Informant personality is associated with ratings of memory problems in older adults

Buchanan, T. and Loveday, C.

This is a copy of the accepted author manuscript of the following article: Buchanan, T. and Loveday, C. (2018) Informant personality is associated with ratings of memory problems in older adults, *American Journal of Alzheimer's Disease and Other Dementias,* Volume: 33 issue: 7, page(s): 479-489, DOI: 10.1177/1533317518790540.

The final definitive version is available from the publisher Sage at:

https://dx.doi.org/10.1177/1533317518790540

Informant personality is associated with ratings of memory problems in older adults.

Tom Buchanan and Catherine Loveday

University of Westminster

Author Note

Tom Buchanan, Department of Psychology, University of Westminster, London, U.K.

Correspondence concerning this article should be addressed to Tom Buchanan, Department of Psychology, University of Westminster, 115 New Cavendish St, London W1W 6UW, United Kingdom. E-mail: T.Buchanan@westminster.ac.uk

Disclosure: The authors have reported no conflicts of interest

Abstract

Memory complaints are a key diagnostic criterion for dementia and Mild Cognitive Impairment. Rating scales can be used to capture information about individuals' memory problems from informants such as family members. However, problems with scale reliability suggest that individual differences influence the ratings informants provide. This project tested whether informants' neuroticism was associated with their ratings of an older adult's memory. In an online study, 293 volunteers completed a Five Factor personality questionnaire and used two memory questionnaires to provide ratings of memory problems in an older individual they knew well. Rater neuroticism correlated positively with estimates of memory problems: more neurotic informants provided higher estimates of memory difficulties in the person they were rating. A second study replicated this finding with 786 volunteers and another widely used memory measure, the AD8. In both studies, exploratory analyses suggested the effect size was large enough to impact on clinical practice.

Keywords: informant report; neuroticism; memory complaint; personality; dementia; AD8

Introduction

Dementia is now widely regarded as one of the biggest health concerns in the industrialised world, with significant consequences for society and the economy[1]. Prompt and accurate diagnosis plays an important part in managing the condition, but because there are no reliable biological markers this requires a series of time-consuming cognitive and medical assessments. Early diagnosis offers the potential to provide timely intervention to those in the first stages, and may also provide vital clues about aetiology and progression. However, with increasing prevalence and shrinking resources, this can present a significant clinical challenge.

One approach is to offer a triage service, which allows front-line clinicians to identify the "worried well" and to divert essential resources to those who are most likely to need it. To be effective this screening approach needs to extend to Mild Cognitive Impairment (MCI), a prodromal condition that affects up to a fifth of the older population.[2] MCI is characterised by the presence of marked age-related cognitive decline in the absence of functional impairment in everyday activities. While a full diagnosis must be supported by an objective cognitive assessment, a critical first step is that a change in cognitive function must be reported by either the patient themselves, or a knowledgeable informant or clinician.[3]

A range of instruments have been developed to specifically capture everyday difficulties that might signal cognitive decline. Typically these use a semi-structured interview (e.g. the Clinical Dementia Rating scale[4]) or questionnaire (e.g. the AD8[5]) to ask individuals or more usually their close relatives to report on a range of cognitive and behavioural functions such as memory, orientation, or personal care. An important issue with this approach is that a key diagnostic feature of both amnestic MCI and Alzheimer's Disease is memory failures.[2] However, memory failures are also the most common complaint in healthy ageing. This has led to the

development of instruments such as Sabbagh et al.'s Alzheimer's Questionnaire,[6] designed to have sufficient sensitivity to distinguish between typical ageing, MCI, and Alzheimer's disease.

A variety of informant report measures have thus been widely used to evaluate memory problems in clinical practice, public screening, and research settings. These tools vary considerably, not just in time taken to complete them, but also in terms of sensitivity and predictive power. In fact, there are indications of potential problems with some informant report measures. Thompson et al[7] showed that informant reports obtained using a modified version of the Prospective and Retrospective Memory Questionnaire did not distinguish MCI patients from controls, and argued that informant reports had limited validity as an indicator of memory problems. This is consistent with other work querying the reliability of informant reports. For example, the Dysexecutive Questionnaire (DEX[8]) can be used to obtain ratings of a patient's cognitive problems – in this with case with executive function – from an independent non-clinician rater such as a family member. Barker et al[9] showed that evaluations of the same person by multiple raters were only moderately correlated, which raises questions about the reliability of individual ratings. Some factor(s) other than actual degree of impairment must therefore be contributing to informant ratings.

There are some indications of this in work by Jorm et al,[10] who found that for a community-dwelling sample of older adults, informant reports of some aspects of memory and intellectual decline were correlated with the informants' own levels of anxiety and depression. Jorm et al concluded that while informant reports have validity as indicators of a target person's cognitive status, they may also be influenced by the informant's own affective state.

A review by Jorm of informant-based measures[11] reported that across three different instruments for rating a target person's cognitive problems, correlations with informant anxiety

and depression ranged from r=.06 to r=.23. Jorm argued that such relationships could arise from cognitive distortions associated with depression influencing the ratings (e.g. more depressed people perceive the problems of the person they are rating as worse). Alternatively, the informant's anxiety and depression could actually be caused by the problems the rated person (with whom they were likely to have a close relationship) was having. Jorm argued that further work was required to evaluate the extent to which such phenomena could affect the accuracy of diagnoses, and necessitate pre-screening of informants on the basis of affective state as well as other characteristics such as the quality of the informant - target relationship.

This work is suggestive but leaves some important questions open (not least, whether such effects would be observed with the psychometric instruments in clinical use today). As well as being reactive to life events, affective states can be influenced by more temporally stable individual differences. For example, there is evidence that trait neuroticism predicts the development of both anxiety and depression.[12] Thus, it is possible that the personality of the informant could influence the ratings they make. This suggestion closely parallels a body of work on self-report measures.

