RESEARCH ARTICLE

The relationship between the frequency of number-puzzle use and baseline cognitive function in a large online sample of adults aged 50 and over

Helen Brooker[1] | Keith A. Wesnes[1,2] | Clive Ballard[1] | Adam Hampshire[3] |
Dag Aarsland[4,6] | Zunera Khan[7] | Rob Stenton[5] | Maria Megalogeni[4] | Anne Corbett[1]

[1]University of Exeter Medical School, University of Exeter, Exeter, UK

[2]Wesnes Cognition Ltd, UK

[3]Department of Medicine, Imperial College London, London, UK

[4]Institute of Psychiatry, Psychology, & Neuroscience, King's College London, London, UK

[5]Manta Software Ltd, Cambridge, UK

[6]Mental Health of Older Adults and Dementia Clinical Academic Group, South London and Maudsley NHS Foundation Trust, London, UK

[7]Wolfson Centre for Age-Related Diseases, King's College London, London, UK

Correspondence
Helen Brooker, University of Exeter Medical School, University of Exeter, Exeter, UK.
Email: h.brooker@exeter.ac.uk

Objective: Establishing affordable lifestyle interventions that might preserve cognitive function in the aging population and subsequent generations is a growing area of research focus. Data from the PROTECT study has been utilised to examine whether number-puzzle use is related to cognitive function in older adults.

Methods: Data from 19 078 healthy volunteers aged 50 to 93 years old enrolled on the online PROTECT study were evaluated for self-reported frequency of performing number puzzles. Two cognitive-test batteries were employed to assess core aspects of cognitive function including reasoning, focussed and sustained attention, information processing, executive function, working memory, and episodic memory. Analysis of covariance was used to establish the differences between the six frequency groups.

Results: Highly statistically significant main effects of the frequency of performing number puzzles were seen on all 14 cognitive measures, with *P* values of less than 0.0004. Interestingly, participants who reported engaging in number puzzles more than once a day had superior cognitive performance on 10 core measures compared with all other frequency groups, although not all were statistically significant.

Conclusions: This study has identified a close relationship between frequency of number-puzzle use and the quality of cognitive function in adults aged 50 to 93 years old. In order to determine the value of these findings as a potential intervention, further research should explore the type and difficulty of the number puzzles. These findings further contribute to the growing evidence that engaging in mentally stimulating activities could benefit the brain function of the ageing population.

KEYWORDS

ageing, attention, cognition, memory, number puzzles, PROTECT

1 | INTRODUCTION

There is increasing awareness of the need to find affordable interventions that might reduce the risk of dementia and Alzheimer disease (AD) and enable older people to proactively manage their risk of cognitive decline.[1] The ageing global population presents one of the greatest challenges for global health sectors and economies.[2] Dementia affects nearly 50-million people around the world, and with ageing populations, it is anticipated that up to 132 million could be affected by 2050. Risk reduction interventions are therefore a priority for

research to ensure current and future generations are able to take responsibility for their own brain health.[1]

Whilst some functional and cognitive decline is a natural consequence of ageing, cognitive loss may also lead to more severe impairment and dementia in later life. Dementia is one of the leading chronic conditions affecting older people. The condition has a devastating impact on an individual's quality of life and the lives of those around them. More research is needed to define the intricacies and hallmarks of early onset of preclinical cognitive deficits in order to facilitate early diagnosis and to target effective preventative strategies towards at-risk groups. Key at-risk individuals are those with existing early cognitive deficits, including mild cognitive impairment (MCI) and age-associated cognitive decline (AACD),[3] both of which are associated with an increased likelihood of future conversion to dementia. Published criteria for MCI and AACD enable identification of these groups, and this is proving valuable for research into preventative approaches.

