Self-report measures of executive function problems correlate with personality, not performance-based executive function measures, in non-clinical samples

Tom Buchanan

Department of Psychology, Faculty of Science and Technology, University of Westminster

This is the final author version of an article which has been accepted for publication, 28th May, 2015, in Psychological Assessment, www.apa.org/pubs/journals/pas. This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Self-report measures of executive function problems correlate with personality, not performance-based executive function measures, in non-clinical samples

Tom Buchanan

University of Westminster

Author Note

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

For Study 3, the assistance of Catherine Loveday with selection of tests, and Maria Flynn with data collection, is gratefully acknowledged.

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

Abstract

Researchers and clinicians often measure executive function in patients and normal samples. In addition to cognitive tests that objectively measure executive function, several instruments have been developed that address individuals' everyday experience of executive problems. Such self-report measures of executive problems may have value, but there are questions about the extent to which they tap objectively-measurable executive problems or are influenced by variables such as personality. Relationships between self-reported executive problems, personality, and cognitive test performance were assessed in three separate, well-powered, methodologically distinct correlational studies using non-clinical samples. These studies used multiple measures of personality and self-reported executive function problems. Across all three studies, self-reported executive function problems were found to correlate with neuroticism and with low conscientiousness, with medium to large effect sizes. However self-reported problems did not correlate with performance on Trail Making, Phonemic Fluency, Semantic Fluency or Digit Span tests tapping executive function. A key implication of these findings is that in non-clinical samples, self-report questionnaires may not be proxies for executive functioning as measured by neuropsychological tests.

Keywords: executive function, self-report, dysexecutive syndrome, neuroticism, conscientiousness, personality

Introduction

Self-report measures of executive problems may provide useful information in research and clinical contexts. They are relatively easy to deploy, and have the potential to provide useful and ecologically-valid information about respondents' experiences. However, questions may be asked about what such instruments actually measure. This project set out to ascertain whether responses to two self-report measures actually reflected objectively-measured executive function, or instead were better accounted for by specific aspects of the personality of respondents. If the latter, there are potentially significant problems for the interpretation of data derived from self-report measures of executive function.

Executive Function and its Measurement

The term 'executive function' refers to a collection of processes that play critical roles in information processing, making up the central executive component of the working memory model (Baddeley, 2003). Executive functions are involved in planning, task coordination, impulse control and attentional control. They have been variously described, measured and explained by numerous researchers (Salthouse, 2005). Drawing on the work of Miyake, Friedman, Emerson, Witzki, Howerter, and Wager (2000) and others, Diamond (2013) notes there is general agreement that three core executive functions exist: Working Memory, Cognitive Flexibility, and Inhibitory Control. Miyake and Friedman (2012) refer to these respectively as "*updating* (constant monitoring and rapid addition/deletion of working-memory contents); *shifting* (switching flexibly between tasks or mental sets); and *inhibition* (deliberate overriding of

dominant or prepotent responses)" (p. 9). While failures of executive function are most pronounced in clinical groups, they are also seen in non-clinical populations (Banich, 2009; Chan, 2001). There are also individual differences in executive function (e.g. Miyake & Friedman, 2012; Williams, Suchy, & Rau, 2009) that relate to important life outcomes (e.g. Diamond, 2013). Executive functions are known to change over the life course, from childhood to old age (e.g. Diamond, 2013).

Measurement of executive function is important in both clinical and research settings. Objective neuropsychological measures such as the Trail Making Test (TMT; Reitan, 1955) or computerized cognitive assessment batteries are the gold-standard tools for assessing executive function. However, there are also strong arguments that self-report measures of executive problems are of value (Barkley & Fischer, 2011; Johnco, Wuthrich & Rapee, 2014; Kamradt, Ullsperger & Nikolas, 2014). A number of such measures have been developed, including the Dysexecutive Questionnaire (DEX; Wilson, Alderman, Burgess, Emslie, & Evans, 1996), the Frontal Systems Behavior Scale (FrSBe; Grace & Malloy, 2001), and Webexec (Buchanan, Heffernan, Parrott, Ling, Rodgers & Scholey, 2010), which was developed for use on the internet.

Value of Self-report Cognitive Measures

It has been argued that cognitive self-assessment questionnaires are "indispensible instruments for the study of everyday cognition" (Rabbitt, Maylor, McInnes, Bent & Moore, 1995, p. S149). This is in part because neuropsychological tests may not always be sensitive to problems people feel they are experiencing. There are multiple reports in the literature of studies where people report executive impairments, but their self-reports do not appear to correlate with scores on batteries of executive neuropsychological tests (e.g. Barkley & Fisher, 2011; Laws,

Patel & Tyson, 2008).

In addition to value with clinical groups, self-report measures are also useful as research tools with non-clinical populations. For example, Heffernan and O'Neill (2013) used a self- report measure (as described by Buchanan et al., 2010) to examine links between executive deficits and exposure to second-hand smoke. Chan (2001) argues that members of the normal population may experience executive failures that may not be serious enough to be considered as clinical impairments, but nonetheless have the capacity to impact on everyday life. Chan showed that in a Hong Kong Chinese community sample there was evidence of everyday executive problems to which a translated version of the self-report DEX was sensitive, but objective tests were not.

As with many self-report research tools, numerous researchers are interested in deploying self-report cognitive measures on the internet for web based research projects. For example, Friedman-Kraus, Raver, Neuspiel and Kinsel (2014) used the Webexec measure (Buchanan et al., 2010) in an online study of executive function and stress in teachers, alongside computerized versions of reverse letter span and trail making tests. Self-report measures are much easier to deploy online than objective tests. With the advent of widely available, sophisticated online survey tools, their creation requires little if any technical knowledge. However, they may not always work as expected (Buchanan, Ali et al., 2005). This is especially the case if self-reports are taken as proxy measures of actual cognitive function. For example, while Buchanan et al. (2010) found that scores on the Webexec measure correlated with objective measures of executive function obtained in a laboratory setting, when Friedman-Kraus et al. (2014) used the same measure online they found no associations with the objective tests they used. This raises questions about what the test was actually measuring in each context.

Influence of Other Variables on Self-report Measures

Buchanan et al. (2010) acknowledge the possibility that factors other than actual executive function may influence responses to self-report measures such as Webexec. They note Rabbitt et al.'s (1995) argument that self-reports may be influenced by factors such as depression and anxiety. In addition, it is possible that personality may be another factor influencing self- reports of executive functioning. For example, Gerstorf, Siedlecki, Tucker-Drob, and Salthouse (2008) found that neuroticism was among factors associated with scores on the DEX in a community-dwelling sample. Questions therefore arise concerning whether self-report measures of executive problems actually do provide indices of executive function, or whether they are influenced more by individual difference variables. A 'pure' index of any psychological construct or function is unlikely to be achievable. The key question is whether a measure captures sufficient variance in the construct of interest to be of value. Is that the case for self-report executive measures of executive problems?

Personality and Objectively Measured Executive Functions

Consideration of the effects that personality may have on self-reporting of executive problems is complicated by the fact that there may also be genuine associations between personality and different aspects of objectively measurable executive function (Unsworth et al., 2009; Williams et al., 2009). For example, DeYoung, Hirsh, Shane, Papademetris, Rajeevan, and Gray (2010) showed that conscientiousness covaried with volume of a brain area (lateral prefrontal cortex) involved in planning and voluntary control of behavior; including the middle frontal gyrus which they note is involved in maintaining information in working memory. Indeed, the literature contains numerous examples of links between personality variables, scores on specific cognitive tests, specific neural substrates shared with executive functions, and life

outcomes theoretically reflecting executive function. However, relatively few studies (e.g. Murdock, Oddi & Bridgett, 2013; Unsworth et al., 2009) have evaluated relationships between broad personality taxonomies and multiple executive function constructs. Furthermore, evidence from such studies has been equivocal, with inconsistent findings. We currently lack a well-replicated account of relationships between these aspects of individual differences. Accordingly, the studies reported here will consider the role personality plays, but no specific hypotheses are advanced regarding links between personality and objectively measured executive functions.

