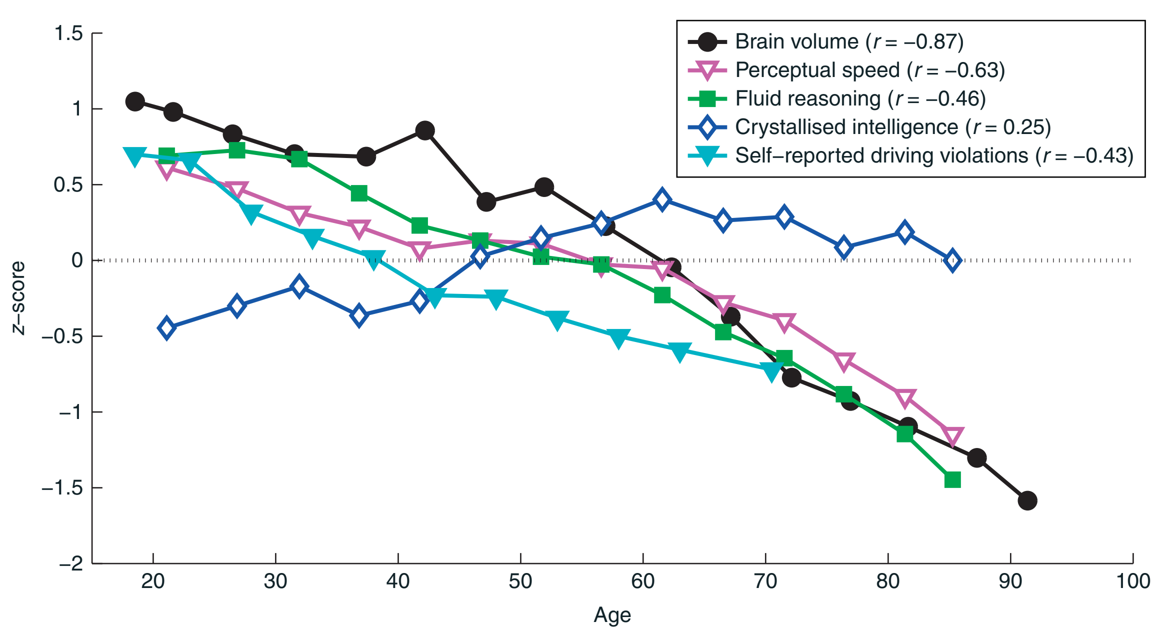

Figure 1. Whole brain volume normalised for head size, self-reported driving violations, and scores in perceptual speed, fluid reasoning and crystallised intelligence, as a function of age. Brain volume data taken from Marcus et al. (2007; raw data available from the Open Access Series of Imaging Studies 2007; $N = 416$, of which 138 were younger than 30 years, 80 were between 30 and 60 years, and 198 were older than 60 years). Perceptual speed, fluid reasoning, and crystallised intelligence data graphically extracted from Salthouse (2009). Driving violations data from Parker et al. (1995). All data are cross-sectional.

(e.g. Borella, Carretti, and De Beni 2008: $r = 0.00$, $N = 304$, age range: 20–86; Bridger, Johnsen, and Brasher 2013: $r = -0.085$, $N = 535$, mean age: 40.9; Weaver Cargin et al. 2008: $r = 0.11$, $N = 100$, mean age: 69.0).

Plausible explanations of the age paradox

Several explanations have been advanced for this paradoxical decrease (and/or non-increase) of CFQ scores with age. Our current list, developed below, elaborates primarily upon the works of Reason (1993) and Rabbitt and Abson (1990), who have each previously highlighted and discussed this paradox. Our overarching aim here is to identify mechanisms, and to gather, examine and synthesise the diverse arguments and critical discussions that might explain the lack of a logically anticipatable positive correlation between age and the CFQ. As might be imagined, several such accounts and explanations have been offered, which we examine prior to offering our own account.

The lenient lifestyle hypothesis

The lenient lifestyle hypothesis (also called 'activity hypothesis' by Reason 1993). Rabbitt and Abson (1990) argued that 'individuals' life-styles and environments may change even faster than their abilities' (14). Elsewhere, Rabbitt and Abson (1991) proposed that the lifestyle of older people is 'less taxing' (143) and characterised by more 'lenient demands' (139), such that older individuals experience a lower (environmental) risk for making cognitive failures, than younger people in the first place – essentially this being a differing baseline effect.

To take an analogy from driving: Older people are disproportionally less frequently involved in accidents simply because they drive many fewer miles than younger persons, that is, a 'mere exposure' effect. But note that the annual crash involvement per driver *per mile* still exhibits a substantial increase from the age of 65 years and beyond (Brar and Rickard 2013). Perhaps ageing is partly characterised by a suppression of novelty or search for comforting regularity, precisely in order that potentially destructive failures and errors are suppressed.

The compensation hypothesis

As proposed by Reason (1993), it may be that people are able to self-regulate their task demands. For example, older people may compensate for their deteriorating cognitive abilities by relying on external mnemonics, such as a notebook to prevent forgetting names (Cheyne, Carriere, and Smilek 2013; Lovelace and Twohig 1990; Moscovitch 1982; Uttl and Kibreab 2011). Additionally, older (i.e. in their 60s and 70s) participants are reportedly more motivated and take tasks more seriously than younger (in their 20s) participants (Aberle et al. 2010; Kliegel et al. 2001; Kvavilashvili and Fisher 2007; Rendell and Craik 2000). Also, older people may invest greater attentional resources in the prevention of trivial errors – what Reason and Lucas (1984) called 'a custodial role for the attentional monitor' (121). If the 'lenient lifestyle' explanation concerns environmental regulation, this 'compensation hypothesis' involves primarily self-regulation.

By analogy: In car driving, older persons drive more slowly than younger people as a way to compensate for their self-assessed physical and cognitive decline (Fuller 2005). In another analogical comparison: It is known that for certain time-cued prospective and habitual memory tasks, such as sending a postcard to the experimenter or calling back at a pre-arranged time, old persons outperform young persons (Uttl, Greene, and Uttl 2008). So, receiving a timely Christmas card from an older person – a prototypical prospective-memory task (Craik and Kerr 1996) – may have been because older persons generally live more structured lives and are motivated to perform well (cf. Aberle et al. 2010; Phillips, Henry, and Martin 2008; Rendell and Craik 2000; Rendell and Thomson 1999). Younger individuals, on the other hand (groups of undergraduate students in many studies), may participate in experiments for course credit (Maylor 1993), thereby exhibiting comparatively low levels of motivation and underperformance (Aberle et al. 2010). Individual differences in motivational state may prove to be a powerful covariate here.

Non-recalled cognitive failures

In the strictest sense, people (both young and old) *cannot* remember all of their own errors because some errors happen unconsciously. Accordingly, a non-zero correlation between CFQ scores and external criteria ought not to be expected: 'Unconscious errors may be hard to remember precisely because they are unconscious; they are not something we want or plan to do' (Bjørnskau and Sagberg 2005, 137). Rabbitt (1990) argued that a declining self-monitoring with age as found in laboratory experiments may, to a certain extent at least, explain the paradox of decreasing self-reported cognitive failures in older age:

> The decline in self-report of lapses may, at least in part, be due to a decline in efficiency of self-monitoring with advancing age. This would explain why 60 and 70 years-olds remember fewer of their lapses than do 50 year-olds (Rabbitt 1990, 1301).

It is also possible that because most of the cognitive failures listed in the CFQ usually have no serious consequences, they may well be considered insignificant and thus go unreported. As noted by Rabbitt et al. (1995), 'lapses of the elderly are less likely to occur in public or to have consequences for others' (S135). Again by analogy: People tend to have a detailed 'flashbulb memory' for events that are of personal importance, are surprising to them or evoke an emotional reaction (Brown and Kulik 1977). Landen and Hendricks (1995) found, for example, that injuries without lost workdays were more underreported (43%) than injuries leading to lost workdays (22.5%). Perhaps similarly, the inconsequential and private 'laughable' (cf. Broadbent et al. 1982) cognitive failures of older people may not be arousing enough to be registered in memory, let alone be reported.

Fluid versus crystallised intelligence

Age exhibits a strong negative correlation with fluid intelligence and speed of information processing (Figure 1). However, age exhibits a moderately *positive* correlation with crystallised intelligence: Older persons are more experienced in many facets of life, leading to greater knowledge (Hedden, Lautenschlager, and Park 2005; Maylor 1994) and perhaps greater wisdom (Grossmann et al. 2010; Staudinger 1999). Tucker-Drob and Salthouse (2008) noted that verbal knowledge, or 'semantic memory', increases over the years (Figure 1; see Park et al. 2002; Schaie and Willis 1993; and various other researchers, for similar results). It is thus possible that the CFQ addresses aspects of crystallised intelligence for which older persons actually perform better than young people do. As noted by Broadbent et al. (1982), the CFQ does not measure resource-limited failures in ability. Rather, it attempts to measure incidents of clumsiness and/or forgetfulness that can be regarded as 'laughably simple' (Broadbent et al. 1982, 1). To pursue again the analogy with car driving, older drivers have much more driving experience than younger drivers, which is a reason why any biological deficits (in terms of speed of processing, motor skills) do not become manifest in accident rates until considerable old age (Peck 1993).