One area of promising, and rapidly growing, research is the use of cognitive activities to reduce the risk of cognitive decline. The evolution of the internet and its place in society has opened doors for different types of cognitive training and brain-stimulating activities, which are available to individuals online 24 hours a day. In the same way in which physical activity has been promoted as essential to a healthy lifestyle, there has been an increase in public interest in brain health and how these approaches might be used to maintain it. This complies with the simple concept of "use it or lose it," which suggests that the brain must continue to be used in a way that it is stimulated and challenged throughout life.[4] Additionally, data is accumulating to support the case that sustaining an active cognitive lifestyle can contribute to favourable cognitive stability in ageing.[5] This evidence, combined with evidence that cognitive reserve formed by a combination of education attainment, career, and sustained cognitive activity can also play a role in cognitive trajectory,[6] further strengthens the value of exploring affordable interventions in this field.

Number puzzles are an example of a cognitively stimulating mental exercise, which are widely accessible. An example is the popular Japanese game sudoku, which requires the correct placement of nine nonrepeating digits and has task demands that stimulate information processing and working memory. The value of number-puzzle use in older adults who later develop dementia has previously been shown in studies in which older adults who regularly complete puzzles perform at a higher level on cognitive tasks compared with those who do not,[6] and a number of additional studies have reported benefit to working memory, attention, and accuracy in cognitive tests.[7-9] Paulraj et al investigated the relationship between regular use of puzzles and performance on the Wisconsin Card-Sorting Test (WCST) in healthy adults, showing that regular sudoku and/or crossword use was associated with better executive function.[10] Overall higher frequency puzzle users were able to complete more task categories, which suggests superior cognitive function and problem-solving skills. A further study in 65 000 individuals aged 18 to 90 years old and who answered a question about the frequency of

Key points

- This paper evaluated the relationship between the frequency of engaging in number puzzles, such as sudoku, and cognitive function in a large online sample of 19 078 individuals aged 50 to 93.

- Cognitive function was associated with the frequency of number-puzzle use, with individuals who never or only occasionally use puzzles showing poorer reasoning, attention, information processing, working memory, and episodic memory.

performing number puzzles showed favourable cognitive performance in those who engaged in daily puzzle use compared with those who never used them.[11] More recently a large-scale cross-sectional analysis conducted in the Survey of Health, Ageing and Retirement in Europe (SHARE) also reported more favourable cognitive function in adults reporting regular use of sudoku and other cognitively stimulating activities.[12] This literature points towards a potential risk-reduction relationship between cognition in healthy adults and use of number puzzles. However, to date, studies have either used small cohorts or broad cognitive-test paradigms to investigate this relationship. Novel online cognitive-test systems now offer the opportunity to explore the impact of number-puzzle use in large numbers of cognitively healthy adults on specific aspects of cognitive function and individual cognitive domains.

A large scale online longitudinal study in the United Kingdom, PROTECT, provides a tailor-made environment to explore research questions relating to cognitive health in ageing and to address the current gaps in the literature around the impact of lifestyle factors on dementia risk.[13-15] The cohort of over 20 000 adults aged 50 years old and above complete two independent annual online assessments of core aspects of cognitive function and lifestyle behaviours, offering a valuable and evolving dataset of independently validated cognitive variables with which to explore the factors, which influence brain function as we age. The analysis reported here utilises baseline data from the PROTECT study cohort to explore the relationship between core measures of cognitive function and the self-reported frequency of number-puzzle use.

2 | METHODS

2.1 | Study design

This is a cross-sectional analysis of data from the ongoing online PROTECT study (http://www.protectstudy.org.uk/), which was launched[13] in November 2015. The study received ethical approval from the UK London Bridge National Research Ethics Committee (Ref: 13/LO/1578).

2.2 | Participants and eligibility criteria

PROTECT participants are aged 50 years old and above, have access to a computer and the Internet and do not have a diagnosis of dementia. Enrolment to the study was completed via the study website following national publicity and signposting through partner cohorts and organisations. Participants gave their electronic informed consent through the online registration process.

2.3 | Demographic and lifestyle data collection

All participants detailed their demographic information at baseline through an online questionnaire adapted from the Office of National Statistics, which included age, sex, ethnicity, and education level. Education level was categorised from secondary education (GCSE/O-Levels) (score of 1) to doctorate (PhD) (score of 6). Participants also completed a questionnaire that captured information about lifestyle items. The question of interest for this analysis was "How frequently do you engage in number puzzles, eg, sudoku?" The six possible responses to the study question were "more than once a day," "once a day," "once a week," "once a month," "occasionally," and "never."