Aims

This project set out to assess the relationship of self-reported executive problems with both objectively-measured executive function and with self-report measures of aspects of personality. Multiple studies and measures were used, to test the extent to which any relationships found were robust and replicable across different methodologies and assessment instruments.

Three correlational studies are reported in this paper. Study 1 is a very large web-based study examining links between Webexec scores and personality. Using a well-established personality taxonomy (Costa & McCrae's 1992 Five Factor Model), this allows examination of the extent to which self-reported problems with executive function are associated with indices of five major dimensions of personality.

Study 2 is again a web-based study measuring self-reported problems using Webexec. It also adds an objective measure of executive function (updating / monitoring Working Memory) to test whether self-reported problems are associated with actual performance on a cognitive test. Inclusion once again of personality scales allows replication of tests of the extent to which personality influences self-reports, and exploration of links with cognitive performance.

Study 3 is a laboratory-based study, involving administration of two different self-report measures of executive problems; multiple objective measures of two distinct executive functions (Cognitive Flexibility and Working Memory); and two different measures of the Five Factor personality dimensions. This allows direct and conceptual replications of findings linking self-report measures of executive problems and personality from Studies 1 and 2; a finer-grained evaluation of links between self-report measures and executive tests; and further exploration of links between personality and executive tests.

The main focus of the project is on understanding what drives responses to self-report measures of executive problems. Do they reflect real executive dysfunction, or personality characteristics?

Study 1

Study 1 was exploratory in nature, with no specific hypotheses being advanced. The primary aim was to evaluate the extent to which responses to the Webexec questionnaire were associated with personality.

Method

Materials. Study 1 was conducted wholly online. Ethical approval for the study came from the host University's Psychology Research Ethics Committee. Data were acquired using an established personality testing website, www.personalitytest.org.uk. This hosts an online Five Factor personality inventory providing indices of Extraversion, Neuroticism, Openness to Experience, Agreeableness and Conscientiousness, as operationalized in the Five Factor Model of Costa and McCrae (1992). This 41-item inventory was derived from an International Personality Item Pool measure (IPIP; Goldberg, 1999) that correlates well with Costa and McCrae's domains. It has been validated for use on the internet (Buchanan, Johnson & Goldberg,

2005). In this inventory, Extraversion is assessed by 9 items such as ‘‘Am skilled in handling social situations’’ (Buchanan et al., 2005, report Cronbach’s alpha reliability of .88). Agreeableness is assessed by 7 items such as ‘‘Have a good word for everyone’’ (alpha = .76). Conscientiousness is assessed by 10 items such as ‘‘Pay attention to details’’ (alpha = .84). Neuroticism is assessed by 8 items such as ‘‘Have frequent mood swings’’ (alpha = .83). Openness to Experience is assessed by 7 items such as ‘‘Believe in the importance of art’’ (alpha = .74). Participants are asked to rate the accuracy of statements about their typical or general behaviour on a 5-point scale ranging from 1 ‘‘very inaccurate’’ to 5 ‘‘very accurate’’. The website has existed for a number of years, and attracted over a thousand users per week at the time the study was conducted. No attempt is made to recruit respondents or otherwise attract them to the site--they are referred by other sites or find it through search engines. Many complete the test as part of some class, being asked to do so by their teacher or professor.

For the purposes of the current study, additional questions were added to the site. These comprised the Webexec measure described by Buchanan et al. (2010), which consists of six items tapping experience of problems with different aspects of executive function (keeping attention on a particular task; concentrating on a task; carrying out more than one task at a time; losing one’s train of thought; seeing tasks through to completion; controlling impulsivity). Participants respond to each of these on a 4-point scale anchored at 1 (no problems experienced) and 4 (a great many problems experienced). A single score is computed by summing responses to the six items: the instrument was designed as a general measure tapping multiple aspects of executive function, rather than focusing on any specific function in particular. Participants progressed through the site in the manner described below.

Procedure. Participants first saw a page with a brief description of the inventory,

explaining that it was part of a research project. On clicking a button to indicate that they consented to participate, they then saw a second page with brief instructions and the 41 items of the inventory. Radio button response formats on a 5-point scale ('Very Inaccurate - Very Accurate') were used for the personality items, while all the others used drop-down menus. The other items comprised age group (in 5-year increments); current location (a comprehensive list of nations); gender; highest level of education; main occupational status. These were followed by the six Webexec items as described in Buchanan et al. (2010). Participants were informed that these were included for research purposes, and that they would not receive feedback on their scores on them. Following this, participants were 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). After responding to the items, respondents were then asked to click on another button to submit their data. Those who had completed all the personality items then saw a debriefing page thanking them for their participation, and providing their scores on each of the scales (those who had not were sent back to complete the missing items). In addition, they were given information to help interpret the scores, including a brief description of the meaning of each of the scales, and normative information about their scores relative (top third, middle, bottom third) to others who had completed the inventory to date. Links were provided to contact the researcher, and to information about personality research elsewhere on the internet.

Data screening and processing. In the period from 26th March 2009 to 9th December 2010, 76,177 data submissions were recorded. Data were screened in three ways. First, the file was examined for multiple submissions: instances where a person participated twice, either on purpose or accidentally by clicking the submit button more than once. This was done using a

unique identification number randomly assigned at the first page of the website. Using the SPSS 20 “Identify Duplicate Cases" dialogue, 7983 cases were identified as having identification numbers duplicating earlier records in the file. These were removed, leaving 68,194 cases. Second, any records where the respondent had not indicated that they gave consent for their data to be used were deleted. This left 51,004 cases.

The file was then examined for unrealistic combinations of demographic data (e.g. people claiming to be children with doctoral degrees) that might indicate mischievous or careless responding (sometimes called 'extreme data entry'). Twenty-nine people claiming an age group below the 11-15 category were removed from the sample, as were 6 people claiming to be in the 11-15 age group but giving an occupational status of 'retired'. Anyone in the 11-15 age group claiming post-school education (17) and anyone in the 16-20 group claiming 'some postgraduate' or higher education (44) was also removed. Then, anybody reporting their age as below 16 years (1153) was removed from the sample due to ethical concerns about whether they could be considered to have given valid consent. Finally, procedures were implemented to identify and remove any individuals who had responded to the personality inventory in a potentially inauthentic manner, by simply selecting the middle response option of each item (21 responses) or by only answering with the extreme response options (112 responses).

Following these checks, 49,398 responses remained in the datafile. All further analyses are based on these. All participants had answered all the personality questions (the website ensures this). However, 2,328 left at least one executive question unanswered, so analyses involving the Webexec total score are restricted to the 47,070 usable respondents who answered all six executive items (61.8% of all data submissions initially recorded).

Participants. Demographic characteristics for the 49,398 participants, who came from a

total of 161 countries around the world, are shown in Table 1. The data suggest that the sample is strongly biased towards young North American women of relatively high ability, motivated to complete the inventory as part of a class requirement.

Results and Discussion

Descriptive statistics for all Study 1 variables are shown in Table 2. Links between personality variables and self-reported executive problems were assessed using Pearson's correlations. This was followed by a multiple regression analysis to evaluate the independent effect of each personality variable on Webexec scores.

Correlations between Webexec and personality variables are shown in Table 2. Given the sample size, all correlations are statistically significantly different from zero so examination of the magnitude of the effect is more instructive. Table 2 shows that self-reported executive problems had substantive correlations with personality; Conscientiousness (negative) and Neuroticism in particular. These correlations support the notion that personality may influence responses to self-report measures of cognition.

A multiple regression analysis with simultaneous entry of all predictors was performed to explore the independent effects of personality variables on Webexec scores. Table 3 demonstrates that Extraversion, Agreeableness, Conscientiousness and Neuroticism had independent effects on Webexec scores, while Openness to Experience did not. Again, the very large sample size mandates consideration of effect size, and of whether relationships are sufficiently large to be important (as opposed to being statistically significant but trivial). Ferguson (2009) suggests a threshold of β=.2 for an effect size to be considered 'practically significant' in social science data, while a β of .5 may be considered a moderate effect. In the current dataset, Conscientiousness

and Neuroticism meet the threshold of practical significance, while Extraversion and Agreeableness do not.