Scale anchors

One particular problem that is intrinsic to semantic labelling derives from the potential for different interpretations of attached labels. Interpretation of scale anchors proves to be both subjective and, more critically, possibly age-related. That is, perceived event frequency as well as the meaning of e.g. 'very often' may well be different for different age groups (Reason 1993) and the very notion of urgency and criticality of a cognitive failure may itself actually change across the years.

Fairly large individual differences have been observed in studies in which participants were asked to assign quantities to the meaning of words such as 'very often', 'quite often', 'occasionally' or 'very rarely' (Bass, Cascio, and O'Connor 1974; Hellrung 2010). In Mutebi (2013), participants ($N = 350$) were asked to match various labels (e.g. 'moderate', 'none', 'often', 'sometimes', 'severe') to the most appropriate anchor on a scale from 0 to 10. Results showed that older participants perceived poor states (e.g. in terms of physical functioning, fatigue, pain, anxiety) as less severe than younger participants did. In another study, this time on drivers' hazard perception, Armsby, Boyle, and Wright (1989) found that older (age range = 50–65) persons tended to use the extremes of the rating scale more often than younger (age range = 23–24) persons did, a phenomenon that, according to those authors, may be 'an indication of decisiveness' (58) in older drivers.

Idiosyncratic reference

In addition to any or all of the above explanations, it is the case that self-assessment of cognitive failures relies on idiosyncratic reference. Schwarz (1999) showed that people generally assume that the anchors in the middle of a scale represent 'average' behaviour, whereas the upper and lower extremes represent the extremes of the distribution. Self-assessment typically is achieved via social comparison rather than in longitudinal terms (except if one is explicitly asked to do so, and see Fozard et al. 1994). Cognitive decline, although substantial across one's lifespan (cf. Figure 1), may be too slow and incremental to be self-perceived (Zelinski, Burnight, and Lane 2001). Schmidt, Berg, and Deelman (1999) found that people between 46 and 89 years old (mean age = 61.6) overestimated their memory ability both with respect to their peers and with respect to 25-year-olds. However, 77% of participants reported a decline in cognitive performance when comparing to themselves at the age of 25.

Kaplan and Baron-Epel (2003) asked 381 individuals what type of reference they use when assessing their health status, and found that 50% of individuals above 60 years old with suboptimal health used their age peers as the reference, whereas only 8% of younger (between 18 and 35 years old) individuals with suboptimal health did so. The authors concluded that 'each individual tries to find ways to evaluate his/her health in a more positive light' (Kaplan and Baron-Epel 2003, 1669).

In the absence of an unequivocally defined 'ground truth' which is used by all respondents, a strong correlation between age and CFQ scores ought not then to be expected.

Beliefs rather than ability

Self-reported cognitive complaints may be seen as pure belief about cognitive functioning rather than as a reflection of actual cognitive functioning per se (Gilewski and Zelinski 1986; Sunderland et al. 1986; Wilhelm, Witthöft, and Schipolowski 2010). Many have reported that memory complaints correlate more strongly with a depressed mood rather than with actual memory performance itself (Bolla et al. 1991; Horner, Harvey, and Denier 1999; Knight et al. 2004). Reason (1993) argued that cognitive failures are not necessarily a function of cognitive decline, but they can also be a function of personal style and local circumstances. Subjective cognitive dysfunction is present when a person perceives shortfalls in their own cognitive functioning and when sufficiently annoying, such failures are expressed as complaints (see De Groot et al. 2000).

However, it might also be reasoned that, either by frustration originating in their cognitive decline or based on the layman's predispositions about ageing, older people would be more likely to complain about their cognitive functioning; again leading to higher rather than lower CFQ scores. In a study using a self-evaluation memory questionnaire (Questionnaire d'auto-évaluation de la mémoire – QAM), for example, Langlois and Belleville (2014) found that the main memory complaints were associated with cognitive failures that were detrimental for autonomy and safety. Broadbent et al. (1982; see also Broadbent 1994) proposed that the CFQ represents a 'vulnerability factor' to stress rather than a measure of cognitive performance per se (and see Hancock and Warm 1989).

Social desirability may be another confounder of self-reported cognitive failures. Fastame and Penna (2012) reported a correlation of -0.34 between the CFQ and the Marlowe-Crowne Social Desirability Scale. Unwillingness of older individuals to admit specific cognitive failures that are associated with age-related cognitive loss may thus also explain a decrease of CFQ scores with age (Rabbitt et al. 1995).

Questionnaires on memory and cognitive failures (including the CFQ) generally consist of negatively formulated statements (e.g. 'How many times have you forgot to . . . '). In an attempt to prevent social desirability, acquiescence bias and other types of response bias, Zimprich, Kliegel, and Rast (2011) included both positive (e.g. 'I almost never forget to . . . .') and negative items in the Prospective and Retrospective Memory Questionnaire. The authors found that older participants ($N = 336$) seemed to embrace more strongly those items with negative ($r = 0.29$) as opposed to positive wording ($r = 0.17$). The same type of age-dependency effect of positive versus negative wording may affect measurement invariance of the CFQ.

Total versus item scores

The correlation coefficient between the total CFQ score and age is either close to zero or marginally negative. However, strong positive correlations may exist between age and particular CFQ items or factors. Rast et al. (2009) studied age effects on CFQ factors in 1303 individuals between 24 and 83 years old (mean age = 51.2) and found that self-reported forgetfulness increases with age (Cohen's effect size increasing from $d = 0.28$ for Group 24–33 years vs. Group 34–43 years to $d = 0.84$ for Group 24–33 years vs. Group 74–93 years). In contrast, self-reported distractibility proved relative stable until the age of 60 (and decreased after that age). Similarly, Schmalzried (2012) found that older (mean age = 74.2) people reported more problems with memory and name items, whereas younger (mean age = 20.7) people reported more blunders. Cutler and Grams (1988) also observed an increase in self-reported memory complaints with age. Using a large sample ($N = 14783$), these authors found that 9.7% of the participants between 55 and 59 years old versus 23.6% of the participants between 80 and 84 years old reported having frequently trouble in remembering things. Similarly, Begum et al. (2014) found that the number of self-reported forgetting incidences over the past period increased from 6.3% for participants between 16 and 24 years old to 15.6% for participants older than 75 years ($N = 7403$). A CFQ study by Maylor (unpublished data reported by Rabbitt et al. 1995; see also Maylor 1990) found that older ($N = 3509$, age range = 50–86) people indicated more self-reported tip-of-the-tongue lapses as compared to a sample of 475 students. Reese and Cherry (2006; $N = 96$) analysed four CFQ items separately ('Do you find you forget appointments?', 'Do you leave important letters unanswered for days?', 'Do you find you forget why you went from one part of the house to the other?', 'Do you find you forget what you came to the shops to buy?') and found that young persons' (mean age = 19.6) ratings were higher than those of old persons (mean age = 69.7) for forgetting appointments (mean score = 1.37 vs. 0.73), whereas the opposite trend was noticed for the item about forgetting why one went from one part of the house to the other (mean score = 1.63 vs. 2.02). These age by item interactions imply that the overall pattern of effects might be much more complex than the single overarching score suggests.

Using a similar cognitive failures questionnaire, the Short Inventory of Minor Lapses (SIML), in two samples ($N = 543$, age range 18–93 years and $N = 1656$, age range 17–69 years), Reason (1993) showed that while self-reported cognitive failures generally declined with age, self-reported difficulty in recalling names significantly increased with age. In other words, the decrease (or non-increase) of the CFQ with age may be due to the fact that the CFQ includes multiple cognitive dimensions, and the self-reported increase with age in some of them (e.g. memory failures) may be confounded, confused or masked by the self-reported decrease in other items (e.g. blunders or distraction). As already pointed out by Herrmann (1982): 'The CFQ is not strictly a memory questionnaire. Besides memory, it covers perceptual errors and stress reactions' (437).

The fact that aggregating (e.g. calculating a total score) implies loss of information may be best captured by the following classical joke:

> Three statisticians go hunting and flush a duck. The first shoots high, over the bird's head. The second aims too low and sends a bullet whistling meters below the duck's belly. So the third statistician jumps up and down yelling, "We got him! We done got him!" (Macilwain 2014, 1221).

The 'average man fallacy', which can be traced back to Quetelet (1849, 96, as described in Mulaik 1985), states that a statistical artefact such as a 'total score' should not be regarded as an entity in the real world. Analogously, the fact that the CFQ total scores do not appreciably correlate with biological age could be the consequence of numerous (strong) idiosyncratic effects that cancel out each other, resulting in an overall null effect.

General methodological problems

In addition to the nominal explanations we have already considered, there are some general concerns that may also serve to explain the identified paradox. Thus:

(1) Many of the studies we cited have been conducted with *small sample sizes*. For age-CFQ correlations, we found only a limited number of studies (Borella, Carretti, and De Beni 2008; Bridger, Johnsen, and Brasher 2013; Mecacci and Righi 2006; Rabbitt et al. 1995; Rabbitt and Abson 1990; Rabbitt and Abson 1991; Weaver Cargin et al. 2008) with more than 100 participants. It is well known that small sample sizes undermine credibility through both false positive and false negative claims (e.g. Button et al. 2013). Thus, selection bias can also potentially serve to influence the issue to hand.