2.4 | Cognitive assessment

Two independent online cognitive-test systems were utilised for this study—the PROTECT Cognitive-Test Battery (PCTB) and the CogTrack system. The PCTB assesses aspects of attention and working memory and takes 10 minutes to complete. The four tasks that make up the battery are paired-associate learning, in which participants are shown a series of objects in "windows" and are then asked to select the correct location of each object using a ratchet-style approach; digit span, in which participants were presented with a sequences of digits to encode and recall; spatial working memory in which participants search a series of on-screen boxes to find a hidden symbol; and verbal reasoning assessed with a grammatical reasoning task, in which participants determine the accuracy of a series of grammatical statements about a picture. These tasks are described in full in previous publications.[9,11,13] The outcome measure of each task is the total score of correct responses, which is corrected for errors made.

The CogTrack system battery adopted for the PROTECT study assesses aspects of attention, episodic memory, and reasoning and takes around 15 minutes to complete. The five tasks used in this study were simple reaction time, choice reaction time, digit vigilance, pattern separation, and grammatical reasoning. The first four have been described previously,[14] and the fifth is based on Baddeley's logical-reasoning paradigm.[9,11,13] The tasks have numerous parallel forms to ensure repeat stimuli is not given to participants at each test session. Four validated composite measures are derived from the three attention tasks. The attentional-intensity index is the sum of the speed scores from the three tasks and assesses the ability focus attention. The sustained-attention index is calculated by combining accuracy scores. Cognitive reaction time is the difference between simple–reaction-time median speed and choice–reaction-time median speed and reflects the extra information processing involved in choice reaction time. The fourth measure is the sum of the three coefficients of variance derived from the attention tasks to form the attentional fluctuation index, which measures moment-to-moment fluctuations in attention.[16] Speed and accuracy scores from the object pattern separation (PS) task are analysed separately for the ability to detect the original pictures and the ability to reject the closely similar ones. Finally, the outcome measures of the Grammatical Reasoning task are the percentage of correct responses and the median response time.

The participants were requested to perform the two cognitive test systems up to three times over a period of seven days, leaving at least 24 hours between each testing session.

2.5 | Data analysis

All data available from the two cognitive-test packages were averaged over the three sessions to obtain a single score per participant for each measure at baseline. In order to determine whether the frequency of puzzle use was related to performance on the various cognitive tasks, ANCOVAs were conducted using the MIXED procedure from the software package SAS Version 9.4. The frequency of use of number puzzles was fitted as the main between group factor with six levels, (1) More than once a day to (6) Never. Age, gender, education, and the number of times the tasks had been performed were fitted as covariates. Comparisons between the frequency groups were made using paired t tests, with the residual error terms from the ANCOVAs being used to determine the Cohen's d effect sizes of any differences identified. Cohen's classification of effect sizes was adopted, $d = 0.2$ (small), $d = 0.5$ (medium), and $d = 0.8$ (large).[17]

3 | RESULTS

3.1 | Cohort characteristics

In total, at the time of data analysis, baseline data were available for 21 463 participants in the PROTECT study, covering the period from November 2015 to September 2017. Of these, 19 212 responded to the number-puzzle frequency questionnaire, and 19 078 of these also performed either or both of the PCTB and CogTrack tasks and were included in the analysis. The population included 14 012 females, mean age 61.1 years old (SD 6.9, range 60 to 92) and 5066 males, mean age 63.4 years old (SD 7.7, range 60 to 93). The demographics for the analysed population are presented in Table 1. The population ages varied between the puzzle-frequency groups, with those who performed number puzzles more than once a day being the oldest, those who performed them daily a year younger, and the group who performed puzzles monthly were the youngest at just under 60 years old.