Hints that the personality of the informant may be important come from work done with questionnaires that ask people to rate their own cognitive performance. Such self-reports of memory problems have long been regarded as problematic, given evidence that they may not reflect objectively-measurable memory performance but rather personality characteristics or depression.[13, 14]These findings appear to apply both to people with memory impairments and to people without such diagnoses, both young and old. For example, Buchanan[15] reported that in a student sample, self-reports of executive problems obtained using the DEX were associated with personality rather than objectively measured performance. This finding held for other aspects of

cognitive function (prospective and retrospective memory) as well as executive function.[16] This and other similar work[17,18] indicates that questionnaire self-reports of cognitive problems are primarily associated with personality traits, and the trait of neuroticism in particular. More neurotic people, drawn from non-clinical populations such as student samples, report higher levels of cognitive impairment in themselves. An important characteristic of these findings is that they apply to healthy, unimpaired, younger samples, rather than people drawn from older populations or those otherwise likely to be experiencing cognitive problems. These are the same kind of people who might be completing informant reports about (e.g.) family members.

If unimpaired participants' evaluations of their own cognitive problems are biased by personality, could the same be true of their evaluations of others' problems? If so, then it is possible that the personality characteristics of the informant will influence the likelihood of the rated person being diagnosed as having cognitive problems. This would compromise the validity of informant reports as a diagnostic tool, and could explain findings such as those of Barker et al[9] of poor inter-rater reliability. It could also underpin relationships observed between informants' affective states and the ratings they make.[10]

The purpose of the current project was to ascertain whether an informant's own personality influences their ratings of a target person's memory performance. Specifically, it was hypothesised that people higher in trait neuroticism would rate others as having more memory problems (Hypothesis 1). This was tested in two studies, using different informant report measures that might be encountered in clinical practice.

Study 1

Method

Materials. This study was conducted wholly online, using an established personality

testing website www.personalitytest.org.uk, which attracted over six thousand users per week during the study, and the Qualtrics online research platform. Ethical approval came from the host University's Psychology Research Ethics Committee.

The personality testing website hosts a 41-item Five Factor personality inventory providing measures of Extraversion, Neuroticism, Openness to Experience, Agreeableness and Conscientiousness, as operationalized in the Five Factor Model of Costa and McCrae.[19] This measure was derived from an International Personality Item Pool measure[20] that correlates well with Costa and McCrae's domains, and has been validated for use online.[21] In this inventory, Extraversion is assessed by 9 items such as ''Make friends easily". Agreeableness is assessed by 7 items such as ''Accept people as they are". Conscientiousness is assessed by 10 items such as ''Carry out my plans. Neuroticism is assessed by 8 items such as ''Panic easily''. Openness to Experience is assessed by 7 items such as ''Have a vivid imagination''. Participants respond on a 5-point scale from 1 ''very inaccurate'' to 5 ''very accurate''. They come to the website through a number of routes (see Table 1) and receive brief feedback on their scores.

Information about the target person's memory problems was obtained using the Alzheimer's Questionnaire (AQ[6]). This 21-item measure has been shown to have good concurrent validity and to correlate well with objective neuropsychological tests.[22] The AQ was modified so that 'patient' was replaced with 'person' to reflect the non-clinical setting of the work (e.g. "Does the person become disoriented in unfamiliar places?"). Participants responded to these using a yes / know / don't know answer format. The AQ provides a total score, and also subscales addressing Memory (6 items); Orientation (3 items); Functional Ability (7 items); Visuospatial function (2 items); and Language (3 items). The AQ can be scored in a weighted manner, where certain questions are assigned double weight due to their known sensitivity in

detecting dementia. For the purposes of this project, both weighted and unweighted scores for the AQ total and subscales were computed (Table 3).

Further memory ratings were obtained using a modified version of the 16-item Prospective and Retrospective Memory Questionnaire (PRMQ[23]). The wording of the questionnaire was again changed to refer to the target person. Thus, retrospective memory (ability to remember things that have happened in the past) was measured by 8 items such as "Do they fail to recall things that have happened to them in the last few days?", while prospective memory (ability to remember things one needs to do in the future) was measured by 8 items such as "Do they fail to mention or give something to a visitor that they were asked to pass on?". Participants responded using a 5-point scale from "very often" to "never".

Procedure. On accessing the personality testing website, participants first saw a page describing the inventory, and giving details of the ethical approval of the research project. On indicating their consent, they moved on to a second page with brief instructions and the test items. Radio button response formats on a 5-point scale ('Very Inaccurate - Very Accurate') were used for the personality items. Participants also answered several other items using drop-down menus: age group (in 5-year increments); current location (a comprehensive list of nations); gender; highest level of education; main occupational status. They were also asked how they came to be taking the test (e.g. as part of a class). Finally, participants were asked whether their data could be used in analyses (they were instructed to answer 'no' if they had not answered the questions seriously, or did not give consent). They then saw a debriefing page providing their scores on each of the scales, along with information to help them interpret the scores. Links were provided to contact the researcher, and to information about personality research elsewhere on the internet.