(2) Due to *age-range restriction*, correlations are attenuated (Hunter, Schmidt, and Le 2006). That is, several studies have compared a group of students (i.e. individuals in their 20s) with a group of older individuals (i.e. individuals in their 60s or 70s; e.g. Brehmer, Westerberg, and Bäckman 2012; Fastame and Penna 2012; Hester, Fassbender, and Garavan 2004; Kane et al. 1994; Rabbitt et al. 1995; Reese and Cherry 2006; Simon and Corbett 1996). In such comparisons, correlations necessarily cannot be expected to be strong.
An illustration of the phenomenon of range restriction is provided in Figure 2. Suppose that a population of individuals, with age uniformly distributed between 20 and 100, exhibit a correlation of -0.60 between age and a cognitive performance score, see Figure 2 (top). Now assume that people above a certain age are not included in the study. As Figure 2 (bottom) shows, the observed correlation declines considerably when the age range shrinks. For example, discarding all participants older than 60 years old restricts the correlation to a value of only -0.35.

(3) *Measurement error* is an issue in any form of behavioural assessment (Schmidt and Hunter 1999). The CFQ, frequently being a measurement taken at only one particular moment of time, will never be perfectly reliable, especially when analyses are carried out at the item level (Rushton, Brainerd, and Pressley 1983). As noted by Schmidt and Hunter (1999), the human nervous system is inherently 'noisy'. Responses of the same individual to similar questions may differ substantially because of external and internal disturbances, such as misreading/misinterpreting a word/sentence, a momentary distraction, imprecise decision making about what to report, a writing/typing error, or otherwise. The inter-item correlations of self-report questionnaires are usually not greater than 0.2 or 0.3. But even the CFQ *total* score has imperfect reliability, with test-retest correlations between $r = 0.49$ and 0.84, as indicated above. Because reliability sets an upper limit to validity, no strong correlations between the CFQ and age should be expected.
One way to reduce random response error is to average responses not only across multiple items but also across multiple occasions. For example, in a study that developed multivariate models based on aggregated data from an entire psychometric test battery and diaries kept across multiple days to assess everyday cognitive failures, Unsworth, Brewer, and Spillers (2012) found an average correlation of -0.26 ($N = 100$) between the total number of everyday cognitive failures as reported in the diaries and the latent cognitive ability factors. This suggests that fairly strong correlations between subjective cognitive failures and objective tests of cognitive performance are possible, if (and, perhaps, only if) collecting and aggregating a large amount of data.

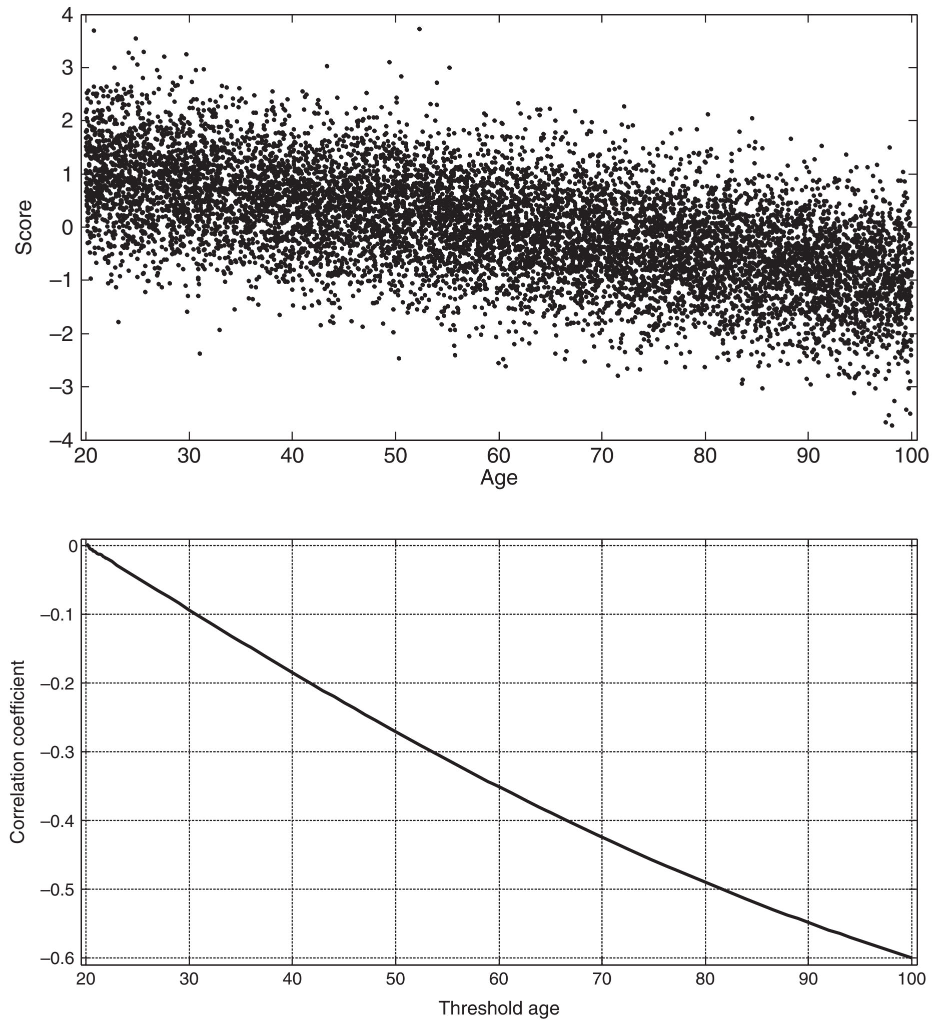

Figure 2. Top: Scatter plot illustrating a Pearson correlation coefficient of -0.60 between age and a performance score for a sample of 10000 people, with age uniformly distributed between 18 and 90. Bottom: Correlation coefficient when excluding participants older than a threshold age. Simulations based on a population of $N = 10000000$).

(4) The majority of studies that have investigated age effects on CFQ scores are *cross-sectional rather than longitudinal*. In a study measuring longitudinal changes in cognitive complaints (using the CFQ and twelve neuropsychological tests), Hohman et al. (2011) found that higher CFQ scores were associated with a steeper decline in verbal memory performance. Cross-sectional research into age-related cognitive decline is methodologically challenging because people have grown up in different environments (cf. the Flynn effect of secular rises in Intelligence Quotient). Longitudinal trends in cognitive functioning may be also obscured because performance on a second testing occasion is positively affected by experience gained during the first testing occasion. Cross-sectional studies may therefore approximate age trends in cognitive functioning better than longitudinal studies (Salthouse 2014). It is possible that longitudinal studies based on self-report questionnaires are susceptible to similar practice effects.

(5) Recruitment of participants and creating *representative samples* is a challenge for much of the gerontology and cognitive ageing research communities (Buckwalter 2009; Christensen et al. 1992; Mody et al. 2008). Older participants in CFQ research may not be representative of the general population as they tend to have an above-average health status. The younger reference group often consists of students (e.g. Mecacci and Righi 2006; Rabbitt et al. 1995; Schmalzried 2012), that is, individuals with above-average cognitive abilities, but possibly not highly motivated, as explained earlier. Non-random sampling introduces potential confounds and causes range restriction, also possibly undermining the validity of the age-CFQ correlations (but see Zelinski, Burnight, and Lane 2001).

Summary and conclusions

The CFQ is a self-report tool that captures peoples' beliefs of their own cognitive functioning. Here, we have focused particularly on the puzzling phenomenon that CFQ scores do not increase (or even show a slight decline) with age; something that runs clearly counter to well-established age-related cognitive declines as captured through psychometric testing. However, the CFQ does not measure 'just noise' because fairly strong test-retest correlations have been established by former investigators. Also, it is untrue that self-reports are inherently invalid (see Hoffman and Hancock 2014). Some types of self-reported cognitive failures, such as traffic violations, exhibit clear associations with biological age (see Figure 1, for self-violations measured with the Driver Behaviour Questionnaire; DBQ).

What can we make of these collective patterns of data and their associated speculative 'explanations'? Clearly, it is possible that the cognitive failures truly do not co-vary with age because people compensate for their biological deterioration through increasing experience, knowledge and wisdom (Explanation 4); also they may adapt their lifestyle (Explanation 1), or find other ways to compensate for their cognitive decline (Explanation 2). Apparently, by virtue of these mechanisms, people are able to function effectively despite their monotonically shrinking brain (Figure 1). It is also possible that the CFQ actually measures something else than resource-limited cognitive ability (Explanation 4), and that clumsy failures do not become manifest easily, except maybe on such tasks as recalling names and facts, where people may (embarrassingly) be confronted with their own limitations (Explanation 8).