TABLE 1 Demographics for the participants who performed the two cognitive test systems according to reported frequency of number-puzzle use.

Question: How frequently do you engage in number puzzles, eg, sudoku?								
		Age (years)			Education		Females	Males
Response	n	Mean	SD	Range	Mean	SD	n	n
PCTB tasks								
1 More than once a day	1214	64.9	7.2	50-91	3.26	1.4	886	328
2 Once a day	2870	63.9	7.0	50-89	3.41	1.3	2157	713
3 Once a week	2918	61.6	6.8	50-91	3.37	1.4	2233	685
4 Once a month	894	59.7	6.5	50-84	3.39	1.4	658	236
5 Occasionally	5170	60.4	7.0	50-89	3.29	1.4	3806	1364
6 Never	5766	61.3	7.3	50-92	3.22	1.5	4092	1674
Total	18 832	61.7	7.2	50-92	3.30	1.4	13 832	5000
CogTrack tasks								
1 More than once a day	1094	65.1	7.2	50-93	3.31	1.4	786	308
2 Once a day	2622	64.0	7.0	50-89	3.45	1.3	1945	677
3 Once a week	2623	61.7	6.8	50-91	3.41	1.3	2004	619
4 Once a month	806	59.7	6.4	50-82	3.42	1.4	588	218
5 Occasionally	4663	60.5	7.0	50-89	3.32	1.4	3396	1267
6 Never	5137	61.5	7.4	50-92	3.25	1.5	3570	1567
Total	16 945	61.8	7.2	50-93	3.30	1.4	12 289	4656

Abbreviations: PCTB, PROTECT Cognitive-Test Battery; SD, standard deviation.

3.2 | Association of cognitive performance on the PCTB with usage of number puzzles

All four PCTB tasks showed highly significant ($P < 0.00001$) main effects of the frequency of number-puzzle use (Table 2). The pattern was for less frequent puzzle use to be associated with poorer performance on the tasks. Cognitive performance in the group who never performed number puzzles was notably poorer than in all other groups (Table 3, each comparison $P < 0.00001$), with Cohen's d effect sizes for paired-associate learning ranging from 0.23 to 0.41, for digit span from 0.17 to 0.29, for spatial working memory from 0.28 to 0.47, and for verbal reasoning from 0.21 to 0.4. Cognitive performance in the group who performed puzzles occasionally was significantly poorer when compared with each of the three highest frequency groups on all four tasks ($P = 0.0003$ to <0.00001). The other major difference was that the group who performed puzzles monthly was significantly poorer than each of the three highest frequency groups for paired-associate learning ($P = 0.0054$ to <0.0001) and verbal reasoning ($P = 0.0004$ to <0.00001).

3.3 | Association of cognitive performance on the CogTrack system with usage of number puzzles

As with the PCTB tasks, each of the 10 scores from the CogTrack tasks showed significant main effects of number-puzzle–use frequency, these being $P < 0.00001$ for each score, apart from the speed

score for the original stimuli in the pattern separation (PS) task ($P = 0.004$). The pattern of results for the CogTrack attention tasks showed a more variable pattern of improvement with puzzle frequency over the four measures. The "never" group again performed significantly more poorly than each other frequency group for the attentional-intensity index (all $P < 0.00001$, Cohen's d from 0.15 to 0.27), the sustained-attention index ($P = 0.0054$ to <0.00001; $d = 0.09$ to 0.18), and the attentional-fluctuation index (all $P < 0.00001$; $d = 0.11$ to 0.18). Thereafter, there were only occasional differences for these three measures. For cognitive reaction time, the three lowest frequency groups were poorer than each of the three highest frequency groups, with only the difference between the "monthly" and "weekly" groups missing significance ($P = 0.0657$). The peak difference between the groups on this measure had an effect size of 0.14.