Respondents indicating that their data could be used for research purposes also saw an invitation to take part in the second stage of the study, described as "a project looking at factors that affect our perceptions of memory problems that other people might have". People who followed the link to the second part, which was hosted on the Qualtrics online research platform, saw a further participant information / consent page outlining what was involved in this second stage. Those giving consent were then asked to think about "a living older person who you know well. For example, this could be a family member who is above retirement age". They were asked for this person's age group (5 year bands), sex, and whether they had been diagnosed as having memory problems. They were also asked to note their relationship to the target person. Participants then proceeded to rate the target person's memory using the modified Alzheimer's Questionnaire and PRMQ. They were finally asked once again for consent to use their data, debriefed and thanked. Previous work[24] has suggested that participants recruited through this method may be higher in Openness to Experience than people recruited in other ways. This is unlikely to present a problem for the current study, given that there are (a) no hypotheses around Openness to Experience, and (b) no evidence the Openness score distributions were appreciably restricted in range (Tables 3 and 6).

Data Screening and Processing. Over the period 12th-25th April 2017, 13222 responses were collected by the www.personalitytest.org.uk website. Of these, 9446 indicated their data could be used for research, and were therefore invited to take part in the second phase of the study. Three hundred and thirty-eight proceeded to the second phase, and 301 went on to participate. In this second phase, multiple participation (where individuals respond more than once, either accidentally or deliberately) was prevented using Qualtrics' proprietary procedures. Two hundred and ninety nine participants progressed to the end of the study and indicated their

data could be used for analysis.

Anybody reporting their age as below 16 years (4 individuals) was removed from the sample due to ethical concerns about whether they could be considered to have given valid consent. Data quality was further assured by examining the datafile for unrealistic combinations of demographic data (e.g. people claiming to be children with doctoral degrees) that might indicate mischievous or careless responding. None were found. Finally, two people who indicated that the person they were reporting on was 'me' or 'myself' were also excluded, leaving 293 in the sample.

Participants. Sample size was initially planned to exceed 200 participants, on the basis that this would give over 80% power to detect an effect size of r=.2. This threshold is suggested by Ferguson[25] as a minimum practical effect size, defined as an effect that will have real-world importance, as opposed to just statistical significance. The achieved sample size in fact conferred 93% power to detect such an effect. Demographic characteristics for the 293 participants are shown in Table 1. Age group of respondent was fairly evenly distributed, with around 10% of the sample in each of the age groups from 16-20 up to 51-55, then slowly trailing off. Participants reported coming from 44 separate countries, with the greatest number being located in the USA, followed by the UK. Overall Table 1 suggests the sample is biased towards well-educated North American women, who had most often found the personality inventory by following a link from some other site. The people whose memory performance they rated were typically over 70 years of age and in most cases had not been diagnosed with memory problems (Table 2).

[Table 1 around here]

[Table 2 around here]

Results

Descriptive statistics for all variables are shown in Table 3. Links between personality variables, and the evaluations of the target person's memory problems, were assessed using Pearson's correlations and are shown in Table 4. In line with Hypothesis 1, the neuroticism scores of participants correlated positively with their ratings of the extent to which the target person had memory problems. This was the case for participants' evaluations of the target person's prospective memory, retrospective memory, total score on the AQ, and general memory performance as assessed by the AQ (whether or not the weighted scoring system was used). Simply, the more neurotic a participant was, the more evidence of memory problems they reported perceiving in their target person.

[Table 3 around here]

[Table 4 around here]

The effect sizes found here are relatively low (all below r=.20), which may lead one to speculate whether the effect of rater personality is sufficiently large to be worth considering. However, even relatively small effects might matter in clinical settings, particularly when large numbers of assessments are being carried out. A further exploratory analysis was conducted to test whether, in the current dataset, neuroticism of the rater was associated with the category a previously un-diagnosed person fell into.

Malek-Ahmadi et al[26] suggest that a total (weighted) AQ score of 5 or above would indicate the possibility of mild cognitive impairment. It may be unwise to use conventional cutoff points with a sample tested via the internet, as score distributions may differ from those obtained with traditional paper-and-pencil administrations of a rating scale[27]. The analysis that follows should therefore be treated as 'proof of concept' that rater neuroticism could be

associated with a diagnosis, not as indicating the percentage of rated individuals who had clinically-significant memory problems.

Participants reported in 240 instances that the people they were rating had not been diagnosed with memory problems. Of these 240, 102 had AQ ratings below 5, while 126 had ratings of 5 or above and thus fell into the range where MCI might be suspected (only 15 had AQ ratings greater than 14, the range where Alzheimer's Disease might be suspected).

To explore the relationship between rater characteristics and categorisation of a previously un-diagnosed rated person as having potential memory problems, as indicated by an AQ score of 5 or above, a binary logistic regression was performed using all personality variables plus age group and sex as predictors (Table 5). This indicated that, overall, rater characteristics were not statistically significantly related to categorisation ($X^2(7, N=221)=13.44$, $p=.062$). However, the main point of interest here is the effect size for neuroticism, which was statistically significant ($p=.023$) as an individual predictor of 'diagnostic category'. With an Odds Ratio of 1.06, the implication is that for every one-point increase in rater neuroticism, the person being rated has a 6% higher chance of being classified as potentially having MCI.

[Table 5 around here]

Discussion

The correlations between neuroticism and ratings of memory performance in others, parallel other findings of correlations with ratings of participants' own cognitive performance.[13,15,16] A likely explanation for the phenomenon is that people scoring higher on neuroticism are more prone to worry and experience anxiety. They may thus pay more attention to evidence of cognitive failures, in both themselves and others. An alternative possibility is that

more emotionally stable individuals are prone to underestimate the real levels of difficulty experienced by the people they are rating.