The link with Neuroticism is consistent with typical observations that more neurotic people are likely to report more problems of various sorts. For example, Bruce, Bruce, Hancock & Lynch, 2010, found that Multiple Sclerosis patients higher on Neuroticism reported more problems with memory. Jang, Haley, Mortimer and Graves (2002) found that older adults scoring higher on Neuroticism were more likely to report mobility problems.

Beyond this, the correlation with Conscientiousness indicates that people who report experiencing more executive problems tend to see themselves as less conscientious. This may reflect something more than a response bias. The Conscientiousness scale used here includes items such as ‘pay attention to details’, ‘get chores done right away’, ‘find it difficult to get down to work’, and ‘am always prepared’. It is possible that a person with poorer executive function may really experience problems in areas such as these—and self-awareness of that fact could well lead to lower scores on a self-report measure of Conscientiousness. Furthermore, there are suggestions that Conscientiousness may be related to multiple executive functions (Murdock et al., 2013), and that Conscientiousness and executive function may be underpinned by the same neural substrates (e.g. DeYoung et at al, 2010) and have multiple important relationships (Williams et al., 2009).

An alternative explanation for correlations between the personality variables and self-report of executive problems might involve item overlap between the measures (that is, the questions in Webexec and the personality scales are asking about the same things). While Neuroticism, with items like "Often feel blue", "Dislike myself" and "Have frequent mood swings", has little if any overlap with the content of the Webexec scale, there are similarities as

noted above between the experiences and behaviors queried by the Conscientiousness and Webexec scales. To resolve this question, measures of one or the other construct that do not rely on self-reports would be required.

These findings suggest that self-reports of executive problems may be related to personality in multiple ways. A question then arises as to whether self-report scales actually measure anything beyond personality. If personality were controlled for, would scores on measures such as Webexec still correlate with objective measures of executive function, as reported by Buchanan et al. (2010)?

Study 2

Study 2 set out to test whether Webexec scores correlated with an objective cognitive task when personality was controlled for. The Webexec questionnaire was originally validated in a laboratory setting (Buchanan et al., 2010), where it was shown that scores were associated with performance on a number of cognitive tests believed to be influenced by executive function. These included a reverse digit span task, which Webexec scores correlated significantly with (r=-.42, n=77, p=.000). This correlation is in the range between 'medium' and 'large' effect sizes as discussed by Cohen (1992), and exceeds the threshold Ferguson (2009) suggests for effects to be regarded as 'practically significant'.

Reverse digit span tasks (named in e.g. the Wechsler tests as “Digit Span Backward”) require participants to repeat increasingly longer strings of digits, but in reverse of the presentation order. This requires use of one of the core executive functions described by Diamond (2013); Working Memory. Davis, Marra, Najafzadeh, and Liu-Ambrose (2010) argue that performance on Digit Span Backward requires use of the central executive component of working memory. Consistent with this, Tamez, Myerson, Morris, White, Baum and Connor

(2011) note that Digit Span, and particularly Backward span, is commonly used as a measure of executive function in stroke patients. Wechsler (2008) notes that Digit Span Backward involves use of working memory, as well as mental manipulation and transformation of information (along with visuospatial imaging). Thus, a reverse digit span task is used here as a measure of executive function, and specifically the Working Memory component.

It has been shown (Tractenberg & Freas, 2007) that a text-based digit span task can be implemented as a web-based test. Thus, Study 2 examined the relationships between self-reported executive problems, reverse digit span, and personality in a web-based format. It was hypothesised that there would be a negative correlation between reverse digit span scores, and executive problems self-reported via Webexec (Hypothesis 1). It was further hypothesized that this relationship would remain when personality variables were controlled for (Hypothesis 2).

Method

Materials. Study 2 was conducted wholly online. Ethical approval for the study came from the host University's Psychology Research Ethics Committee. The personality inventory (Buchanan et al., 2005), demographic items, and Webexec questionnaire (Buchanan et al., 2010) were all as described in Study 1. The digit span task was implemented online using the Qualtrics online survey platform, which permits timed presentation of stimuli. It was administered using the instructions described by Tractenberg and Freas (2007) who validated an online version of the task against an offline version. Rather than the randomized stimuli and staircase algorithm used by Tractenberg and Freas, the current study used the traditional format (as used for example in the WAIS-III: Wechsler, 1997), where strings of digits are presented in order of increasing size. This was done to maximise comparability of current findings with existing literature on digit span tasks and executive function. Precedents for this come from work by Logie and

Maylor (2009) and Germine, Nakayama, Duchaine, Chabris, Chatterjee, and Wilmer (2012), who report successful use of an online digit span task adapted from the WAIS-III.

Prior to each digit span trial, the participant saw the messages “Ready” and then “Go” for 0.5 seconds each. They were then shown a sequence of digits between 1 and 9, one at a time, for one second each. At the end of the sequence, they saw a cue reading either ‘FORWARD’ or ‘BACKWARD’. They then typed the numbers they had seen, in sequence either forward or backward as directed, putting an ‘x’ in the place of any number they could not recall. The trials began with forwards recall, starting with two digits and rising to a maximum of 9 digits. There were two trials for each sequence length. The forward trials terminated when either all had been successfully completed, or the participant had failed twice at a given sequence length. The reverse recall trials that followed these again began with two digits and rose to a maximum sequence length of 8 digits. For both forward and reverse trials, the score was the participant’s total number of correct responses across that set of trials.

Procedure. Participants were recruited via the same personality testing website www.personalitytest.org.uk as used in Study 1, though without the Webexec items. On completing the inventory and receiving their feedback, website users who had indicated their data were valid for analysis were shown a message inviting their participation in Study 2. Clicking a link in this recruitment message led them to a participant information / consent page for Study 2. This told them what they would be asked to do, and indicated that if they participated the personality data they had already supplied would be used in the analysis. Those giving consent then saw the six items of the Webexec scale. Having completed these, they then progressed to the digit span task. On termination of that task, debriefing pages gave them more information about the study.

Data screening and processing. The online personality inventory was completed 9023 times during data collection for this project. In 6401 of these instances, respondents indicated their data could be used for analysis, and were subsequently invited to take part in the digit span study. Of the 6401 invitations, 671 individuals accessed the data collection website and saw at least the first question. Of these 671 respondents, 354 completed the study and gave consent for their data to be used. A further 7 completed it but did not give consent. The remaining 310 individuals did not complete the study, leaving 354 who completed it and gave consent for their data to be used at both the beginning and end of the procedure. To assure data quality, Qualtrics' proprietary techniques to prevent multiple participation were used. Further, the datafile was examined for implausible combinations of demographic data (e.g. people claiming to be very young but have doctoral degrees). None were found. Next, eight respondents reporting their age as below 16 years were excluded due to ethical concerns about their capacity to give valid consent. Finally, procedures were implemented to identify and exclude from analysis any respondents who had responded to the personality inventory in a potentially inauthentic manner, by simply selecting the middle response option of each item (no respondents) or by only answering with the extreme response options (one respondent).

Participants. Demographic details of the 345 participants are shown in Table 1. The sample is in general slightly older and better educated than that in Study 1, but is otherwise broadly comparable.

Results

Descriptive statistics for all Study 2 variables are shown in Table 4. Links between personality variables, self-reported executive problems and digit span scores were assessed using Pearson's correlations. This was followed by multiple regression analyses to evaluate the

independent effect of each personality variable on Webexec and digit span scores.

Table 4 indicates that in this sample, internal reliability for the Openness to Experience scores just failed to meet the conventional .7 threshold for Cronbach's alpha. While the scale has been retained in the analysis, findings obtained with it must be regarded as speculative.

Pearson's correlations between Webexec, personality and digit span scores were computed and are shown in Table 5. Webexec scores did not correlate with performance on the reverse digit span task, so the findings are not consistent with Hypothesis 1. This further implies that Hypothesis 2 – that a relationship would persist when controlling for personality – cannot be true. In fact a partial correlation between Webexec and reverse digit span score, controlling for the personality variables Webexec correlated with (Extraversion, Agreeableness, Conscientiousness and Neuroticism), was not significantly different from zero ($r_{partial}$=-.067, df=339, p=.22).