For the PS task, the accuracy and speed scores for the original stimuli as well as the closely similar stimuli are presented. To facilitate comparisons between the two types of stimuli, the accuracy and speed scores for each are plotted over the same range. The pattern was again for a relationship of frequency to quality of performance, although the differences between the frequency groups were notably greater for the closely similar stimuli as opposed to the original stimuli. For the ability to correctly identify the stimuli, the main effect was for the group who never performed number puzzles to be significantly poorer than each other group, both for the original stimuli ($P = 0.0005$ to <0.00001; $d = 0.07$ to 0.15) and the closely similar stimuli (all $P < 0.00001$; $d = 0.13$ to 0.25). For the speed of correctly

TABLE 2 The outcomes of the ANCOVAs of the 14 cognitive scores, including the LSMeans and SEM for each Group.

Measure (Units)	ANCOVA Main Effect of Group			Group 1 More than once per day		Group 2 One a day		Group 3 Weekly		Group 4 Monthly		Group 5 Occasionally		Group 6 Never	
	F	df	P	LSMean	sem	LSMean	sem	LSMean	sem	LSMean	sem	LSMean	sem	LSMean	sem
PCTB tasks															
Paired-associate learning (Total score)	108.17	5.19000	<0.00001	4.62	0.02	4.64	0.01	4.61	0.01	4.52	0.03	4.48	0.01	4.31	0.01
Digit span (Total score)	45.76	5.19000	<0.00001	7.56	0.04	7.52	0.03	7.53	0.03	7.45	0.05	7.38	0.02	7.12	0.02
Spatial working memory (Total score)	118.43	5.19000	<0.00001	7.94	0.07	7.81	0.04	7.83	0.04	7.69	0.08	7.50	0.03	6.87	0.03
Verbal reasoning (Total score)	94.12	5.19000	<0.00001	33.33	0.25	33.38	0.16	33.02	0.16	31.86	0.29	31.70	0.12	29.90	0.11
CogTrack tasks															
Attentional-intensity index (msec)	33.1	5.17000	<0.00001	1330	3.9	1332	2.5	1336	2.5	1341	4.6	1346	1.9	1364	1.8
Sustained-attention index (msec)	12.6	5.17000	<0.00001	94.1	0.14	94.3	0.09	94.3	0.09	94.4	0.16	94.0	0.07	93.7	0.06
Cognitive reaction time (msec)	8.29	5.17000	<0.00001	170.9	1.7	172.0	1.1	173.6	1.1	177.7	1.9	176.7	0.8	178.8	0.8
Attentional-fluctuation index (CV%)	15.78	5.17000	<0.00001	24.1	0.38	23.8	0.24	24.2	0.24	23.7	0.43	24.6	0.18	26.0	0.17
PS-task original stimuli accuracy (%)	10.51	5.17000	<0.00001	91.5	0.21	91.3	0.14	91.4	0.14	91.6	0.25	91.0	0.10	90.5	0.10
PS-task novel stimuli accuracy (%)	29.39	5.17000	<0.00001	70.4	0.47	69.6	0.30	70.3	0.30	69.7	0.54	68.6	0.23	66.6	0.21
PS-task original stimuli speed (msec)	3.46	5.17000	0.004	1212	7.7	1218	5.0	1230	5.0	1244	8.9	1236	3.7	1234	3.5
PS-task novel stimuli speed (msec)	19.27	5.17000	<0.00001	1424	10.1	1452	6.5	1472	6.5	1484	11.7	1491	4.9	1510	4.6
Grammatical-reasoning accuracy (%)	26.91	5.17000	<0.00001	91.1	0.24	91.4	0.16	91.6	0.15	91.3	0.28	90.9	0.12	89.8	0.11
Grammatical-reasoning speed (msec)	82.37	5.17000	<0.00001	3169	25	3191	16	3257	16	3362	29	3406	12	3519	11

Abbreviations: CV%, coefficient of variance percentage; *df*, degrees of freedom; *F*, *F* Value; LSMean, least square mean; msec, millisecond; PCTB, PROTECT Cognitive-Test Battery; SD, standard deviation; sem, standard error of mean.

TABLE 3 The results of the comparisons between the frequency groups for the 14 cognitive scores, including the LSMeans, P values, and Cohen's d effect sizes.