The relationships between memory ratings and extraversion, where people higher in extraversion rated their targets as having lower levels of memory problems (Table 4), are harder to explain. One possibility is that they are artefacts arising from a correlation between extraversion and neuroticism – while theoretically orthogonal, in practice personality trait measures often correlate. In this case, extraversion did correlate significantly with neuroticism (r=-.29, p<.0005, n=293), and the logistic regression (which of course controls for such relationships) indicated extraversion was not associated with diagnostic category. However, in one case extraversion was associated with a variable that neuroticism was not: the AQ Orientation subscale. More extraverted people reported lower levels of orientation problems in the people they rated. The reason for this is not known. Given that we had no hypotheses regarding extraversion, this should be treated as an exploratory finding. However, it does suggest that there may be a relationship worthy of further research.

While neuroticism was associated with total scores on the AQ, this was not the case for all of the AQ subscales. This raises the possibility that some measures may be more susceptible to influence by personality than others. Thus, it is desirable to test whether these findings can be replicated, in particular with measures likely to be widely used in practice. This was the purpose of Study 2.

Study 2

The purpose of Study 2 was to establish whether the findings of Study 1 held true for a different instrument. In this instance, the measure used was the AD8[5]. This is an established test that is widely used for dementia screening. It forms part of the Cognitive Assessment Toolkit

developed by the Alzheimer's Association for use in Medicare assessments.[28] As well as being used in clinical practice, it can easily be found online at a number of locations, ranging from clinician resources to pharmaceutical company websites. It is also possible to download smartphone applications for self-screening purposes. Thus, it is very likely that people worried about their own cognitive status, or that of a loved one, will use it unsupervised by a clinician prior to seeking professional help if the instrument indicates a problem may be present. As in Study 1, it was hypothesised that people higher in trait neuroticism would rate others as having more memory problems (Hypothesis 1).

Method

Materials. This study was conducted wholly online, using substantially the same materials and procedure as Study 1. The key difference was that in Study 2 the AQ and PRMQ were not used. Instead, information about the target person's memory problems was obtained using the Washington University Dementia Screening Test (AD8). The AD8 consists of 8 items, where informants are asked whether they have seen problems in the person they are reporting on. Items address a number of domains, such as "Less interest in hobbies / activities" and "Trouble remembering appointments". They are responded to using the options 'Yes, a change'; 'No, no change' and 'N/A, don't know'. Participants are prompted that "Answering 'Yes, a change' indicates that there has been a change in the last several years caused by cognitive (thinking and memory) problems." The final score is the number of items where informants indicate they have seen a change, with scores of 2 or above suggesting an impairment in cognition.[29]

Procedure. The procedure used, and experience of participants, was exactly as described in Study 1, up to the point where they were asked to think about "a living older person who you know well" in the second phase of the study. On this occasion, participants were asked to rate the

target person's memory using only the AD8. Following this they were asked once again for consent to use their data, debriefed and thanked.

Data Screening and Processing. Over the period 26th June-14th November 2017, 33083 responses were collected by the www.personalitytest.org.uk website. Of these, 23377 indicated their data could be used for research, and were therefore invited to take part in the second phase of the study. Eight hundred and seventy-nine proceeded to the second phase, and 812 went on to participate. In this second phase, multiple participation (where individuals respond more than once, either accidentally or deliberately) was prevented using Qualtrics' proprietary procedures. Eight hundred and five participants progressed to the end of the study and indicated their data could be used for analysis. Anybody reporting their age as below 16 years (9 individuals) was removed from the sample for ethical reasons. No unrealistic combinations of demographic data (e.g. people claiming to be children with doctoral degrees) that might indicate mischievous or careless responding were found. Finally, 10 people who indicated that the person they were reporting on was 'me', 'self' or 'myself' were also excluded, leaving 786 in the sample.

Participants. Sample size was again planned to exceed 200 participants to give over 80% power to detect an effect size of r=.2. The achieved sample size in fact conferred 99.99% power to detect such an effect. Demographic characteristics for the 786 participants are shown in Table 1 (right side). Around 10% of the sample reported being in each of the age groups from 16-20 up to 51-55, and trailing off thereafter. While participants reported coming from 64 different countries, around half reported being located in the USA, followed by the UK. The sample is biased towards well-educated North American women, who had most often found the personality inventory by following a link from some other site, and were somewhat younger on average than those in Study 1. The people whose memory performance they rated were typically over 70

years of age, with the largest proportion being 70-74 years of age. In most cases they had not been formally diagnosed with memory problems. In general, the individuals being rated were very similar to those rated in Study 1 in terms of age distribution, gender, and memory problem diagnoses (Table 2, right side).

Results

Descriptive statistics for all variables are shown in Table 6. In line with Hypothesis 1, the Pearson's correlation between neuroticism scores of participants and their AD8 ratings of the extent to which the target person had memory problems was positive and statistically significant (r=.097, p=.007). While statistically significant, this effect size is small. This may in part be due to the distributional properties of the AD8 scores, which were positively skewed with the majority of observations towards the lower end of the scale (0 being the most frequent score). This potentially attenuated correlations due to restriction of range and floor effects.

[Table 6 around here]

As in Study 1, further exploratory analyses evaluated the extent to which the relationship found might have implications for diagnoses. Galvin and Zweig[29] suggest that AD8 scores of 2 or above may be indicative of cognitive impairment. This threshold was exceeded by 534 (68.5%) of the individuals reported on. When cross-tabulated against those cases for whom a formal diagnosis of memory problems was reported (Table 7), it is clear that there are many individuals for whom cognitive impairment might be suspected, but for whom there has not been a formal diagnosis of memory problems. To what extent might a participant's neuroticism level influence their likelihood of giving someone an AD8 rating that exceeds the cut-off score of 2?

[Table 7 around here]

Exploratory analyses were again performed to assess the effect of personality on AD8 scores for that subgroup where informant ratings might be most important – people who do not already have a clinical diagnosis of memory problems. Among these individuals a stronger correlation between rater neuroticism and AD8 scores was observed (r=.13, n=584, p=.001).