In a final exploratory analysis, multiple regressions (with straightforward simultaneous entry of all predictor variables) were performed to examine the extent to which personality variables predicted variance in Webexec and digit span scores (Table 6). This showed that Webexec scores were significantly predicted by Neuroticism and low Conscientiousness, but not other personality variables. Effect sizes were broadly consistent with those in Study 1. Both forward and reverse digit span scores were predicted by Openness to Experience, while forward span was also predicted by Agreeableness (for reverse digit span, Agreeableness was not a significant predictor, p=.077).

Discussion

The present findings vary from Buchanan et al.'s (2010) demonstration that Webexec scores correlated significantly with reverse digit span. A parsimonious explanation for this is

that it is a simple failure to replicate. However, the present study had ample power to detect an effect of the magnitude reported by Buchanan et al. This suggests the possibility that the original finding may have been a false positive (though this is somewhat militated against by the fact that Buchanan et al. report a second study showing links between Webexec scores and use of cannabis, which it has been argued impacts on executive function – McHale & Hunt, 2008).

A third possibility is that performance on the reverse digit span task may be influenced by different factors when administered in a text-based format in an online, unsupervised environment than when administered normally under standard conditions. This view is supported somewhat by the similar finding of Friedman-Kraus et al. (2014) that the Webexec measure did not correlate with objective executive tests (including reverse letter span) administered online. The relationship shown in Table 6 between Openness to Experience and both forward and reverse digit span, and between Agreeableness and forward span, further hints that non-cognitive factors may account for some of the variance in online digit span performance. Examining the frequency distributions of forward and reverse digit span scores, it was notable that an unexpectedly high proportion of respondents had zero scores (12.2% for forward span, 13% for reverse span). This suggests they either had a problem with the task, or perhaps did not make a serious attempt to complete it. When analyses are re-run excluding any participant with a zero score on reverse digit span, the correlation between Webexec and reverse span only rises very slightly from r=-.08 to r=-.10, and remains non-significant (p=.07). This implies that the high proportion of zero-scoring respondents does not present a problem for the present analysis, or explain the failure to replicate Buchanan et al.'s (2010) findings. However, it does raise questions about how people interact with web-based cognitive tests. This would benefit from exploration in future research.

Beyond this, the pattern of correlations between personality and Webexec scores is the same as in Study 1, with substantive correlations with Conscientiousness (negative) and Neuroticism (positive). As in Study 1, only these two variables exceed Ferguson's (2009) threshold (r=.2) for a practically significant association. Furthermore, the regression analysis (Table 6) indicates that it is only these two variables that are associated with Webexec scores. The lack of associations with the other personality dimensions gives some confidence that the relationships are genuine and not artifacts of, say, a general yea-saying bias.

It is clear from Studies 1 and 2 that there are relationships between personality and executive problems as self-reported using this measure. This raises the question of whether the finding is idiosyncratic to these questionnaires, or something that would apply to other measures of executive problems and personality as well.

Study 3

Study 3 set out to test these possibilities in a controlled laboratory setting. Previous research on associations between self-report executive measures with objective tests on non-clinical participants in laboratory settings is equivocal. For example, while Laws et al. (2008) found that the self-report DEX did not correlate with executive measures, Buchanan et al. (2010) found that Webexec did. The first main aim of Study 3 was to extend that work by examining whether performance on these measures correlated with performance on a range of executive tasks. Study 2 looked at a single task – reverse digit span, which addresses working memory. However, the multiplicity of abilities and variables that are commonly described as executive functions (e.g. Salthouse, 2005) implies that a better design would incorporate multiple measures of multiple executive functions. The second aim was to further explore the role of personality as a determinant of self-reports of executive problems. Again, using multiple measures of the

personality constructs Study 2 found to affect self-reports of problems would increase confidence in the generalizability of findings.

It was hypothesized that self-reports of executive problems, assessed via Webexec and DEX, would be associated negatively with several objective measures of executive function: reverse Digit Span (Hypothesis 1), Trail Making Test Part B (Hypothesis 2), Controlled Oral Word Association test (Hypothesis 3) and Semantic Fluency (Hypothesis 4). It was hypothesized that self-reports of executive problems would correlate positively with Neuroticism (Hypothesis 5) and negatively with Conscientiousness (Hypothesis 6). It was further hypothesized that relationships between Webexec, DEX and the objective measures would persist when personality variables were controlled for (Hypothesis 7).

Method

Study 3 was conducted in a face-to-face, laboratory setting. Ethical approval for the study came from the host University's Psychology Research Ethics Committee. In a correlational design, participants completed measures of self-reported executive problems, personality, and objective cognitive tests believed to index executive functions.

Materials. Self-reported experience of executive problems was assessed using two different questionnaire measures. The first was the 6-item Webexec measure used in Studies 1 and 2. This was administered in a paper-and-pencil format for this study, rather than online. The second measure used was the self-report DEX scale drawn from the Behavioural Assessment of the Dysexecutive Syndrome battery (BADS; Wilson et al., 1996). This is a 20-item Likert-style questionnaire that can be administered in informant or self-report formats, and gives an overall score that assesses the frequency of behaviors associated with executive problems.

Personality was assessed using two established inventories that tap the domain constructs of the Five Factor Model. The first was the NEO-FFI (Costa & McCrae, 1992), a widely used short form of the ubiquitous NEO-PI-R, which is a standard measure of these constructs. The second was the 41-item IPIP questionnaire, as used in Studies 1 and 2, though in paper and pencil format on this occasion. Both measures provide indices of Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness to Experience.

Cognitive function was assessed using sets of tests thought to be influenced by two distinct executive processes. The processes addressed were cognitive flexibility and working memory.

The Trail Making Test (TMT; Reitan, 1955) Part B is believed to measure cognitive flexibility (Arbuthnot and Frank, 2000; Davis et al., 2010; Strauss, Sherman and Spreen, 2006). It comprises circles that are either numbered (1-13) or lettered (A-L). Participants draw lines connecting them in ascending order, but switching between letters and numbers (1, A, 2, B and so on).

The Controlled Oral Word Association (COWA) Test (Benton, Hamsher & Sivan, 1994) also measures cognitive flexibility (Diamond, 2013; Johnco et al., 2014). Often described as a measure of Phonemic Fluency or Verbal Fluency, in this task participants must produce as many examples as they can of words beginning with a specific letter. In the current study, the 'FAS' version of the test was administered, giving participants one minute each to generate words beginning with the letters F, A and S. The COWA was supplemented by a Semantic (or Category) Fluency task, where participants were asked to name as many animals as they could in one minute. Again, this measures cognitive flexibility (Diamond, 2013).

The final cognitive measure was the Digit Span test from the WAIS-IV, administered according to the standard instructions (Wechsler, 2008). This was similar in principle to the digit span task used in Study 2, though here it was presented in oral format. It comprises forward and reverse digit span tasks and a Sequencing task, in which participants are required to repeat strings of digits having arranged them into the correct numerical order. Digits Backwards and Digits Sequencing both tap working memory (Davis et al., 2010; Diamond, 2013; Wechsler, 2008).

Procedure. Participants were tested singly in quiet cubicles. They first completed the Webexec and self-report DEX measures. They then did the cognitive tests: WAIS IV Digit Span (forwards, backwards and sequencing); Trail Making Test (A and B); Controlled Oral Word Association (Phonemic Fluency) and Semantic Fluency tests. A stopwatch was used for timing. They then completed the NEO-FFI and IPIP personality inventories, were thanked and debriefed.

Data screening and processing. One hundred and four people participated. Initial inspection of the data indicated that a number of cases (11) had missing data, particularly on individual personality items. These were retained in the sample, but excluded from any analyses involving variables on which they had missing scores. Examining score distributions for personality variables suggested one participant, an 18 year-old male, had responded in a possibly inauthentic manner, endorsing only the extremes or midpoints of all items on both personality measures. He was excluded from the analyses that follow.

Participants. The sample comprised 103 psychology students recruited through a research participation scheme in return for course credit. Demographic data are shown in Table 1. In comparison with Studies 1 and 2, the sample is again young and well-educated. However, it is

more homogenous in terms of international makeup, route to participation, and other participant characteristics. Men are even more under-represented than in Studies 1 and 2.