Group Comparisons
1 = More than once a day 2 = Once a day 3 = Once a week 4 = Once a month 5 = Occasionally 6 = Never

		1 vs 2	1 vs 3	1 vs 4	1 vs 5	1 vs 6	2 vs 3	2 vs 4	2 vs 5	2 vs 6	3 vs 4	3 vs 5	3 vs 6	4 vs 5	4 vs 6	5 vs 6
PCTB tasks																
Paired-associate learning (total score)	LSMean	-0.023	0.005	0.095	0.135	0.311	0.028	0.118	0.158	0.335	0.090	0.130	0.306	0.040	0.216	0.177
	P	0.3714	0.8611	0.0054	<0.00001	<0.00001	0.1655	<0.0001	<0.00001	<0.00001	0.0021	<0.00001	<0.00001	0.1518	<0.00001	<0.00001
	Cohen's d	-0.03	0.01	0.12	0.18	0.41	0.04	0.15	0.21	0.44	0.12	0.17	0.40	0.05	0.28	0.23
Digit span (total score)	LSMean	0.046	0.032	0.116	0.178	0.442	-0.014	0.069	0.132	0.395	0.084	0.146	0.410	0.062	0.326	0.264
	P	0.3765	0.5432	0.089	0.0003	<0.00001	0.7212	0.2409	0.0003	<0.00001	0.1534	<0.0001	<0.00001	0.2603	<0.00001	<0.00001
	Cohen's d	0.03	0.02	0.08	0.12	0.29	-0.01	0.05	0.09	0.26	0.05	0.10	0.27	0.04	0.21	0.17
Spatial working memory (total score)	LSMean	0.129	0.111	0.255	0.440	1.074	-0.018	0.126	0.311	0.944	0.145	0.329	0.963	0.185	0.818	0.634
	P	0.0969	0.1553	0.0114	<0.00001	<0.00001	0.7592	0.1508	<0.00001	<0.00001	0.0971	<0.00001	<0.00001	0.0251	<0.00001	<0.00001
	Cohen's d	0.06	0.05	0.11	0.19	0.47	-0.01	0.06	0.14	0.42	0.06	0.14	0.42	0.08	0.36	0.28
Verbal reasoning (total score)	LSMean	-0.05	0.31	1.47	1.63	3.43	0.36	1.52	1.68	3.48	1.15	1.31	3.12	0.16	1.97	1.81
	P	0.8637	0.2843	0.0001	<0.00001	<0.00001	0.1065	<0.00001	<0.00001	<0.00001	0.0004	<0.00001	<0.00001	0.6039	<0.00001	<0.00001
	Cohen's d	-0.01	0.04	0.17	0.19	0.40	0.04	0.18	0.20	0.41	0.13	0.15	0.36	0.02	0.23	0.21
CogTrack tasks																
Attentional-Intensity index (msec)	LSMean	-2.2	-6.0	-11.2	-15.7	-34.5	-3.8	-9.0	-13.5	-32.2	-5.2	-9.7	-28.5	-4.5	-23.3	-18.8
	P	0.6288	0.1963	0.063	0.0003	<0.00001	0.2901	0.0862	<0.00001	<0.00001	0.3184	0.0022	<0.00001	0.3633	<0.00001	<0.00001