A binary logistic regression (Table 8) was again performed to assess degree of association between rater characteristics and categorisation of the target person as having cognitive problems (as indicated by an AD8 score of 2 or above). The analysis showed that rater characteristics were associated with target categorisation ($\boldsymbol{X}^2(7, N=561)=21.58$, p=.003), though with a small effect size (Cox & Snell R^2=.038). The individual effect of neuroticism is once again the key point of interest. Neuroticism was significantly associated with categorisation based on AD8 scores, with an Odds Ratio of 1.05 indicating that for every one-point increase in rater neuroticism, the person being rated had a 5% higher chance of being classified as potentially having cognitive impairment.

[Table 8 around here]

Discussion

Study 2 conceptually replicated the main finding of Study 1, with a different assessment instrument. Informant neuroticism once again correlated positively with ratings of memory problems in an older adult. Among individuals who did not already have a diagnosis of memory problems, higher rater neuroticism was associated with increased likelihood of an AD8 score exceeding the cutoff for suspected problems. Study 1 had also reported a correlation between rater extraversion and the memory problem ratings they provided. This was not observed in Study 2 (Table 6).

The logistic regression analysis here also (but not in Study 1) found that age group was associated with AD8 classification – the older an informant was, the less likely they were to give a participant an AD8 score of two or above. Given that the informants were relatively young, it may be that younger participants were less aware of the natural process of cognitive aging and were thus more extreme in their judgements. Alternatively, their relationships with the people being rated may have been more distant (e.g. grandparents, rather than spouses). If this is a genuine replicable effect and not an anomaly, further work would be needed to understand the phenomenon.

General Discussion

Across two studies, using three different informant report measures, evidence has been presented that indicates the neuroticism level of an informant is associated with the ratings they assign to others on memory scales. Raw correlations are low in magnitude (especially in Study 2, where they may however have been affected by distributional properties of the outcome variable). However, given the way that assessment instruments are used in practice, it may be more informative to consider the association between personality and assignment to potential diagnostic categories.

For both studies, a one-point increase in participant neuroticism was associated with around a 5% increase in the likelihood of providing a score consistent with a potential diagnosis of cognitive problems. When one considers that the observed range of rater neuroticism scores was 32 points, this suggests that if your memory is being rated by a person high on neuroticism you have a higher probability being classified as potentially having cognitive impairment, than if your rater is more emotionally stable.

The current findings reinforce Barker et al's finding[9] that raters may vary, and suggest this may be associated with individual differences in rater neuroticism. The phenomenon described here could also underpin Jorm et al's observation[10] that anxiety and depression were associated with informant ratings, given that people high on neuroticism may be more likely to report experiencing those states. A possible explanation for all of these observations is that individuals high on neuroticism worry and pay more attention to cognitive failures they perceive in others, thus providing inflated reports.

A further possibility is that neuroticism is not associated with a bias in symptom reporting, but rather the degree of accuracy with which symptoms are observed. It is possible that more neurotic people do not over-report problems in others, but are better at detecting them. In the literature on links between neuroticism and symptom self-reporting, some researchers have hypothesised that neuroticism may be associated with greater vigilance for health threats or attention to symptoms[30]. Thus, it is possible that more neurotic informants are more attentive to markers of problems in a patient, and therefore report more errors being made. In this scenario, more neurotic informants would provide better quality information. To test this hypothesis, objective data about the cognitive performance of the target is required so that the accuracy of the informant ratings can be assessed.

Exploratory analyses in Study 1 indicated that extraversion may also be associated with ratings on some aspects of cognitive performance, independent of neuroticism. Its importance is likely to be lower than neuroticism's, given that it had no relationship with AD8 scores in Study 2, and was not associated with categorisation of the target as potentially having memory problems in either Study 1 or Study 2. In common with the neuroticism findings however, the observation that extraversion was associated with some but not all components of the measures

used reinforces the notion that personality may have different effects on different types of rating scale item or construct measured.

The work reported here has a number of limitations. A key issue is the nature of the sample – highly self-selected volunteers – and the target persons they were rating. The nature of the target persons, other than being 'a living older person', and the rater's relationship to them, was not specified. This may have led to a sample of rated individuals who were less likely to display cognitive failures than those who might be assessed in settings such as memory clinics. In real life, only individuals suspected of having memory problems would be rated using instruments such as the AQ or AD8. This means that the current sample are likely to receive ratings of memory problems lower than would be the case in a real assessment context, potentially attenuating correlations with rater personality, and leading to an underestimate of the effect size.

Furthermore, the relationships between raters and targets varied considerably – in some cases people were rating their spouses, in other cases parents or grandparents, in yet others friends, neighbours or work colleagues (though in each study, over 90% of the targets were described as being family members). The extent to which raters were familiar with the behaviors and experiences of the people they were rating was thus varied. Further work is required with a more realistic 'patient' sample, and with samples where degree of familiarity is more closely controlled, to get a better idea of the magnitude of the phenomenon.

Another question that can only be answered by more clinically-focused work, is which end of the trait neuroticism continuum is associated with more accurate assessments. Is it the case that people high on neuroticism overestimate problems in others, or that people low on neuroticism underestimate them? This question needs to be answered, by work in which

informant ratings can be compared with clinical judgements or objective test data, in order to fully understand the implications of the findings.