Another, potentially even more illuminating possibility is that people are simply unable to rate themselves with respect to some unstated reference value. In trying to do so, people seek to find a norm in their own idiosyncratic environment (Explanation 6) and simply report what they believe about themselves (Explanation 7) or what they believe that the scale anchors ought to represent (Explanation 5). This would certainly be an interesting possibility and suggests that the dissociation between objective and subjective performance could well be a rich source of information (cf. Hancock 1996; Yeh and Wickens 1988) for which recent empirical observations are beginning to coalesce (and see Reinerman et al. 2014). For example, when a person claims to make only few cognitive failures while psychometric tests clearly demonstrate otherwise, this could be indicative of delusional thought.

Yet, at the heart of the age-CFQ paradox may be the idea that people cannot veridically reflect on their own failures. If the frequency of forgetting what one has forgotten increases over age then the inherent limitation of the individual is inevitably commingled in the method itself. This is an apparently insurmountable flaw in the reliance on the subjective reaction itself. One potential avenue for improvement would seem lies in the latency question. That is, forgetting itself is a dynamic process that varies across time. Perhaps if CFQs were administered on a moment-by-moment basis then more accurate reporting would be achieved (cf. Lange and Süß 2014, reporting on the ambulatory use of the Questionnaire for Cognitive Failures in Everyday Life via a mobile phone). However, we have here a case of *what* is being tested being a critical arbiter of *how* the test is conducted. Conflating and conflicting those two inherent functions can lead to some of the concerns we have pointed to. An important question remaining is whether such contamination affects all forms of human subjective report.

The principle behind 'forgetting that one forgets' can be further illustrated with some basic algebra. Suppose a person completes a number (N) of cognitive tasks and fails in a certain fraction (a, between 0 and 1) of those. The total number of failed tasks is therefore Na, and the total number of successfully completed tasks is its complement, $N - Na$, or $N(1 - a)$. Now suppose that one forgets a fraction (b) of the failed cognitive tasks. The number of remembered cognitive failures is then $Na(1 - b)$. If this effect holds true, then the number of self-reported cognitive failures represents a veridical linear function of one's actual performance (Figure 3). However, it is possible that the probability of remembering equals the probability of task success. That is to say, people who (due to e.g. cognitive decline) make many cognitive failures are also likely to forget these failures. The number of forgotten cognitive failures is then Na^2, and the number of remembered cognitive failures equals $Na(1 - a)$, or $Na - Na^2$. This means that people who have an intermediate cognitive performance (i.e. $a = 0.5$) report the highest number of cognitive failures because they make a fair amount of cognitive failures and also remember a fair share of these failures. People who make lots of failures appear to themselves to have a good memory: Although they make lots of failures, they forget these failures (Figure 4). This is an extreme position, but it is plausible, considering that all cognitive processes are strongly correlated with each other (e.g. Jensen and Weng 1994). In reality, considering an imperfect, but strongly positive, correlation between cognitive failures and forgetting (which is in essence a cognitive failure too), the CFQ-age relationship may lie in-between the dashed and dotted line in Figure 3. However, even when the assumptions of our model are loosened this way, it could still explain why self-reported cognitive failures decrease with age.

Clearly, the age-related cognitive decline, particularly when not acknowledged by the ageing individual himself, has implications for HF/E and the design of products (Rabbitt 1992). Without the subjective dimension of the human operator's participation, HF/E devolves to pure engineering and we re-embrace the fallacies inherent in radical behaviourism.

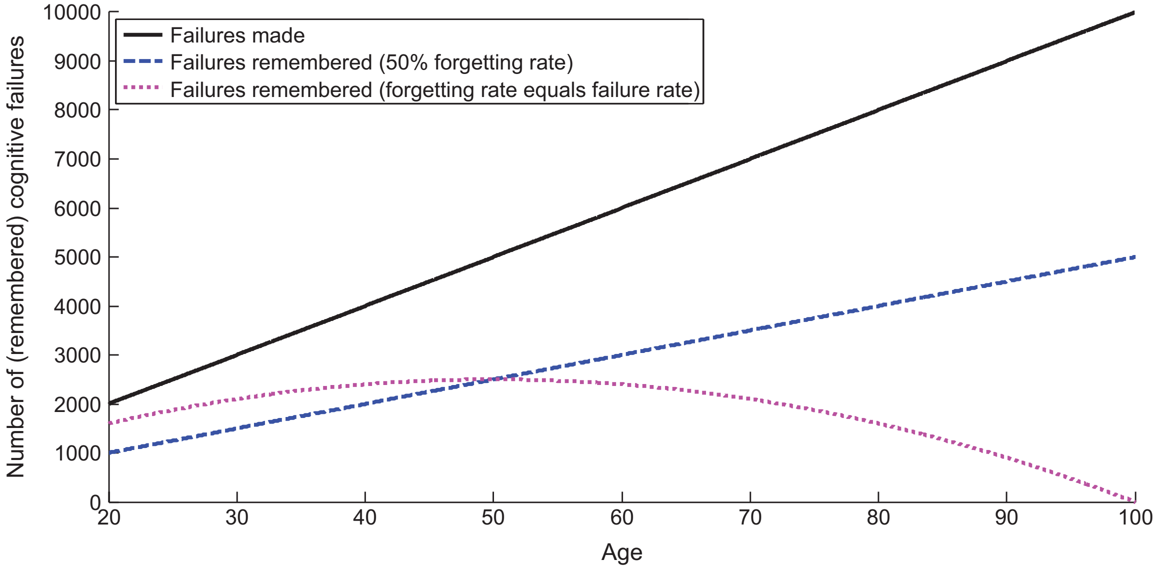

Figure 3. Predicted number of cognitive failures made (solid black line; the failure rate equals age/100) and remembered (dashed blue and dotted magenta lines) as a function of age. The dashed blue line illustrates the case in which the forgetting rate is 50% for all ages, whereas the magenta line corresponds to the case of increasing forgetting rate with age (the forgetting rate equals the failure rate). In this analysis, a total of 10000 tasks were performed.

Nevertheless, we must continue to ensure that the methods we use to elicit that conscious participation of the human element in the overall work system, are themselves not inherently flawed. This is an on-going challenge that we must embrace and conquer if HF/E is to have the impact in the world that we aspire for it.

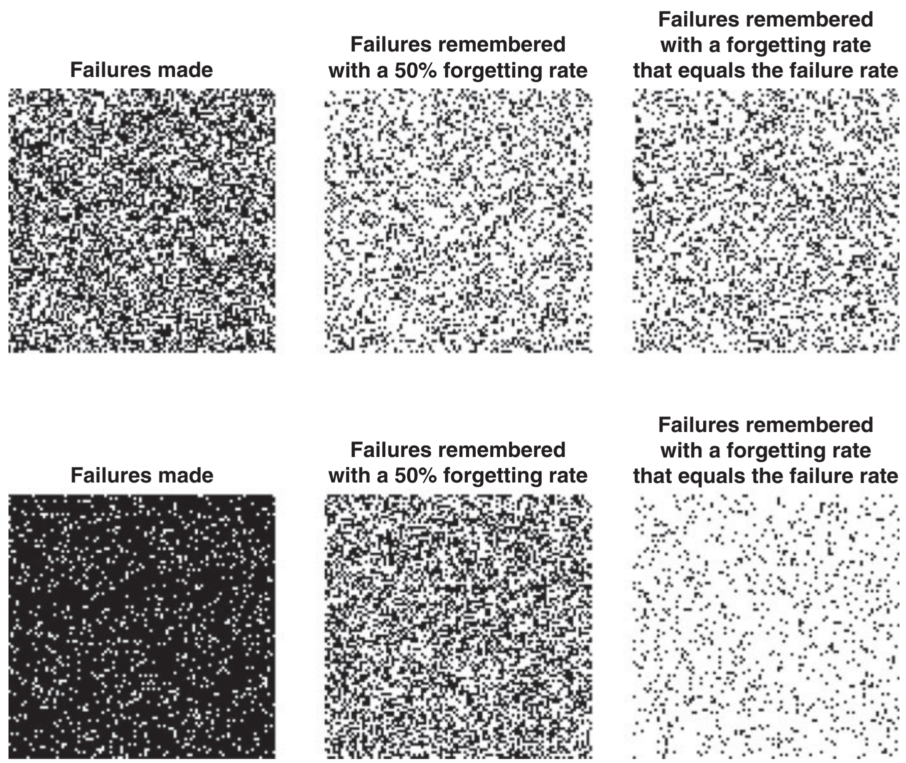

Figure 4. Illustration of cognitive failures made (black pixels in first column) and cognitive failures remembered when a person forgets 50% of his failures ($b = 0.5$; black pixels in second column) and when a person's forgetting rate equals the failure rate (black pixels in third column). The top row corresponds to a person who fails on 50% of the occasions ($a = 0.5$). The bottom row corresponds to a person who fails on 90% of the occasions ($a = 0.9$). In all six figures, the cognitive performance is depicted as a 100 × 100 matrix (=10000) of tasks.

Acknowledgements

We thank OASIS for the cross-sectional data used in this article. OASIS is supported by grant numbers: P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, R01 MH56584.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Aberle, I., P. G. Rendell, N. S. Rose, M. A. McDaniel, and M. Kliegel. 2010. "The Age Prospective Memory Paradox: Young Adults May Not Give Their Best Outside of the Lab." *Developmental Psychology* 46 (6): 1444–1453. doi:10.1037/a0020718.