	Cohen's d	-0.02	-0.05	-0.09	-0.12	-0.27	-0.03	-0.07	-0.10	-0.25	-0.04	-0.07	-0.22	-0.03	-0.18	-0.15
Sustained-attention index (%)	LSMean	-0.257	-0.224	-0.371	0.068	0.422	0.033	-0.114	0.325	0.679	-0.147	0.291	0.646	0.439	0.793	0.355
	P	0.115	0.1718	0.0797	0.6594	0.0054	0.7913	0.5341	0.0037	<0.00001	0.4195	0.0085	<0.00001	0.011	<0.00001	0.0001
	Cohen's d	-0.06	-0.05	-0.08	0.01	0.09	0.01	-0.03	0.07	0.15	-0.03	0.06	0.14	0.10	0.18	0.08
Attentional-fluctuation index (CV%)	LSMean	0.28	-0.04	0.41	-0.48	-1.85	-0.32	0.13	-0.76	-2.13	0.45	-0.44	-1.81	-0.89	-2.26	-1.37
	P	0.5232	0.9278	0.4761	0.2542	<0.00001	0.3434	0.7988	0.0126	<0.00001	0.3639	0.1473	<0.00001	0.0589	<0.00001	<0.00001
	Cohen's d	0.02	0.00	0.03	-0.04	-0.15	-0.03	0.01	-0.06	-0.17	0.04	-0.04	-0.15	-0.07	-0.18	-0.11
Cognitive reaction time (msec)	LSMean	-1.07	-2.68	-6.75	-5.76	-7.86	-1.61	-5.68	-4.69	-6.79	-4.08	-3.09	-5.19	0.99	-1.11	-2.10
	P	0.5884	0.1782	0.0086	0.002	0.00002	0.2922	0.0108	0.0006	<0.00001	0.0657	0.0216	<0.0001	0.6368	0.0594	0.0593
	Cohen's d	-0.02	-0.05	-0.12	-0.10	-0.14	-0.03	-0.10	-0.09	-0.12	-0.07	-0.06	-0.09	0.02	-0.02	-0.04
PS-task original stimuli accuracy (%)	LSMean	0.23	0.08	-0.09	0.53	1.02	-0.15	-0.32	0.30	0.79	-0.17	0.45	0.94	0.62	1.11	0.50
	P	0.3665	0.7479	0.7838	0.0269	<0.00002	0.4513	0.2629	0.0848	<0.00001	0.5439	0.0095	<0.00001	0.021	<0.00003	0.0005
	Cohen's d	0.03	0.01	-0.01	0.08	0.15	-0.02	-0.05	0.04	0.11	-0.02	0.06	0.13	0.09	0.16	0.07
PS-task novel stimuli accuracy (%)	LSMean	0.81	0.07	0.68	1.81	3.77	-0.73	-0.13	1.00	2.96	0.60	1.73	3.69	1.13	3.09	1.96
	P	0.1442	0.8951	0.3442	0.0005	<0.00001	0.0846	0.8361	0.0082	<0.00001	0.3277	<0.00001	<0.00001	0.053	<0.00001	<0.00001
	Cohen's d	0.05	0.00	0.04	0.12	0.25	-0.05	-0.01	0.07	0.19	0.04	0.11	0.24	0.07	0.20	0.13