Results

Descriptive statistics for all Study 3 variables are shown in Table 7. Links between self-reported executive problems, personality variables, and executive tests were assessed using Pearson's correlations. This was followed by a multiple regression analysis to evaluate the independent effects of each self-report measure and of each personality variable on composite measures of executive functions.

In this sample, datasets from a number of the tests failed to show acceptable levels of internal consistency, with alphas below .7. While these variables have been retained in the analysis for completeness, findings obtained with them should be regarded as speculative. However, there were substantial correlations between the IPIP and corresponding NEO-FFI personality measures, even in cases where one of the pair being examined had an alpha of below .7. Pearson's *r* correlations of .73 (Neuroticism), .63 (Extraversion), .69 (Openness to Experience), .66 (Agreeableness), and .85 (Conscientiousness) were all significantly different from zero at the p=.000 level, providing some evidence that the inventories are measuring similar constructs, and some reassurance that the measures are providing usable data. The same is true of the correlation between DEX and Webexec (r=.58, p=.000).

While the correlations indicate that variance is shared between these conceptually-related pairs of variables, there is also substantive variance unique to each. For example, the correlation between DEX and Webexec indicates that only 34% of the variance in one is accounted for by the other. This vindicates the decision to use multiple measures, in case findings were idiosyncratic to the non-shared elements of the variance in these measures.

Table 8 indicates that neither DEX nor Webexec correlated significantly with any of the objective cognitive measures, despite the fact that the study was sufficiently powered to detect a correlation of the magnitude reported by Buchanan et al. (2010). However, both DEX and Webexec correlated significantly with neuroticism, low agreeableness and low conscientiousness.

Of the five cognitive tests used, Digits Forward and Trails A are not generally considered to map onto executive functions (they are included in Tables 6 and 8 for completeness). However, Digits Backwards and Digits Sequencing relate to working memory. Trails B, Phonemic Fluency and Semantic Fluency relate to cognitive flexibility. For the purposes of further exploratory analyses, each of these two groups of measures was combined into an overall score by converting raw scores to *z* scores then computing the mean for each group. This resulted in two composite scores, Cognitive Flexibility and Working Memory.

For each of the five personality constructs measured (N, E, O, A, C) two indices were available, drawn respectively from the NEO-FFI and the IPIP measures. Again, composite measures were created by transforming to *z* scores then computing the mean for each pair of variables. This resulted in five composite personality scores.

For each of the composite EF variables, a regression analysis was carried out examining the effect of self-reported executive problems on EF while controlling for personality. Separate analyses were performed for DEX and Webexec. While they share variance, for applied measurement purposes the specific relationship between each and the EF variables is of interest. These are shown in the top and bottom halves of Table 9 respectively. Neither DEX nor Webexec, nor any of the composite personality variables, predicted scores on either Working

Memory or Cognitive Flexibility.

Discussion

Of the hypotheses proposed, findings were only consistent with the predictions that self-reported executive problems would correlate with neuroticism (H5) and conscientiousness (H6). No evidence was found that self-report measures of executive problems correlated with any of the objective measures (H1, H2, H3, H4, H7). Thus, in the present sample these two self-report measures of executive problems do not appear to reflect actual cognitive performance. Rather, they have substantive correlations with personality variables. The fact that both executive self-report measures (DEX and Webexec) have the same pattern of correlations with both sets of personality variables (NEO-FFI and IPIP) indicates it is the core construct addressed by each measure that is important, rather than variance idiosyncratic to each.

General Discussion

The general finding across Studies 1-3 is that at least two self-report measures of executive problems fail to correlate with objectively measured executive function in non-clinical (largely student) samples. Instead, self-reports of problems are associated with the personality dimensions neuroticism and conscientiousness (negatively). The associations are stable across all three studies (the only exception being the non-significant correlation between IPIP Neuroticism and Webexec – which had low reliability – in Study 3). The effect sizes are well above the threshold Ferguson (2009) describes as being ‘practically significant’. Across all studies, the mean correlation between both indices of neuroticism and both indices of self-reported executive problems was r=.31 (a medium effect size according to Cohen, 1992); the corresponding value for conscientiousness was r=-.52 (classified by Cohen, 1992, as a large effect size). While generally comprising well-educated participants, the samples differed in how they were recruited

and tested, and in demographic characteristics such as sex, nationality and age profile. The fact that key relationships were replicated across the somewhat disparate samples permits some confidence regarding their robustness and generalizability. These findings may have practical implications, and also point to the need for further research in a number of areas.

Predictors of Self-reported Executive Problems

If these self-report scales do not reflect real executive dysfunction in normal samples, then what influences responses to them? One possibility is that they largely reflect personality.

Everybody makes mistakes. Minor cognitive failures in everyday life may be commonplace in the general population (e.g. Chan, 2001), while not reaching levels of severity or frequency that would be viewed as a clinically significant problem or measurable using objective tests. Awareness of these lapses may be associated with low self-reported conscientiousness. One possible mechanism is that awareness of everyday executive failures (e.g. realizing one has failed to plan ahead for something) influences individuals' views of themselves as being conscientious.

Neuroticism is known to be a predictor of self-reports of various types of problem (e.g. Bruce et al., 2010; Jang et al. 2002; Buchanan et al., 2005). As noted earlier, one possibility is that the substantive associations reported here reflect the tendency of less emotionally stable people to report more problems in general. However, neuroticism's relationship with general negative affectivity raises the possibility of another mechanism linking these constructs.

Heffernan and O'Neill (2013) found that both anxiety and depression were associated with self-reported executive problems (Webexec) scores in a student sample. Work by Johnco et al. (2014) suggests that self-ratings of cognitive flexibility correlate with anxiety and depression in a non-clinical sample of older adults. These results strike a chord with Gerstorf et al.'s (2008)

finding that trait anxiety, depression and general negative affect were associated with DEX scores, and their suggestion that in community dwelling adults the DEX "does not tap directly tap into executive functioning, but rather indexes aspects of unpleasant affect that may accompany the experience of everyday executive functioning problems." (p. 442). It is possible that when more neurotic people become aware of their everyday mistakes, they experience more negative affect, and this leads them to rate themselves as having more problems. In the absence of controlled experimentation, however, this must remain a tentative hypothesis.

Study 3 found that Agreeableness has a weak negative association with self-reported executive problems on the DEX and possibly also Webexec (Table 8). However, Agreeableness did not meaningfully predict Webexec scores in Studies 1 and 2. In Study 3, a partial correlation between IPIP Agreeableness and Webexec, controlling for DEX scores (with which Webexec correlates) was not significantly different from zero ($r_{partial}$=-.11, df=98, p=.27). This suggests that Agreeableness may be associated with DEX scores, but the relationship may not be generalizable to the other self-report measure.

Another possible explanation for links between personality and self-reported problems, as noted in Study 1, is that there is overlap between the content of the scales (they are asking about the same things, or have similar items). This explanation is however not compelling, because multiple (but not all) personality variables correlated with the self-report scales, and because two different inventories measuring the personality constructs correlated with two different scales measuring the experience of executive problems.

Personality and Executive Function

While the studies reported here were not designed to test hypotheses about links between personality and objectively measured executive functions, exploratory analyses were performed. However, few links were found.

In Study 3, the only significant correlation between a personality variable and a cognitive test was between IPIP Neuroticism and Digits Backward. This relationship was not significant for NEO-FFI Neuroticism and Digits Backward; nor was the correlation between IPIP Neuroticism and Digits Backward significant in Study 2. Furthermore, in the Study 3 regression analyses, the composite measure of Neuroticism did not predict the composite measure of Working Memory (which incorporated Digits Backward). None of this inspires confidence that the finding reflects a genuine relationship rather than a false positive error.

In Study 2, Openness to Experience predicted both Digits Forward (not an executive task) and Digits Backwards. However, it did not correlate with either of these variables in Study 3. Unsworth et al. (2009) found a link between Openness to Experience and Fluency. Murdock et al. (2013) reported that Openness was associated with both cognitive flexibility and updating / monitoring (which in their implementation was based on similar fluency tasks to the cognitive flexibility variable in the current Study 3). Taken together, their results could be interpreted as suggesting that Openness to Experience should be related to cognitive flexibility. However, here the relationship was with a different function, and Openness did not correlate with either of the raw fluency tasks or predict the composite Cognitive Flexibility variable they contributed to. If the current findings tell us anything about Openness, it is likely to be something to do with factors affecting responses to internet-based cognitive tests versus those administered in person, rather than anything to do with executive function. This is especially the case as the correlations in Study 2 were with both an executive and a non-executive test.