Ackerman, P. L., M. E. Beier, and K. R. Bowen. 2002. "What We Really Know About Our Abilities and Our Knowledge." *Personality and Individual Differences* 33 (4): 587–605. doi:10.1016/S0191-8869(01)00174-X.

Ackerman, P. L., R. Kanfer, and M. Goff. 1995. "Cognitive and Noncognitive Determinants and Consequences of Complex Skill Acquisition." *Journal of Experimental Psychology: Applied* 1 (4): 270–304. doi:10.1037/1076-898X.1.4.270.

Armsby, P., A. J. Boyle, and C. C. Wright. 1989. "Methods for Assessing Drivers' Perception of Specific Hazards on the Road." *Accident Analysis & Prevention* 21 (1): 45–60. doi:10.1016/0001-4575(89)90048-1.

Bass, B. M., W. F. Cascio, and E. J. O'Connor. 1974. "Magnitude Estimations of Expressions of Frequency and Amount." *Journal of Applied Psychology* 59 (3): 313–320. doi:10.1037/h0036653.

Beaudoin, M., and O. Desrichard. 2011. "Are Memory Self-Efficacy and Memory Performance Related? A Meta-Analysis." *Psychological Bulletin* 137 (2): 211–241. doi:10.1037/a0022106.

Begum, A., M. Dewey, A. Hassiotis, M. Prince, S. Wessely, and R. Stewart. 2014. "Subjective Cognitive Complaints Across the Adult Life Span: A 14-Year Analysis of Trends and Associations Using the 1993, 2000 and 2007 English Psychiatric Morbidity Surveys." *Psychological Medicine* 44 (9): 1977–1987. doi:10.1017/S0033291713002444.

Bennett-Levy, J., and G. E. Powell. 1980. "The Subjective Memory Questionnaire (SMQ). An Investigation into the Self-Reporting of 'Real-Life' Memory Skills." *British Journal of Social and Clinical Psychology* 19 (2): 177–188. doi:10.1111/j.2044-8260.1980.tb00946.x.

Bjørnskau, T., and F. Sagberg. 2005. "What Do Novice Drivers Learn During the First Months of Driving? Improved Handling Skills or Improved Road User Interaction?" In *Traffic and Transport Psychology, Theory and Application*, edited by G. Underwood, 129–140. London: Elsevier.

Bolla, K. I., K. N. Lindgren, C. Bonaccorsy, and M. L. Bleecker. 1991. "Memory Complaints in Older Adults: Fact or Fiction?." *Archives of Neurology* 48 (1): 61–64. doi:10.1001/archneur.1991.00530130069022.

Borella, E., B. Carretti, and R. De Beni. 2008. "Working Memory and Inhibition Across the Adult Life-Span." *Acta Psychologica* 128 (1): 33–44. doi:10.1016/j.actpsy.2007.09.008.

Brar, S. S., and D. P. Rickard. 2013. *Teen and Senior Drivers* (Final Report No. RSS-13-240). Sacramento, CA: California Department of Motor Vehicles.

Brehmer, Y., H. Westerberg, and L. Bäckman. 2012. "Working-Memory Training in Younger and Older Adults: Training Gains, Transfer, and Maintenance." *Frontiers in Human Neuroscience* 6: 63. doi:10.3389/fnhum.2012.00063.

Bridger, R. S., S. Å. K. Johnsen, and K. Brasher. 2013. "Psychometric Properties of the Cognitive Failures Questionnaire." *Ergonomics* 56 (10): 1515–1524. doi:10.1080/00140139.2013.821172.

Broadbent, D. E. 1994. "Individual Differences in Response to Low Discretion Jobs." *Ergonomics* 37 (4): 641–653. doi:10.1080/00140139408963679.

Broadbent, D. E., P. F. Cooper, P. FitzGerald, and K. R. Parkes. 1982. "The Cognitive Failures Questionnaire (CFQ) and its Correlates." *British Journal of Clinical Psychology* 21 (1): 1–16. doi:10.1111/j.2044-8260.1982.tb01421.x.

Brown, R., and J. Kulik. 1977. "Flashbulb Memories." *Cognition* 5 (1): 73–99. doi:10.1016/0010-0277(77)90018-X.

Buckwalter, K. C. 2009. "Recruitment of Older Adults." *Research in Gerontological Nursing* 2 (4): 265–266. doi:10.3928/19404921-20090816-01.

Button, K. S., J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, and M. R. Munafò. 2013. "Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience." *Nature Reviews Neuroscience* 14 (5): 365–376. doi:10.1038/nrn3475.

Cheyne, J. A., J. S. Carriere, and D. Smilek. 2013. "Age Differences in Attention Lapses Mask Age Differences in Memory Failures: A Methodological Note on Suppression." *Frontiers in Psychology* 4: 99. doi:10.3389/fpsyg.2013.00099.

Christensen, K. J., J. Moye, R. R. Armson, and T. M. Kern. 1992. "Health Screening and Random Recruitment for Cognitive Aging Research." *Psychology and Aging* 7 (2): 204–208. doi:10.1037/0882-7974.7.2.204.

Craik, F. I., and S. A. Kerr. 1996. "Prospective Memory, Aging, and Lapses of Intention." In *Prospective memory: Theory and applications*, edited by M. Brandimonte, G. O. Einstein, and M. A. McDaniel, 227–237. Mahwah, NJ: Erlbaum.

Crooks, V. C., J. G. Buckwalter, D. B. Petitti, K. K. Brody, and R. L. Yep. 2005. "Self-Reported Severe Memory Problems as a Screen for Cognitive Impairment and Dementia." *Dementia* 4 (4): 539–551. doi:10.1177/1471301205058310.

Cutler, S. J., and A. E. Grams. 1988. "Correlates of Self-Reported Everyday Memory Problems." *Journal of Gerontology* 43 (3): S82–S90. doi:10.1093/geronj/43.3.S82.

De Groot, J. C., M. Oudkerk, J. V. Gijn, A. Hofman, J. Jolles, and M. M. Breteler. 2000. "Cerebral White Matter Lesions and Cognitive Function: The Rotterdam Scan Study." *Annals of Neurology* 47 (2): 145–151. doi:10.1002/1531-8249(200002)47:2<145:AID-ANA3>3.0.CO;2-P.

De Winter, J. C. F., A. Zadpoor, and D. Dodou. 2014. "The Expansion of Google Scholar Versus Web of Science: A longitudinal Study." *Scientometrics* 98 (2): 1547–1565. doi:10.1007/s11192-013-1089-2.

Deaton, J. E., and R. Parasuraman. 1993. "Sensory and Cognitive Vigilance: Effects of Age on Performance and Subjective Workload." *Human Performance* 6 (1): 71–97. doi:10.1207/s15327043hup0601_4.

Dixon, R. A., D. F. Hultsch, and C. Hertzog. 1987. "The Metamemory in Adulthood (MIA) Questionnaire." *Psychopharmacology Bulletin* 24: 671–688.

Dul, J., R. Bruder, P. Buckle, P. Carayon, P. Falzon, W. S. Marras, J. R. Wilson, and B. van der Doelen. 2012. "A Strategy for Human Factors/Ergonomics: Developing the Discipline and Profession." *Ergonomics* 55 (4): 377–395. doi:10.1080/00140139.2012.661087.

Dunning, D., C. Heath, and J. M. Suls. 2004. "Flawed Self-Assessment: Implications for Health, Education, and the Workplace." *Psychological Science in the Public Interest* 5 (3): 69–106. doi:10.1111/j.1529-1006.2004.00018.x.

Dunning, D., K. Johnson, J. Ehrlinger, and J. Kruger. 2003. "Why People Fail to Recognize Their Own Incompetence." *Current Directions in Psychological Science* 12 (3): 83–87. doi:10.1111/1467-8721.01235.

Fastame, M. C., and M. P. Penna. 2012. "Does Social Desirability Confound the Assessment of Self-Reported Measures of Well-Being and Metacognitive Efficiency in Young and Older Adults?" *Clinical Gerontologist* 35 (3): 239–256. doi:10.1080/07317115.2012.660411.

Fitts, P. M., and R. E. Jones. 1947. *Analysis of Factors Contributing to 460 'Pilot-Error' Experiences in Operating Aircraft Controls* (Report No. TSEAA-694-12). Wright-Patterson Air Force Base, OH: Aero Medical Laboratory.

Flehmig, H. C., M. Steinborn, R. Langner, and K. Westhoff. 2007. "Neuroticism and the Mental Noise Hypothesis: Relationships to Lapses of Attention and Slips of Action in Everyday Life." *Psychology Science* 49: 343–360.

Fozard, J. L., M. Vercruyssen, S. L. Reynolds, P. A. Hancock, and R. E. Quilter. 1994. "Age Differences and Changes in Reaction Time: The Baltimore Longitudinal Study of Aging." *Journal of Gerontology: Psychological Sciences* 49 (4): P179–P189. doi:10.1093/geronj/49.4.P179.

Fuller, R. 2005. "Towards a General Theory of Driver Behaviour." *Accident Analysis & Prevention* 37 (3): 461–472. doi:10.1016/j.aap.2004.11.003.