The phenomenon documented in this paper could potentially lead to either over- or under- referral of patients with suspected memory problems for further assessment. If high neuroticism leads raters to exaggerate memory problems, this would increase the rate of false-positive referrals, and waste clinical time and resources. Conversely, if lower neuroticism leads to underestimation of memory problems, individuals who should be referred for further testing might not be. Of the two possibilities, the latter probably has the most severe consequences for individuals (failure to obtain a timely diagnosis, and hence delay in whatever treatment or management options are available). However, both scenarios may have implications for clinical practice. If the effect size is substantive enough to actually influence the likelihood of an accurate clinical diagnosis, then clinicians may therefore need to take informant personality into account. A factor that has even a small effect on the likelihood of false positive or false negative diagnoses could be important in a context where very large numbers of assessments are taking place. But how could this be addressed in practice?

Jorm[11] suggested one might 'prescreen' informants to eliminate those with characteristics that would lead them to provide less valid information. However, in clinical practice this is unlikely to be feasible or desirable. Another potential approach would be 'correcting' informant ratings on the basis of informant neuroticism scores, though collecting the data required for this would increase the assessment burden on informants.

Perhaps more realistically, the findings could drive selection or development of questionnaires and rating scales that are less likely to be biased by informant characteristics. Some evidence that this might be feasible comes from Table 4: there were a number of AQ

subscales that were not at all associated with rater personality. It is possible that particular abilities or item types are differentially influenced. For example, items asking about observable behaviour (e.g. 'are they often late for appointments?') may function differently from those requiring inferences about subjective experience (e.g. 'do they find themselves forgetting about appointments?'). The latter might be more susceptible to being influenced by traits or states of the person doing the rating. Further research on this question would be of value to inform the construction of 'personality proof' informant rating measures.

While the present studies have a number of limitations, the findings do suggest that this previously un-documented effect is worth knowing about. Given the potential growing importance of questionnaires and rating scales for screening memory problems in an aging population, more work in this vein is likely to be required to improve maximise their usefulness. A critical question is how much this phenomenon actually matters, if at all. In real-life clinical practice, how many patients would be misdiagnosed (either false positive or false negative) due to the effects of informant personality? The scale of this issue clearly needs to be addressed by further research. However, with very large numbers of screening assessments being carried out, there is potential for the effect we report here to be a real problem.

References

1. Alzheimer's Association. 2017 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2017;13 (4):325-373.

2. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA*. 2014;312 (23):2551-2561.

3. Albert MS, DeKosky ST, Dickson D et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia*. 2011;7 (3):270-279.

4. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*. 1993;43 (11):2412.

5. Galvin JE, Roe CM, Xiong C, Morris JC. Validity and reliability of the AD8 informant interview in dementia. *Neurology*. 2006;67 (11):1942-1948.

6. Sabbagh MN, Malek-Ahmadi M, Kataria R et al. The Alzheimer's questionnaire: a proof of concept study for a new informant-based dementia assessment. *J Alzheimers Dis*. 2010;22 (3):1015-1021.

7. Thompson CL, Henry JD, Rendell PG, Withall A, Brodaty H. How valid are subjective ratings of prospective memory in mild cognitive impairment and early dementia? *Gerontology*. 2015;61 (3):251-257.

8. Wilson BA, Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. *Behavioural Assessment of the Dysexecutive Syndrome (BADS)*. Bury St. Edmunds, U.K: Thames Valley Test Company; 1996.

9. Barker LA, Morton N, Morrison TG, McGuire BE. Inter-rater reliability of the Dysexecutive Questionnaire (DEX): comparative data from non-clinician respondents-all raters are not equal. *Brain Inj*. 2011;25 (10):997-1004.

10. Jorm AF, Christensen H, Henderson AS, Korten AE, Mackinnon AJ, Scott R. Complaints of cognitive decline in the elderly: a comparison of reports by subjects and informants in a community survey. *Psychol Med*. 1994;24 (02):365.

11. Jorm AF. Assessment of cognitive impairment and dementia using informant reports. *Clinical Psychology Review*. 1996;16 (1):51-73.

12. Zinbarg RE, Mineka S, Bobova L et al. Testing a Hierarchical Model of Neuroticism and Its Cognitive Facets: Latent Structure and Prospective Prediction of First Onsets of Anxiety and Unipolar Mood Disorders During 3 Years in Late Adolescence. *Clinical Psychological Science*. 2016;4 (5):805-824.

13. Pearman A, Hertzog C, Gerstorf D. Little evidence for links between memory complaints and memory performance in very old age: longitudinal analyses from the Berlin Aging Study. *Psychol Aging*. 2014;29 (4):828-842.

14. Slavin MJ, Brodaty H, Kochan NA et al. Prevalence and Predictors of "Subjective Cognitive Complaints" in the Sydney Memory and Ageing Study. *The American Journal of Geriatric Psychiatry*. 2010;18 (8):701-710.

15. Buchanan T. Self-report measures of executive function problems correlate with personality, not performance-based executive function measures, in nonclinical samples. *Psychol Assess*. 2016;28 (4):372-385.

16. Buchanan T. Self-assessments of memory correlate with neuroticism and conscientiousness, not memory span performance. *Personality and Individual Differences*. 2017;105 19-23.

17. Uttl B, Kibreab M. Self-report measures of prospective memory are reliable but not valid. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*. 2011;65 (1):57-68.

18. Williams PG, Rau HK, Suchy Y, Thorgusen SR, Smith TW. On the validity of self-report assessment of cognitive abilities: Attentional control scale associations with cognitive performance, emotional adjustment, and personality. *Psychological Assessment*. 2017;29 (5):519-530.

19. Costa PT, McCrae RR. *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources; 1992.

20. Goldberg LR. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In: I. Mervielde, I.J. Deary, F. De Fruyt, FO, eds. *Personality Psychology in Europe Vol. 7*. Tilburg, The Netherlands: Tilburg University Press; 1999:7-28.