Overall, the present data say little about relationships between personality and executive functions. In particular, findings from other research (e.g. Murdock et al., 2013; Unsworth et al., 2009) were not replicated.

Limitations

The program of research reported here has a number of limitations. One is the narrow range of executive functions measured. Within the literature the definition of ‘executive function’ and the abilities it includes has been somewhat varied (Miyake & Friedman, 2012, note that it has been difficult both to define and to measure). While Studies 2 and 3 set out to measure distinct core executive functions, of the three outlined by Diamond (2013), only Working Memory and Cognitive Flexibility were directly addressed. There was no measure of Inhibitory Control (although Miyake & Friedman, 2012, suggest that inhibition may be a factor that overlaps with all other aspects of executive function). Moreover, the three-function framework advocated by Diamond may not be exhaustive – Miyake and Freidman (2012) for example note the possibility of other functions such as Dual Tasking. Thus, the tests used did not measure all processes that could be classed as 'executive functions'. Furthermore, for those functions that were measured, a limited range of tests was employed. There are a variety of measures of Working Memory and Cognitive Flexibility in addition to those used here (Diamond, 2013; Miyake & Friedman, 2012). It may therefore be possible that the self-report questionnaires tap an aspect of executive function not measured in Studies 2 and 3. If questionnaires such as Webexec and DEX seek to assess executive function in general, then the present studies’ focus on a limited subset of executive function measures may contribute to the failure to find expected relationships.

The large sample sizes available in Studies 1 and 2 may also be seen as an issue: given the very high power available, even trivial correlations will emerge as significantly different from zero. This necessitates the consideration of effect sizes, and benchmarks of what effect size should be considered meaningful, when evaluating findings. The converse is true of Study 3. The sample in that study was sufficiently large to reliably detect effects of the size found in Studies 1 and 2 (Cohen, 1992), but it may be that more subtle relationships between variables were not detected (although that then raises the question as to whether such small effects would be practically meaningful).

It should be noted that the present findings were obtained from relatively high- performing (largely student) samples. This group is less likely than many others to be experiencing significant enduring cognitive impairment. Tombaugh (2004) presents a range of norms for the Trail Making Test. For older educated samples more likely to be experiencing cognitive impairment, the time taken to complete Trails A and B was considerably higher than found in Study 3. For example, in a group with 12+ years of education, aged 75-79 years, Tombaugh reports mean Trails B time as over 100 seconds – double the time obtained with the present sample. Restriction of range in cognitive performance would attenuate correlations between the self-report and objective measures in the present study. Furthermore, neuropsychological measures are not generally created to assess normal range variability in the abilities of higher functioning populations. While such tests often are used with non-clinical groups, and there is demonstrable variability in the scores obtained (e.g. in the present samples), they may lack sensitivity when it comes to measuring subtle differences in higher-level cognitive performance. Questions might also be raised about the accuracy with which high-functioning groups report their experience of executive failures. They may overestimate the problems they

experience, in comparison to participants with higher levels of impairment. It might therefore be expected that self-report measures would have more utility among moderately impaired groups. Such individuals might have sufficient problems that their reports of experiences did map on to real cognitive failures, but not so many that their ability to accurately report them was overly affected.

An outstanding question is whether the relationships between personality and self-reported cognitive status are unique to these measures, or are more widely applicable to other self-report measures of cognitive function or problems. If the latter, there are implications for how such measures are used in research. However, replication of this work—including conceptual replication with other measures—is required before such conclusions can be drawn.

Conclusions

In summary, the studies reported in this paper suggest that responses to two self-report measures of executive problems are driven by personality rather than objectively-measurable executive function in normal samples. Laws et al. (2008) found that the DEX did not correlate with executive measures in non-clinical participants. Friedman-Kraus et al. (2014) found that Webexec did not correlate with executive measures in teachers tested online. The current findings reinforce both these results. For researchers, there is a strong implication that in non- clinical samples these questionnaires may not be suitable as proxies for actual executive function as measured by neuropsychological tests.

It is important to be clear that the current findings say nothing about the validity of the measures for use with clinical samples, where they may well tap into lived experience of real problems. However, it is possible that self-reports of problems may be influenced by personality in executively-impaired samples too, and this should be a focus of future investigation. This is an

important question, because a number of authors have argued that ratings (including self-ratings) of executive problems have greater ecological validity than cognitive tests, which often do not correlate with real-world problems experienced by various impaired samples (e.g. Barkley & Fischer, 2011; Kamradt et al., 2014). The argument that self-ratings are more ecologically valid may only hold true if those self-ratings do reflect executive problems. If instead they substantially reflect neuroticism or other personality traits, their usefulness may be compromised.

References

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, *22*(4), 518-528. doi:10.1076/1380-3395(200008)22:4;1-0;FT518

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839. doi:10.1038/nrn1201

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*(2), 89-94. doi:10.1111/j.1467-8721.2009.01615.x

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology*, *36*(2), 137-161. doi:10.1080/87565641.2010.549877

Benton, A. L., Hamsher, K. D., & Sivan, A. B. (1994). *Multilingual aphasia examination: manual of instructions*. Iowa City, Iowa.: AJA Assoc.

Bruce, J. M., Bruce, A. S., Hancock, L., & Lynch, S. (2010). Self-reported memory problems in multiple sclerosis: Influence of psychiatric status and normative dissociative experiences. *Archives of Clinical Neuropsychology*, *25*(1), 39-48. doi:10.1093/arclin/acp092

Buchanan, T., Ali, T., Heffernan, T. M., Ling, J., Parrott, A. C., Rodgers, J., & Scholey, A. B. (2005). Nonequivalence of on-line and paper-and-pencil psychological tests: The case of the prospective memory questionnaire. *Behavior Research Methods*, *37*(1), 148-154. doi:10.3758/BF03206409

Buchanan, T., Heffernan, T. M., Parrott, A. C., Ling, J., Rodgers, J., & Scholey, A. B. (2010). A short self-report measure of problems with executive function suitable for administration

via the Internet. *Behavior Research Methods*, *42*(3), 709-714. doi:10.3758/BRM.42.3.709

Buchanan, T., Johnson, J. A., & Goldberg, L. R. (2005). Implementing a Five-Factor Personality Inventory for Use on the Internet. *European Journal of Psychological Assessment*, *21*(2), 115-127. doi:10.1027/1015-5759.21.2.115

Chan, R. C. K. (2001). Dysexecutive symptoms among a non-clinical sample: A study with the use of the Dysexecutive Questionnaire. *British Journal of Psychology*, *92*(3), 551 - 565. doi:10.1348/000712601162338

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155 - 159. doi:10.1037/0033-2909.112.1.155

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Davis, J. C., Marra, C. A., Najafzadeh, M., & Liu-Ambrose, T. (2010). The independent contribution of executive functions to health related quality of life in older women. *BMC geriatrics*, *10*(1), 16. doi:10.1186/1471-2318-10-16

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science*, *21*(6), 820-828. doi:10.1177/0956797610370159

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168. doi:10.1146/annurev-psych-113011-143750

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532-538. doi:10.1037/a0015808

Friedman-Krauss, A. H., Raver, C. C., Neuspiel, J. M., & Kinsel, J. (2014). Child Behavior

Problems, Teacher Executive Functions, and Teacher Stress in Head Start Classrooms. *Early Education & Development, 25*(5), 681-702. doi:10.1080/10409289.2013.825190

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857. doi:10.3758/s13423-012-0296-9

Gerstorf, D., Siedlecki, K. L., Tucker-Drob, E. M., & Salthouse, T. A. (2008). Executive dysfunctions across adulthood: measurement properties and correlates of the DEX self-report questionnaire. *Aging, Neuropsychology, and Cognition*, *15*(4), 424-445. doi:10.1080/13825580701640374.