Garavan, H., R. Hester, K. Murphy, C. Fassbender, and C. Kelly. 2006. "Individual Differences in the Functional Neuroanatomy of Inhibitory Control." *Brain Research* 1105 (1): 130–142. doi:10.1016/j.brainres.2006.03.029.

Gehanno, J. -F., L. Rollin, and S. Darmoni. 2013. "Is the Coverage of Google Scholar Enough to be Used Alone for Systematic Reviews." *BMC Medical Informatics and Decision Making* 13: 7. doi:10.1186/1472-6947-13-7.

Gilewski, M., and E. Zelinski. 1986. "Questionnaire Assessment of Memory Complaints." In *Handbook for Clinical Memory Assessment of Older Adults*, edited by L. W. Ponn, T. Crook, K. L. Davis, C. Eisdorder, B. J. Gurland, A. W. Kaszniak, and L. W. Thompson, 93–107. Washington, DC: American Psychological Association.

Grossmann, I., J. Na, M. Varnum, D. C. Park, S. Kitayama, and R. E. Nisbett. 2010. "Reasoning About Social Conflicts Improves Into Old Age." *Proceedings of the National Academy of Sciences* 107 (16): 7246–7250. doi:10.1073/pnas.1001715107.

Gulian, E., G. Matthews, A. I. Glendon, D. R. Davies, and L. M. Debney. 1989. "Dimensions of Driver Stress." *Ergonomics* 32 (6): 585–602. doi:10.1080/00140138908966134.

Hale, A. R., J. Stoop, and J. Hommels. 1990. "Human Error Models as Predictors of Accident Scenarios for Designers in Road Transport Systems." *Ergonomics* 33 (10–11): 1377–1387. doi:10.1080/00140139008925339.

Hancock, P. A. 1996. "Effects of Control Order, Augmented Feedback, Input Device and Practice on Tracking Performance and Perceived Workload." *Ergonomics* 39 (9): 1146–1162. doi:10.1080/00140139608964535.

Hancock, P. A., and J. S. Warm. 1989. "A Dynamic Model of Stress and Sustained Attention." *Human Factors* 31: 519–537.

Hancock, P. A., J. L. Weaver, and R. Parasuraman. 2002. "Sans Subjectivity – Ergonomics is Engineering." *Ergonomics* 45 (14): 991–994. doi:10.1080/00140130210166898.

Hannon, R., P. Adams, S. Harrington, C. Fries-Dias, and M. T. Gipson. 1995. "Effects of Brain Injury and Age on Prospective Memory Self-Rating and Performance." *Rehabilitation Psychology* 40 (4): 289–298. doi:10.1037/0090-5550.40.4.289.

Harty, S., R. G. O'Connell, R. Hester, and I. H. Robertson. 2013. "Older Adults Have Diminished Awareness of Errors in the Laboratory and Daily Life." *Psychology and Aging* 28 (4): 1032–1041. doi:10.1037/a0033567.

Hedden, T., G. Lautenschlager, and D. C. Park. 2005. "Contributions of Processing Ability and Knowledge to Verbal Memory Tasks Across the Adult Life-Span." *The Quarterly Journal of Experimental Psychology Section A* 58 (1): 169–190. doi:10.1080/02724980443000179.

Hellrung, J. L., II. 2010. *Labels for Anchors: The Good, the Bad, and the Endpoints* (Master thesis). Bowling Green, Kentucky: Western Kentucky University.

Herrmann, D. J. 1982. "Know Thy Memory: The Use of Questionnaires to Assess and Study Memory." *Psychological Bulletin* 92 (2): 434–452. doi:10.1037/0033-2909.92.2.434.

Hester, R., C. Fassbender, and H. Garavan. 2004. "Individual Differences in Error Processing: A Review and Reanalysis of Three Event-Related fMRI Studies Using the GO/NOGO Task." *Cerebral Cortex* 14 (9): 986–994. doi:10.1093/cercor/bhh059.

Hoffman, R. R., and P. A. Hancock. 2014. "Words Matter." *Human Factors and Ergonomics Society Bulletin* 57: 3–7.

Hohman, T. J., L. L. Beason-Held, M. Lamar, and S. M. Resnick. 2011. "Subjective Cognitive Complaints and Longitudinal Changes in Memory and Brain Function." *Neuropsychology* 25 (1): 125–130. doi:10.1037/a0020859.

Hoorens, V., and P. Harris. 1998. "Distortions in Reports of Health Behaviors: The Time Span Effect and Illusory Superiority." *Psychology and Health* 13 (3): 451–466. doi:10.1080/08870449808407303.

Horner, M. D., R. T. Harvey, and C. A. Denier. 1999. "Self-Report and Objective Measures of Cognitive Deficit in Patients Entering Substance Abuse Treatment." *Psychiatry Research* 86 (2): 155–161. doi:10.1016/S0165-1781(99)00031-1.

Hunter, J. E., F. L. Schmidt, and H. Le. 2006. "Implications of Direct and Indirect Range Restriction for Meta-Analysis Methods and Findings." *Journal of Applied Psychology* 91 (3): 594–612. doi:10.1037/0021-9010.91.3.594.

Jensen, A. R., and L. J. Weng. 1994. "What is a Good G?." *Intelligence* 18 (3): 231–258. doi:10.1016/0160-2896(94)90029-9.

Kane, M. J., L. Hasher, E. R. Stoltzfus, R. T. Zacks, and S. L. Connelly. 1994. "Inhibitory Attentional Mechanisms and Aging." *Psychology and Aging* 9 (1): 103–112. doi:10.1037/0882-7974.9.1.103.

Kaplan, G., and O. Baron-Epel. 2003. “What Lies Behind the Subjective Evaluation of Health Status?” *Social Science & Medicine* 56 (8): 1669–1676. doi:10.1016/S0277-9536(02)00179-X.

Kliegel, M., M. Martin, M. A. McDaniel, and G. O. Einstein. 2001. “Varying the Importance of a Prospective Memory Task: Differential Effects Across Time-and Event-Based Prospective Memory.” *Memory* 9 (1): 1–11. doi:10.1080/09658210042000003.

Knight, R. G., J. McMahon, T. J. Green, and C. M. Skeaff. 2004. “Some Normative and Psychometric Data for the Geriatric Depression Scale and the Cognitive Failures Questionnaire from a sample of healthy older persons.” *New Zealand Journal of Psychology* 33: 163–170.

Kruger, J. 1999. “Lake Wobegon be Gone! The Below-Average Effect and the Egocentric Nature of Comparative Ability Judgments.” *Journal of Personality and Social Psychology* 77 (2): 221–232. doi:10.1037/0022-3514.77.2.221.

Kruger, J., and D. Dunning. 1999. “Unskilled and Unaware Of It: How Difficulties in Recognizing One’s Own Incompetence Lead to Inflated Self-Assessments.” *Journal of Personality and Social Psychology* 77 (6): 1121–1134. doi:10.1037/0022-3514.77.6.1121.

Kvavilashvili, L., and L. Fisher. 2007. “Is Time-Based Prospective Remembering Mediated by Self-Initiated Rehearsals? Role of Incidental Cues, Ongoing Activity, Age, and Motivation.” *Journal of Experimental Psychology: General* 136 (1): 112–132. doi:10.1037/0096-3445.136.1.112.

Landen, D. D., and S. Hendricks. 1995. “Effect of Recall on Reporting of At-Work Injuries.” *Public Health Reports* 110: 350–354.

Lange, S., and H. M. Süß. 2014. “Measuring Slips and Lapses When They Occur – Ambulatory Assessment in Application to Cognitive Failures.” *Consciousness and Cognition* 24: 1–11. doi:10.1016/j.concog.2013.12.008.

Langlois, A. S., and S. Belleville. 2014. “Subjective Cognitive Complaint in Healthy Older Adults: Identification of Major Domains and Relation to Objective Performance.” *Aging, Neuropsychology, and Cognition* 21 (3): 257–282. doi:10.1080/13825585.2013.795928.

Larson, G. E., D. L. Alderton, M. Neideffer, and E. Underhill. 1997. “Further Evidence on Dimensionality and Correlates of the Cognitive Failures Questionnaire.” *British Journal of Psychology* 88 (1): 29–38. doi:10.1111/j.2044-8295.1997.tb02618.x.

Larson, G. E., and C. R. Merritt. 1991. “Can Accidents be Predicted? An Empirical Test of the Cognitive Failures Questionnaire.” *Applied Psychology* 40 (1): 37–45. doi:10.1111/j.1464-0597.1991.tb01356.x.

Lesch, M. F., and P. A. Hancock. 2004. “Driving Performance During Concurrent Cell-Phone Use: Are Drivers Aware of Their Performance Decrements?” *Accident Analysis & Prevention* 36 (3): 471–480. doi:10.1016/S0001-4575(03)00042-3.

Lovelace, E. A., and P. T. Twohig. 1990. “Healthy Older Adults’ Perceptions of Their Memory Functioning and Use of Mnemonics.” *Bulletin of the Psychonomic Society* 28 (2): 115–118. doi:10.3758/BF03333979.