21. Buchanan T, Johnson JA, Goldberg LR. Implementing a Five-Factor Personality Inventory for Use on the Internet. *European Journal of Psychological Assessment*. 2005;21 (2):115-127.

22. Budolfson K, Malek-Ahmadi M, Belden CM et al. Neuropsychological Correlates of the Alzheimer's Questionnaire. *Journal of Alzheimers Disease*. 2015;46 (2):389-397.

23. Smith G, Della Sala S, Logie RH, Maylor EA. Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory*. 2000;8 (5):311-321.

24. Buchanan T. Personality biases in different types of 'internet samples' can influence research outcomes. *Computers in Human Behavior*. 2018;86: 235-244.

25. Ferguson CJ. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*. 2009;40 (5):532-538.

26. Malek-Ahmadi M, Davis K, Belden C et al. Validation and diagnostic accuracy of the Alzheimer's questionnaire. *Age Ageing*. 2012;41 (3):396-399.

27. Buchanan T. Personality testing on the Internet: What we know and what we do not. In: Joinson AN, KcKenna KYA, Postmes T, Reips U-D, eds. *The Oxford Handbook of Internet Psychology*. Oxford, UK: Oxford University Press; 2007:447-459.

28. Cordell CB, Borson S, Boustani M et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimers Dement*. 2013;9 (2):141-150.

29. Galvin JE, Zweig Y. The AD8: The Washington University Dementia Screening Test. In: Greenberg SA, ed. *Try this: Best Practice in Nursing Care to Older Adults with Dementia*. New York: The Hartford Institute for Geriatric Nursing, New York University, College of Nursing; 2013.

30. Taga KA, Friedman HS, Martin LR. Early personality traits as predictors of mortality risk following conjugal bereavement. *J Pers*. 2009;77 (3):669-690.

Tables

Table 1
Demographic Data

Measure		Study 1	Study 2
N		293	786
Sex			
	Men	47 (16.0%)	154 (19.6%)
	Women	235(80.2%)	604 (76.8%)
	Unanswered	11 (3.8%)	28 (3.6%)
Age			
	Modal age group	36-40 (13.7%)	16-20 (12.7%)
	Age range	16-75	16-85+
	Unanswered	0	0
Location			
	USA	143 (48.8%)	389 (49.5%)
	UK	27 (9.2%)	111 (14.1%)
	Other	122 (41.7%)	276 (35.1%)
	Unanswered	3 (0.6%)	10 (1.3%)
Route to participation			
	Followed link from another site	205 (70.0%)	388 (49.4%)
	Doing as part of some class	30 (10.2%)	119 (15.1%)
	Found through search engine	17 (5.8%)	134 (17.0%)
	Got link from a friend	14 (4.8%)	36 (4.6%)
	Other	24 (8.3%)	98 (12.5%)
	Unanswered	3 (1.0%)	11 (1.4%)
Highest level of education			
	Primary Education	10 (3.4%)	10 (1.3%)
	Secondary Education	53 (18.1%)	131 (16.7%)
	Vocational / Technical college	33 (11.3%)	77 (9.8%)
	Some college / University	56 (19.1%)	170 (21.6%)
	College / University Graduate	73 (24.9%)	206 (26.2%)
	Some Postgraduate	34 (11.6%)	96 (12.5%)
	Postgraduate / Professional Degree	34 (11.6%)	93 (11.8%)
	Unanswered	0 (0%)	1 (0.1%)
Occupation			
	Employed for Wages	133 (45.4%)	362 (46.1%)
	Self-employed	31 (10.6%)	75 (9.5%)
	Unemployed	20 (6.8%)	44 (5.6%)
	Home-maker	14 (4.8%)	26 (3.3%)
	Student	46 (15.7%)	156 (19.8%)
	Retired	16 (5.5%)	47 (6.0%)
	Unable to work	19 (6.5%)	39 (5.0%)
	Unanswered	14 (4.8%)	37 (4.7%)

Note. Percentages may not sum exactly to 100% due to rounding errors

Table 2
The Person Being Rated

Measure		Study 1	Study 2
N		293	786
Sex			
	Men	105 (35.8%)	263 (33.5%)
	Women	178 (60.8%)	498 (63.4%)
	Unanswered	10 (3.4%)	25 (3.2%)
Has the person been diagnosed with memory problems?			
	Yes	30 (10.2%)	121 (15.4%)
	No	240 (81.9%)	590 (75.1%)
	Don't Know	14 (4.8%)	53 (6.7%)
	Unanswered	9 (3.1%)	22 (2.8%)
Age of person being rated			
	Below 60	16 (5.5%)	58 (7.4%)
	60-64	35 (11.9%)	74 (9.4%)
	65-69	37 (12.6%)	101 (12.8%)
	70-74	53 (18.1%)	157 (20.0%)
	75-79	42 (14.3%)	107 (13.6%)
	80-84	50 (17.1%)	117 (14.9%)
	85-89	30 (10.2%)	67 (8.5%)
	90-94	14 (4.8%)	46 (5.9%)
	95-99	3 (1.0%)	10 (1.3%)
	100 or over	1 (0.3%)	1 (0.1%)
	Unanswered	12 (4.1%)	48 (6.1%)

Note. Percentages may not sum exactly to 100% due to rounding errors.