Goldberg, L. R. (1999). 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, & F. O. (Eds.), *Personality Psychology in Europe Vol. 7* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press. Retrieved from http://projects.ori.org/lrg/PDFs_papers/A broad-bandwidth inventory.pdf

Grace, J., & Malloy, P. (2001). *Frontal Systems Behavior Scale: Professional manual.* Lutz, FL: Psychological Assessment Resources.

Heffernan, T. M., & O'Neill, T. S. (2013). Everyday Prospective Memory and Executive Function Deficits Associated with Exposure to Second-Hand Smoke. *Journal of Addiction, 2013*, 1-7. doi:10.1155/2013/160486

Jang, Y., Mortimer, J. A., Haley, W. E., & Graves, A. B. (2002). The role of neuroticism in the association between performance-based and self-reported measures of mobility. *Journal of Aging and Health*, *14*(4), 495-508. doi:10.1177/089826402237180

Johnco, C., Wuthrich, V. M., & Rapee, R. M. (2014). Reliability and validity of two self-report measures of cognitive flexibility. *Psychological Assessment, 26*(4), 1381-1387. doi:10.1037/a0038009

Kamradt, J. M., Ullsperger, J. M., & Nikolas, M. A. (2014). Executive Function Assessment and Adult Attention-Deficit/Hyperactivity Disorder: Tasks Versus Ratings on the Barkley Deficits in Executive Functioning Scale. *Psychological Assessment, 26*(4), 1095-1105. doi:10.1037/pas0000006

Laws, K. R., Patel, D. D., & Tyson, P. J. (2008). Awareness of everyday executive difficulties precede overt executive dysfunction in schizotypal subjects. *Psychiatry Research, 160*(1), 8-14. doi:10.1016/j.psychres.2007.06.004

Logie, R. H., & Maylor, E. A. (2009). An Internet study of prospective memory across adulthood. *Psychology and Aging, 24*(3), 767. doi:10.1037/a0015479

McHale, S., & Hunt, N. (2008). Executive function deficits in short-term abstinent cannabis users. *Human Psychopharmacology: Clinical and Experimental, 23*(5), 409-415. doi:10.1002/hup.941

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi:10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8-14. doi:10.1177/0963721411429458

Murdock, K. W., Oddi, K. B., & Bridgett, D. J. (2013). Cognitive correlates of personality:

Links between executive functioning and the big five personality traits. *Journal of Individual Differences*, *34*(2), 97. doi:10.1027/1614-0001/a000104

Rabbitt, P., Maylor, E., McInnes, L., Bent, N., & Moore, B. (1995). What goods can self‐assessment questionnaires deliver for cognitive gerontology? *Applied Cognitive Psychology*, *9*(7), S127-S152. doi:10.1002/acp.2350090709

Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology*, *19*(5), 393-394. doi:10.1037/h0044509

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, *19*(4), 532. doi:10.1037/0894-4105.19.4.532

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press. Tamez, E., Myerson, J., Morris, L., White, D. A., Baum, C., & Connor, L. T. (2011). Assessing executive abilities following acute stroke with the trail making test and digit span. *Behavioural Neurology*, *24*(3), 177-185. doi:10.3233/BEN-2011-0328

Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203-214. doi:10.1016/S0887-6177(03)00039-8

Tractenberg, R. E., & Freas, C. E. (2007). Benchmarking a computerized test of immediate verbal memory. *Behavior Research Methods*, *39*(4), 863-869. doi:10.3758/BF03192980

Unsworth, N., Miller, J. D., Lakey, C. E., Young, D. L., Meeks, J. T., Campbell, W. K., & Goodie, A. S. (2009). Exploring the relations among executive functions, fluid intelligence, and personality. *Journal of Individual Differences*, *30*(4), 194-200. doi:10.1027/1614-0001.30.4.194

Wechsler, D. (2008). *WAIS-IV UK administration and scoring manual*. London, UK: Pearson.

Williams, P. G., Suchy, Y., & Rau, H. K. (2009). Individual differences in executive functioning: Implications for stress regulation. *Annals of Behavioral Medicine, 37*(2), 126-140. doi:10.1007/s12160-009-9100-0

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

Tables

Table 1

Demographic Data for Studies 1-3

Measure		Study 1	Study 2	Study 3
N		49,398	345	103
Sex				
	Men	17,031 (34.5%)	105 (30.4%)	6 (5.8%)
	Women	30,969 (62.7%)	233 (67.5%)	97 (94.2%)
	Unanswered	1,398 (2.8%)	7 (2.0%)	0 (0%)
Age				
	Modal age group	16-20	16-20	-
	Age range	16-20 to Over 85	16-20 to 66-70	18-41 years
	Mean Age (*SD*)	-	-	20.19 (3.70)
	Unanswered	0 (0%)	0 (0%)	0 (0%)
Location				
	USA	32,531 (65.9%)	167 (48.4%)	0 (0%)
	UK	6,705 (13.6%)	53 (15.4%)	103 (100%)
	Other	9,871 (19.9%)	122 (35.3%)	0 (0%)
	Unanswered	291 (0.6%)	3 (0.9%)	0 (0%)
Route to participation				
	Doing as part of some class	28,753 (58.2%)	126 (36.5%)	0 (0%)
	Found through search engine	10,353 (21.0%)	118 (34.2%)	0 (0%)
	Got link from a friend	3,658 (7.4%)	26 (7.5%)	0 (0%)
	Followed link from another site	3,324 (6.7%)	48 (13.9%)	0 (0%)
	Other	2,568 (5.2%)	25 (7.2%)	103 (100%)
	Unanswered	742 (1.5%)	2 (0.6%)	0 (0%)
Highest level of education				
	Primary Education	1,582 (3.2%)	7 (2.0%)	0 (0%)
	Secondary Education	13,369 (27.1%)	66 (19.1%)	0 (0%)
	Vocational / Technical college	2,429 (4.9%)	24 (7.0%)	0 (0%)
	Some college / University	17.052 (34.5%)	87 (25.2%)	103 (100%)
	College / University Graduate	8,208 (16.6%)	83 (24.1%)	0 (0%)
	Some Postgraduate	3,496 (7.1%)	38 (11.0%)	0 (0%)
	Postgraduate / Professional Degree	3,174 (6.4%)	39 (11.3%)	0 (0%)
	Unanswered	88 (0.2%)	1 (0.3%)	0 (0%)
Occupation				
	Employed for Wages	18,249 (36.9%)	112 (32.5%)	0 (0%)
	Self-employed	2,073 (4.2%)	30 (8.7%)	0 (0%)
	Unemployed	2,544 (5.2%)	23 (6.7%)	0 (0%)
	Home-maker	1,084 (2.2%)	14 (4.1%)	0 (0%)
	Student	23,703 (48.0%)	147 (42.6%)	103 (100%)
	Retired	321 (0.6%)	4 (1.2%)	0 (0%)
	Unable to work	501 (1.0%)	9 (2.6%)	0 (0%)
	Unanswered	923 (1.9%)	6 (1.7%)	0 (0%)

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

Table 2

Descriptive Statistics and Pearson's r Correlations of Personality with Self-Reported Executive Problems (Webexec scale)

					Range			
Variable	*n*	*M*	*SD*	α	Potential	Actual	Skew	Correlation with Webexec
Extraversion	49398	30.46	7.13	.87	9-45	9-45	-0.37	-.11***
Agreeableness	49398	27.39	4.45	.74	7-35	7-35	-0.74	-.20***
Conscientiousness	49398	35.54	7.16	.85	10-50	10-50	-0.33	-.58***
Neuroticism	49398	20.59	6.51	.83	8-40	8-40	0.42	.37***
Openness to Experience	49398	25.28	5.15	.72	7-35	7-35	-0.28	.02***
Webexec	47070	11.06	3.27	.79	6-30	6-24	0.87	

****p<.0005*

Table 3

Regression of Webexec Scores on Personality

Variable	*B*	*SE B*	*β*	*t*
Constant	16.26	0.14		115.59***
Extraversion	0.02	0.00	0.05	11.46***
Agreeableness	0.01	0.00	0.01	3.30**
Conscientiousness	-0.23	0.00	-0.51	-126.64***
Neuroticism	0.10	0.00	0.20	47.83***
Openness to Experience	0.01	0.00	0.01	1.91