Macilwain, C. 2014. “A Touch of the Random.” *Science* 344 (6189): 1221–1223. doi:10.1126/science.344.6189.1221.

Marcus, D. S., T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner. 2007. “Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults.” *Journal of Cognitive Neuroscience* 19 (9): 1498–1507. doi:10.1162/jocn.2007.19.9.1498.

Martin, M. 1983. “Cognitive Failure: Everyday and Laboratory Performance.” *Bulletin of the Psychonomic Society* 21 (2): 97–100. doi:10.3758/BF03329964.

Maylor, E. A. 1990. “Recognizing and Naming Faces: Aging, Memory Retrieval, and the Tip of the Tongue State.” *Journal of Gerontology* 45 (6): P215–P226. doi:10.1093/geronj/45.6.P215.

Maylor, E. A. 1993. “Minimized Prospective Memory Loss in Old Age.” In *Adult Information Processing: Limits on Loss*, edited by J. Cerella, J. Rybash, W. Hoyer, and M. L. Commons, 529–551. San Diego, CA: Academic Press.

Maylor, E. A. 1994. “Ageing and the Retrieval of Specialized and General Knowledge: Performance of Masterminds.” *British Journal of Psychology* 85 (1): 105–114. doi:10.1111/j.2044-8295.1994.tb02511.x.

Mecacci, L., and S. Righi. 2006. “Cognitive Failures, Metacognitive Beliefs and Aging.” *Personality and Individual Differences* 40 (7): 1453–1459. doi:10.1016/j.paid.2005.11.022.

Merckelbach, H., P. Muris, H. Nijman, and P. J. De Jong. 1996. “Self-Reported Cognitive Failures and Neurotic Symptomatology.” *Personality and Individual Differences* 20 (6): 715–724. doi:10.1016/0191-8869(96)00024-4.

Mody, L., D. K. Miller, J. M. McGloin, M. Freeman, E. R. Marcantonio, J. Magaziner, and S. Studenski. 2008. “Recruitment and Retention of Older Adults in Aging Research.” *Journal of the American Geriatrics Society* 56 (12): 2340–2348. doi:10.1111/j.1532-5415.2008.02015.x.

Moscovitch, M. 1982. “A Neuropsychological Approach to Memory and Perception in Normal and Pathological Aging.” In *Advances in the Study of communication and Affect: Vol. 8. Aging and Cognitive Processes*, edited by F. I. M. Craik and S. Trehub, 55–78. New York: Plenum Press.

Mulaik, S. A. 1985. “Exploratory Statistics and Empiricism.” *Philosophy of Science* 52 (3): 410–430. doi:10.1086/289258.

Mutebi, A. 2013. *Comparing Response Scaling Formats Used in Patient-Reported Outcome (PRO) Instruments* (Doctoral dissertation) The University of Arizona.

Open Access Series of Imaging Studies. 2007. *OASIS: Cross-Sectional MRI Data in Young, Middle Aged, Nondemented and Demented Older Adults*, [Data file]. Retrieved from http://www.oasis-brain.org

Park, D. C., G. Lautenschlager, T. Hedden, N. S. Davidson, A. D. Smith, and P. K. Smith. 2002. “Models of Visuospatial and Verbal Memory Across the Adult Life Span.” *Psychology and Aging* 17 (2): 299–320. doi:10.1037/0882-7974.17.2.299.

Parker, D., J. T. Reason, A. S. Manstead, and S. G. Stradling. 1995. “Driving Errors, Driving Violations and Accident Involvement.” *Ergonomics* 38 (5): 1036–1048. doi:10.1080/00140139508925170.

Peck, R. C. 1993. “The Identification of Multiple Accident Correlates in High Risk Drivers With Specific Emphasis on the Role of Age, Experience and Prior Traffic Violation Frequency.” *Alcohol, Drugs & Driving* 9: 145–166.

Pfeifer, S., J. Van Os, M. Hanssen, P. Delespaul, and L. Krabbendam. 2009. “Subjective Experience of Cognitive Failures as Possible Risk Factor for Negative Symptoms of Psychosis in the General Population.” *Schizophrenia Bulletin* 35 (4): 766–774. doi:10.1093/schbul/sbn004.

Phillips, L. H., J. D. Henry, and M. Martin. 2008. "Adult Aging and Prospective Memory: The Importance of Ecological Validity." In *Prospective Memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives*, edited by M. Kliegel, M. A. McDaniel, and G. O. Einstein, 161–185. Mahwah, NJ: Erlbaum.

Podsakoff, P. M., S. B. MacKenzie, J. Y. Lee, and N. P. Podsakoff. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies." *Journal of Applied Psychology* 88 (5): 879–903. doi:10.1037/0021-9010.88.5.879.

Portet, F., P. J. Ousset, P. J. Visser, G. B. Frisoni, F. Nobili, P. Scheltens, ..., and J. Touchon. 2006. "Mild Cognitive Impairment (MCI) in Medical Practice: A Critical Review of the Concept and New Diagnostic Procedure. Report of the MCI Working Group of the European Consortium on Alzheimer's Disease." *Journal of Neurology, Neurosurgery & Psychiatry* 77 (6): 714–718. doi:10.1136/jnnp.2005.085332.

Quetelet, M. A. 1849. *Letters Addressed to H.R.H. the Grand Duke of Saxe Coburg and Gotha, on the Theory of Probability as Applied to the Moral and Political Sciences.* (O.G. Downes, trans.). London: Charles and Edwin Layton. (Originally written in 1837 but published as *Lettres ... sur la théorie des probabilités*, Brussels, 1846).

Rabbitt, P. 1990. "Age, IQ and Awareness, and Recall of Errors." *Ergonomics* 33 (10–11): 1291–1305. doi:10.1080/00140139008925333.

Rabbitt, P. M. A. 1992. "Cognitive Changes with Age Must Influence Human Factors Design." In *Gerontechnology*, edited by H. Bouma and J. A. M. Graafmans, 113–140. Amsterdam: IOS Press.

Rabbitt, P., and V. Abson. 1990. "'Lost and Found': Some Logical and Methodological Limitations of Self-Report Questionnaires as Tools to Study Cognitive Ageing." *British Journal of Psychology* 81 (1): 1–16. doi:10.1111/j.2044-8295.1990.tb02342.x.

Rabbitt, P., and V. Abson. 1991. "Do Older People Know How Good They Are?" *British Journal of Psychology* 82 (2): 137–151. doi:10.1111/j.2044-8295.1991.tb02389.x.

Rabbitt, P., E. Maylor, L. McInnes, N. Bent, and B. Moore. 1995. "What Goods Can Self-Assessment Questionnaires Deliver for Cognitive Gerontology?" *Applied Cognitive Psychology* 9 (7): S127–S152. doi:10.1002/acp.2350090709.

Rast, P., D. Zimprich, M. Van Boxtel, and J. Jolles. 2009. "Factor Structure and Measurement Invariance of the Cognitive Failures Questionnaire Across the Adult Life Span." *Assessment* 16 (2): 145–158. doi:10.1177/1073191108324440.

Reason, J. 1984. "Lapses of Attention in Everyday Life." In *Varieties of Attention*, edited by R. Parasuraman, R. Davies, and J. B. Beatty, 515–549. New York: Academic Press.

Reason, J. 1993. "Self-Report Questionnaires in Cognitive Psychology: Have They Delivered the Goods?" In *Attention: Selection, Awareness and Control: A Tribute to Donald Broadbent*, edited by A. Baddeley and L. Weiskrantz, 406–423. Oxford: Clarendon Press.

Reason, J. T., and D. Lucas. 1984. "Absent-Mindedness in Shops: its Incidence, Correlates and Consequences." *British Journal of Clinical Psychology* 23 (2): 121–131. doi:10.1111/j.2044-8260.1984.tb00635.x.

Reason, J. T., and K. Mycielska. 1982. *Absent-Minded? The Psychology of Mental Lapses and Everyday Errors*. Englewood Cliffs, NJ: Prentice-Hall.

Reese, C. M., and K. E. Cherry. 2006. "Effects of Age and Ability on Self-Reported Memory Functioning and Knowledge of Memory Aging." *The Journal of Genetic Psychology* 167 (2): 221–240. doi:10.3200/GNTP.167.2.221-240.

Reinerman, L., J. Mercado, R. Leis, J. L. Szalma, and P. A. Hancock. 2014. *Assessing Operator Response on Nuclear Power Plant Control Tasks Through Objective Measures, Subjective Report, and Physiological Reactions*. Manuscript submitted for publication.

Rendell, P. G., and F. I. M. Craik. 2000. "Virtual Week and Actual Week: Age-Related Differences in Prospective Memory." *Applied Cognitive Psychology* 14 (7): S43–S62. doi:10.1002/acp.770.

Rendell, P. G., and D. M. Thomson. 1999. "Aging and Prospective Memory: Differences Between Naturalistic and Laboratory Tasks." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 54B (4): P256–P269. doi:10.1093/geronb/54B.4.P256.

Rushton, J. P., C. J. Brainerd, and M. Pressley. 1983. "Behavioral Development and Construct Validity: The Principle of Aggregation." *Psychological Bulletin* 94 (1): 18–38. doi:10.1037/0033-2909.94.1.18.