Table 3
Participant's Own Personality, and Ratings of Memory of Target Person

Variable		*n*	*M*	*SD*	α	Range Potential	Range Actual	Skew
Own personality								
	Extraversion	293	28.09	7.30	.86	9-45	11-45	-0.04
	Agreeableness	293	28.89	3.96	.72	7-35	14-35	-0.81
	Conscientiousness	293	34.99	6.87	.81	10-50	16-50	-0.27
	Neuroticism	293	21.68	6.90	.85	8-40	8-40	0.19
	Openness to Experience	293	27.56	4.61	.68	7-35	12-35	-0.64
PRMQ								
	Retrospective	289	19.43	6.86	.89	8-40	8-40	0.52
	Prospective	290	18.33	6.90	.90	8-40	8-40	0.61
Alzheimer's Questionnaire								
	Total	280	6.14	5.20	.90	0-21	0-21	0.91
	Memory	287	2.60	1.92	.75	0-6	0-6	0.15
	Orientation	291	0.70	1.00	.70	0-3	0-3	1.13
	Functional Ability	292	1.85	2.07	.81	0-7	0-7	1.07
	Visuospatial	292	0.30	0.59	.55	0-2	0-2	1.78
	Language	289	0.70	0.85	.47	0-3	0-3	0.95
Alzheimer's Questionnaire (Weighted scores)								
	Total	280	7.51	6.54	-	0-27	0-27	0.99
	Memory	287	3.04	2.26	-	0-7	0-7	0.15
	Orientation	291	0.89	1.31	-	0-4	0-4	1.23
	Functional Ability	292	2.10	2.43	-	0-8	0-8	1.09
	Visuospatial	292	0.38	0.81	-	0-3	0-3	2.31
	Language	289	1.10	1.39	-	0-5	0-5	1.07

Table 4
Pearson's r Correlations of Participant's Own Personality with Ratings of Memory of Target Person (n=293).

Variable		Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness to Experience
PRMQ						
	Retrospective	-.17**	.04	-.01	.15*	.07
	Prospective	-.14*	-.02	.00	.17**	.09
Alzheimer's Questionnaire						
	Total	-.14*	.00	-.02	.15*	.07
	Memory	-.10	-.06	.01	.17**	.05
	Orientation	-.19**	.05	-.07	.11	.04
	Functional Ability	-.10	.02	-.04	.10	.11
	Visuospatial	-.09	.01	-.03	.09	.05
	Language	-.08	.01	.04	.07	.01
Alzheimer's Questionnaire (Weighted scores)						
	Total	-.14*	-.01	-.02	.15*	.06
	Memory	-.09	-.07	.00	.18**	.04
	Orientation	-.18**	.05	-.07	.10	.03
	Functional Ability	-.11	.02	-.04	.09	.11
	Visuospatial	-.09	-.01	-.03	.08	.04
	Language	-.08	.00	.04	.07	.01

**p<.05, **p<.01*

Table 5
Binary Logistic Regression: Rater Personality as Predictor of MCI 'diagnosis'

	B	*S.E.*	*Wald*	*df*	*p*	*Exp(B)*	95% CI for *Exp(B)*
Extraversion	-0.03	0.02	1.79	1	.18	0.97	[0.94, 1.01]
Agreeableness	0.00	0.04	0.00	1	.97	1.00	[0.93, 1.08]
Conscientiousness	0.05	0.02	4.81	1	.028	1.05	[1.01, 1.10]
Neuroticism	0.06	0.03	5.18	1	.023	1.06	[1.01, 1.12]
Openness to Experience	0.01	0.03	0.02	1	.88	1.01	[0.94, 1.07]
Age Group	-0.02	0.05	0.09	1	.77	0.99	[0.89, 1.09]
Sex (M=1, F=2)	0.52	0.39	1.73	1	.19	1.68	[0.78, 3.64]
Constant	-3.14	1.98	2.50	1	.11	0.04	

Table 6
Participant's own Personality, and AD8 Rating of Memory of Target Person

Variable		*n*	*M*	*SD*	*α*	Range		Skew	Correlation (*r)* with AD8
						Potential	Actual		
Personality									
	Extraversion	786	28.23	7.66	0.87	9-45	9-45	-0.17	-.03
	Agreeableness	786	28.35	4.50	0.75	7-35	10-35	-0.99	.02
	Conscientiousness	786	35.94	6.93	0.82	10-50	15-50	-0.30	-.04
	Neuroticism	786	21.77	6.87	0.83	8-40	8-40	0.23	.097**
	Openness to Experience	786	27.80	4.79	0.69	7-35	13-35	-0.61	-.01
AD8									
	Total AD8 score	779	3.05	2.40	0.79	0-8	0-8	0.52	-

***p<.01*

Table 7
Cross-tabulation of AD8 'Diagnosis' and Reported Formal Diagnosis of Memory Problems

		Formal diagnosis			
		Yes	No	Don't know	Total
AD8 'diagnosis'	No impairment	2	226	12	240
	Impairment	119	358	41	518
Total		121	584	53	758

Note. AD8 scores of 0 or 1 were classified as 'no impairment', scores of 2 or above as 'impairment'.

Table 8
Binary Logistic Regression: Rater Personality as Predictor of AD8 'Diagnosis' for Rated Individuals Without a Formal Diagnosis of Memory Problems

	B	*S.E.*	*Wald*	*df*	*p*	*Exp(B)*	*95% CI for EXP(B)*
Extraversion	0.01	0.01	0.52	1	.47	1.01	[0.98, 1.04]
Agreeableness	0.03	0.02	1.45	1	.23	1.03	[0.98, 1.07]
Conscientiousness	0.00	0.02	0.05	1	.82	1.00	[0.98, 1.03]
Neuroticism	0.04	0.02	7.85	1	.005	1.05	[1.01, 1.08]
Openness to Experience	0.01	0.02	0.17	1	.68	1.01	[0.97, 1.05]
Age Group	-0.07	0.03	5.08	1	.02	0.93	[0.87, 0.99]
Sex (M=1, F=2)	0.41	0.22	3.46	1	.06	1.50	[0.98, 2.31]
Constant	-1.97	1.12	3.12	1	.08	0.14	