Note. $F_{(5, 47064)}$=5370.08***; R^2=*.36.*
***p<.005, ***p<.0005*

Table 4

Descriptive Statistics for Personality Self-Reported Executive Problems (Webexec scale) and Digit Span

Variable	*n*	*M*	*SD*	α	Range		Skew
					Potential	Actual	
Extraversion	345	29.75	7.24	.86	9-45	11-45	-0.28
Agreeableness	345	27.96	4.33	.72	7-35	11-35	-0.86
Conscientiousness	345	35.10	7.37	.84	10-50	14-50	-0.27
Neuroticism	345	21.16	6.69	.82	8-40	8-40	0.57
Openness to Experience	345	26.72	4.97	.69	7-35	10-35	-0.50
Webexec	345	11.63	3.60	.80	6-30	6-24	0.74
Digit Span Forward	345	9.32	4.31		0-16	0-16	-0.99
Digit Span Backward	345	8.29	4.31		0-14	0-14	-0.67

Table 5

Pearson's r Correlations of Personality with Self-Reported Executive Problems (Webexec scale) and Digit Span

	Webexec	Digit Span Forward	Digit Span Backward
Extraversion	-.17**	-.03	-.03
Agreeableness	-.16**	.18**	.13*
Conscientiousness	-.52***	.01	.02
Neuroticism	.45***	-.09	-.05
Openness to Experience	-.04	.22***	.16**
Webexec	-	-.08	-.08
Digit Span Forward		-	.78***

Note. For all correlations, *N*=345. Findings for Openness to Experience must be regarded as speculative given low internal reliability (alpha=.69).
*p<.05, **p<.005, ***p<.0005

Table 6

Regression of Webexec, Forward and Backward Digit Span scores on Personality

	Webexec				Digit Span Forward				Digit Span Backward			
Variable	*B*	*SE*	β	*t*	*B*	*SE*	β	*t*	*B*	*SE*	β	*t*
(Constant)	15.29	1.96		7.79***	3.23	2.77		1.17	3.06	2.83		1.08
Extraversion	0.02	0.02	0.05	1.01	-0.06	0.03	-0.10	-1.74	-0.05	0.03	-0.08	-1.39
Agreeableness	0.02	0.04	0.02	0.41	0.14	0.06	0.14	2.4*	0.10	0.06	0.10	1.78
Conscientiousness	-0.20	0.02	-0.42	-8.39***	0.00	0.03	0.00	-0.05	0.02	0.04	0.03	0.43
Neuroticism	0.16	0.03	0.29	5.64***	-0.04	0.04	-0.06	-0.98	-0.01	0.04	-0.02	-0.34
Openness to Experience	-0.04	0.03	-0.05	-1.18	0.19	0.05	0.21	3.99***	0.13	0.05	0.15	2.82**
R^2	0.34				0.08				0.04			
F	34.41***				5.89***				2.94*			

*p<.05, **p<.005, ***p<.0005

Table 7

Descriptive Statistics for all Study 3 Measures

Variable	*n*	*M*	*SD*	α	Range		Skew
					Potential	Actual	
Webexec	102	11.87	2.52	.61	6-30	6-19	0.46
DEX	102	21.93	10.32	.86	0-80	5-51	0.73
NEO-FFI Neuroticism	102	24.00	6.36	.72	0-48	9-40	0.24
NEO-FFI Extraversion	101	29.74	5.12	.69	0-48	17-41	-0.04
NEO-FFI Openness to Experience	103	27.00	5.90	.67	0-48	5-40	-0.66
NEO-FFI Agreeableness	102	30.51	6.05	.77	0-48	18-42	-0.27
NEO-FFI Conscientiousness	100	30.36	7.25	.87	0-48	11-48	-0.28
IPIP Neuroticism	100	21.00	5.27	.71	8-40	12-37	0.86
IPIP Extraversion	101	30.86	5.50	.80	9-45	18-45	-0.07
IPIP Openness to Experience	102	24.24	4.48	.64	7-35	15-32	-0.08
IPIP Agreeableness	101	28.15	3.29	.61	7-35	18-35	-0.63
IPIP Conscientiousness	97	34.48	6.61	.86	10-50	22-50	0.16
Digit Span Forward	103	9.93	2.32		0-16	5-15	0.09
Digit Span Backwards	103	8.17	2.15		0-16	4-14	0.51
Digit Span Sequencing	103	8.41	1.89		0-16	4-14	0.05
Trails A (seconds)	103	19.69	5.87			10-49	1.51
Trails B (seconds)	102	47.59	15.80			27-97	1.19
Phonemic fluency	102	31.57	8.62			13-53	0.33
Semantic fluency	102	19.47	5.25			9-33	0.33

Table 8

Pearson's r Correlations Between Self-report Measures, Personality and Cognitive Tests

Variable	Digits Forward	Digits Backwards	Digits Sequencing	Trails A	Trails B	Phonemic fluency	Semantic fluency	Webexec†	DEX
NEO-FFI Neuroticism	.05	.11	.08	.07	.16	.04	-.03	.20*	.45***
IPIP Neuroticism	.15	.21*	.08	.01	.03	-.01	-.08	.08	.31***
NEO-FFI Extraversion†	-.05	-.02	-.04	-.14	-.13	-.01	-.03	-.04	-.09
IPIP Extraversion	.00	-.02	-.15	.00	-.05	.01	.01	.03	-.08
NEO-FFI Openness to Experience†	.03	.14	.14	-.18	.00	-.05	.02	-.05	-.17
IPIP Openness to Experience†	-.10	-.02	.06	-.01	.07	-.17	-.09	.04	-.18
NEO-FFI Agreeableness	-.07	-.11	-.06	.09	.00	.03	-.05	-.05	-.27***
IPIP Agreeableness†	-.10	-.11	.08	.14	.08	.09	.04	-.22*	-.26**
NEO-FFI Conscientiousness	-.09	.05	.05	-.15	-.11	-.06	.07	-.56***	-.48***
IPIP Conscientiousness	.16	.04	-.03	-.08	-.07	-.08	.07	-.52***	-.46**
Webexec†	.00	-.11	-.08	-.02	.15	.04	-.03	-	.58***
DEX	-.02	-.08	-.04	.06	.08	.11	.03	.58***	-

†These variables have Cronbach's alpha <.7 in this sample.
*p<.05, **p<.005, ***p<.0005

Table 9

Regressions of Composite Executive Function Scores on Self-Reported Executive Problems and Composite Personality Scores

	Working Memory					Cognitive Flexibility				
		B	*SE*	β	*t*		*B*	*SE*	β	*t*
(Constant)		0.26	0.27		0.97		-0.15	0.18		-0.82
Composite N		0.22	0.12	0.24	1.80		0.01	0.08	0.01	0.06
Composite E		0.02	0.11	0.01	0.05		-0.02	0.07	-0.03	-0.29
Composite O		0.07	0.12	0.07	0.57		-0.05	0.08	-0.07	-0.61
Composite A		0.01	0.11	0.01	0.10		0.08	0.08	0.13	1.05
Composite C		0.01	0.12	0.01	0.10		0.03	0.08	0.04	0.33
DEX		-0.01	0.01	-0.14	-1.07		0.01	0.01	0.12	0.89
R^2	0.06					0.03				
F	0.86					0.41				
(Constant)		0.76	0.57		1.33		-0.27	0.40		-0.69
Composite N		0.17	0.12	0.18	1.46		0.03	0.08	0.05	0.37
Composite E		0.02	0.11	0.02	0.17		-0.03	0.07	-0.04	-0.34
Composite O		0.10	0.11	0.09	0.84		-0.07	0.08	-0.10	-0.83
Composite A		0.02	0.11	0.02	0.14		0.07	0.08	0.12	0.97
Composite C		-0.04	0.13	-0.04	-0.28		0.03	0.09	0.06	0.39
Webexec[†]		-0.07	0.05	-0.18	-1.37		0.02	0.03	0.09	0.70
R^2	0.07					0.03				
F	0.98					0.36				

[†]Variable has Cronbach's alpha <.7 in this sample.

*p<.05