Salthouse, T. A. 2009. "Decomposing Age Correlations on Neuropsychological and Cognitive Variables." *Journal of the International Neuropsychological Society* 15 (5): 650–661. doi:10.1017/S1355617709990385.

Salthouse, T. A. 2014. "Why are There Different Age Relations in Cross-Sectional and Longitudinal Comparisons of Cognitive Functioning?" *Current Directions in Psychological Science* 23 (4): 252–256. doi:10.1177/0963721414535212.

Salthouse, T. A. 2015. "Individual Differences in Working Memory and Aging." In *Working Memory and Ageing*, edited by R. H. Logie and R. G. Morris, 1–20. East Sussex, UK: Psychology Press.

Sandberg, K., J. U. Blicher, M. Y. Dong, G. Rees, J. Near, and R. Kanai. 2014. "Occipital GABA Correlates with Cognitive Failures in Daily Life." *NeuroImage* 87: 55–60. doi:10.1016/j.neuroimage.2013.10.059.

Schaie, K. W., and S. L. Willis. 1993. "Age Difference Patterns of Psychometric Intelligence in Adulthood: Generalizability Within and Across Ability Domains." *Psychology and Aging* 8 (1): 44–55. doi:10.1037/0882-7974.8.1.44.

Schmalzried, R. C. 2012. *Patterns of Rumination by Young and Older Adults* (Doctoral dissertation). Lawrence, KS: University of Kansas.

Schmidt, I. W., I. J. Berg, and B. G. Deelman. 1999. "Illusory Superiority in Self-Reported Memory of Older Adults." *Aging, Neuropsychology, and Cognition* 6 (4): 288–301. doi:10.1076/1382-5585(199912)06:04;1-B;FT288.

Schmidt, F. L., and J. E. Hunter. 1999. "Theory Testing and Measurement Error." *Intelligence* 27 (3): 183–198. doi:10.1016/S0160-2896(99)00024-0.

Schwarz, N. 1999. "Self-Reports: How the Questions Shape the Answers." *American Psychologist* 54 (2): 93–105. doi:10.1037/0003-066X.54.2.93.

Sedikides, C., R. Meek, M. D. Alicke, and S. Taylor. 2014. "Behind Bars But above the Bar: Prisoners Consider Themselves More Prosocial Than Non-Prisoners." *The British Journal of Social Psychology* 53 (2): 396–403. doi:10.1111/bjso.12060.

Shariff, S. Z., S. A. D. Bejaimal, J. M. Sontrop, A. V. Iansavichus, R. B. Haynes, M. A. Weir, and A. X. Garg. 2013. “Retrieving Clinical Evidence: A Comparison of PubMed and Google Scholar for Quick Clinical Searches.” *Journal of Medical Internet Research* 15 (8): e164–. doi:10.2196/jmir.2624.

Sharit, J. 2006. “Human error.” In *Handbook of Human Factors and Ergonomics*, edited by G. Salvendy. 3rd ed., 708–760. New York: John Wiley & Sons.

Sheridan, T. B. 2008. “Risk, Human Error, and System Resilience: Fundamental Ideas.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50 (3): 418–426. doi:10.1518/001872008X250773.

Simon, F., and C. Corbett. 1996. “Road Traffic Offending, Stress, Age, and Accident History Among Male and Female Drivers.” *Ergonomics* 39 (5): 757–780. doi:10.1080/00140139608964497.

Smilek, D., J. S. Carriere, and J. A. Cheyne. 2010. “Failures of Sustained Attention in Life, Lab, and Brain: Ecological Validity of the SART.” *Neuropsychologia* 48 (9): 2564–2570. doi:10.1016/j.neuropsychologia.2010.05.002.

Stanton, N. A., and S. V. Stevenage. 1998. “Learning to Predict Human Error: Issues of Acceptability, Reliability and Validity.” *Ergonomics* 41 (11): 1737–1756. doi:10.1080/001401398186162.

Staudinger, U. M. 1999. “Older and Wiser? Integrating Results on the Relationship Between Age and Wisdom-Related Performance.” *International Journal of Behavioral Development* 23 (3): 641–664. doi:10.1080/016502599383739.

Sunderland, A., J. E. Harris, and A. D. Baddeley. 1983. “Do Laboratory Tests Predict Everyday Memory? A Neuropsychological Study.” *Journal of Verbal Learning and Verbal Behavior* 22 (3): 341–357. doi:10.1016/S0022-5371(83)90229-3.

Sunderland, A., K. Watts, A. D. Baddeley, and J. E. Harris. 1986. “Subjective Memory Assessment and Test Performance in Elderly Adults.” *Journal of Gerontology* 41 (3): 376–384. doi:10.1093/geronj/41.3.376.

Sundström, A. 2008. “Self-Assessment of Driving Skill – A Review From a Measurement Perspective.” *Transportation Research Part F: Traffic Psychology and Behaviour* 11: 1–9.

Svenson, O. 1981. “Are We all Less Risky and More Skillful Than Our Fellow Driver?.” *Acta Psychologica* 47 (2): 143–148. doi:10.1016/0001-6918(81)90005-6.

Tucker-Drob, E. M., and T. A. Salthouse. 2008. “Adult Age Trends in the Relations Among Cognitive Abilities.” *Psychology and Aging* 23 (2): 453–460. doi:10.1037/0882-7974.23.2.453.

Unsworth, N., G. A. Brewer, and G. J. Spillers. 2012. “Variation in Cognitive Failures: An Individual Differences Investigation of Everyday Attention and Memory Failures.” *Journal of Memory and Language* 67 (1): 1–16. doi:10.1016/j.jml.2011.12.005.

Uttl, B., and E. Greene. 2008. “Transparent Meta-Analysis of Prospective Memory and Aging.” *Plos One* 3 (2): e1568, doi:10.1371/journal.pone.0001568.

Uttl, B., and M. Kibreab. 2011. “Self-Report Measures of Prospective Memory are Reliable But Not Valid.” *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale* 65: 57–68.

Van der Werf-Eldering, M. J., H. Burger, N. Jabben, E. A. Holthausen, A. Aleman, and W. A. Nolen. 2011. “Is the Lack of Association Between Cognitive Complaints and Objective Cognitive Functioning in Patients with Bipolar Disorder Moderated by Depressive Symptoms?” *Journal of Affective Disorders* 130 (1–2): 306–311. doi:10.1016/j.jad.2010.10.005.

Van Yperen, N. W., and B. P. Buunk. 1991. “Illusoire Superioriteit: Het Verband Met Het Belang Van de Verifieerbaarheid Van de Vergelijkingsdimensies [Illusory Superiority: The Relation with Importance and Verifiability of Comparison Dimensions].” In *Fundamentele sociale psychologie*, [Fundamental Social Psychology], edited by J. van der Pligt, W. van der Kloot, and M. Poppe. 5, 186–200. Tilburg: Tilburg University Press.

Vom Hofe, A., G. Mainemarre, and L. C. Vannier. 1998. “Sensitivity to Everyday Failures and Cognitive Inhibition: Are They Related?” *European Review of Applied Psychology/Revue Européenne de Psychologie Appliquée* 48: 49–56.

Wallace, J. C., and S. J. Vodanovich. 2003. “Can Accidents and Industrial Mishaps be Predicted? Further Investigation into the Relationship Between Cognitive Failure and Reports of Accidents.” *Journal of Business and Psychology* 17 (4): 503–514. doi:10.1023/A:1023452218225.

Wallace, J. C., S. J. Vodanovich, and B. M. Restino. 2003. “Predicting Cognitive Failures from Boredom Proneness and Daytime Sleepiness Scores: An Investigation Within Military and Undergraduate Samples.” *Personality and Individual Differences* 34 (4): 635–644. doi:10.1016/S0191-8869(02)00050-8.

Weaver Cargin, J., A. Collie, C. Masters, and P. Maruff. 2008. “The Nature of Cognitive Complaints in Healthy Older Adults With and Without Objective Memory Decline.” *Journal of Clinical and Experimental Neuropsychology* 30 (2): 245–257. doi:10.1080/13803390701377829.

Wilhelm, O., M. Witthöft, and S. Schipolowski. 2010. “Self-Reported Cognitive Failures.” *Journal of Individual Differences* 31 (1): 1–14. doi:10.1027/1614-0001/a000001.

World Health Organization. 1992. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Retrieved from http://www.who.int/classifications/icd/en/bluebook.pdf

Yeh, Y. Y., and C. D. Wickens. 1988. “Dissociation of Performance and Subjective Measures of Workload.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 30: 111–120.

Zelinski, E. M., K. P. Burnight, and C. J. Lane. 2001. “The Relationship Between Subjective and Objective Memory in the Oldest Old: Comparisons of Findings from a Representative and a Convenience Sample.” *Journal of Aging and Health* 13 (2): 248–266. doi:10.1177/089826430101300205.

Zimprich, D., M. Kliegel, and P. Rast. 2011. “The Factorial Structure and External Validity of the Prospective and Retrospective Memory Questionnaire in Older Adults.” *European Journal of Ageing* 8 (1): 39–48. doi:10.1007/s10433-011-0174-8.