(Continues)

TABLE 3 (Continued)

| | | Group Comparisons 1 = More than once a day 2 = Once a day 3 = Once a week 4 = Once a month 5 = Occasionally 6 = Never | | | | | | | | | | | | | | |
		1 vs 2	1 vs 3	1 vs 4	1 vs 5	1 vs 6	2 vs 3	2 vs 4	2 vs 5	2 vs 6	3 vs 4	3 vs 5	3 vs 6	4 vs 5	4 vs 6	5 vs 6
PS-task original stimuli speed (msec)	LSMean	−5.5	−17.5	−31.7	−24.3	−21.6	−12.0	−26.2	−18.8	−16.1	−14.2	−6.8	−4.1	7.5	10.1	2.7
	P	0.5478	0.0568	0.0075	0.0049	0.0112	0.0882	0.0108	0.0028	0.0089	0.1645	0.2769	0.503	0.4402	0.2924	0.6048
	Cohen's d	−0.02	−0.07	−0.13	−0.10	−0.09	−0.05	−0.10	−0.07	−0.06	−0.06	−0.03	−0.02	0.03	0.04	0.01
PS-task novel stimuli speed (msec)	LSMean	−28.1	−48.6	−60.6	−67.6	−86.7	−20.4	−32.5	−39.5	−58.6	−12.0	−19.1	−38.1	−7.0	−26.1	−19.1
	P	0.0183	<0.0001	<0.0001	<0.00001	<0.00001	0.026	0.0154	<0.00001	<0.00001	0.3667	0.0186	<0.00001	0.5773	0.0372	0.0044
	Cohen's d	−0.09	−0.15	−0.18	−0.20	−0.26	−0.06	−0.10	−0.12	−0.18	−0.04	−0.06	−0.12	−0.02	−0.08	−0.06
Grammatical-reasoning accuracy (%)	LSMean	−0.31	−0.49	−0.21	0.24	1.34	−0.18	0.10	0.55	1.65	0.28	0.73	1.83	0.45	1.55	1.10
	P	0.269	0.0851	0.5661	0.3786	<0.00001	0.4174	0.749	0.0049	<0.00001	0.3791	0.0002	<0.00001	0.1369	<0.00001	<0.00001
	Cohen's d	−0.04	−0.06	−0.03	0.03	0.17	−0.02	0.01	0.07	0.21	0.04	0.09	0.23	0.06	0.20	0.14
Grammatical-reasoning speed (msec)	LSMean	−22	−88	−193	−237	−351	−66	−171	−215	−328	−105	−149	−262	−44	−157	−114
	P	0.4498	0.0028	<0.00001	<0.00001	<0.00001	0.0035	<0.00001	<0.00001	<0.00001	0.0015	<0.00001	<0.00001	0.1592	<0.00001	<0.00001
	Cohen's d	−0.03	−0.11	−0.24	−0.29	−0.43	−0.08	−0.21	−0.26	−0.40	−0.13	−0.18	−0.32	−0.05	−0.19	−0.14

Abbreviations: CV%, coefficient of variance percentage; LSMean, least square mean; msec, millisecond; PCTB, PROTECT Cognitive-Test Battery.

identifying the original stimuli, the two highest frequency groups were significantly faster than the three lowest frequency groups ($P = 0.0112$ to $P = 0.0028$; $d = 0.06$ to 0.13). For the speed of identifying the closely similar stimuli, the steep profile of response was reflected by each group being significantly different to each other group, with the exception of group 4, which did not differ from the group at either side. The difference between the group who never performed puzzles and the group who performed them more than once per day had an effect size of 0.26.

For the grammatical-reasoning task, the accuracy score for the "never" group was significantly lower than each other group (all $P < 0.00001$; $d = 0.06$ to 0.23). The other difference was for the "daily" and "weekly" groups to out perform the "occasionally" group ($P < 0.005$, $d = 0.07$ and 0.09). The speed score showed a steep profile over the frequency groups, with 13 of the 15 between group comparisons reaching significance ($P = 0.0028$ to <0.00001; $d = 0.08$ to 0.43); the exceptions were the comparison between the two highest frequency groups and that between the "occasionally" and "monthly" groups.

4 | DISCUSSION

This cross-sectional baseline analysis has identified a strong association with frequency of number-puzzle use and cognitive performance in older adults using two independent cognitive-test batteries. Analysis of data from the four PCTB measures showed consistent superior performance, with effect sizes ranging from $d = 0.17$ to $d = 0.47$. It is important to note that although these effect sizes are not large, current major treatments for AD have an average effect size[18] of 0.28, and thus for a simple intervention and early data, these findings are both promising and comparable with drug interventions. The findings are a further extension to our previously published work using the same puzzle question and the PCTB, which also showed positive effects of the frequency of puzzle use on the spatial working memory, paired-associate learning, and verbal-reasoning tasks, although cognitive performance was only assessed once in this previous study.[11] This analysis also included the in-depth CogTrack-test battery, which further emphasised the impact on cognition, although the frequency-dependent pattern was not as stable across the whole battery. Participants who reported engaging in number puzzles more than once a day did have superior cognitive performance on all 10 measures. However, the frequency association was less linear.

The analysis identified benefit across several cognitive domains including episodic memory, spatial working memory, attention, processing speed, and executive function. This correlates with previous studies, which have also indicated associations with working memory, problem solving, and attention.[7-10] Combined, the literature appears to suggest a global cognitive impact, which is not restricted to a specific cognitive domain. This perhaps may be explained by the multi-domain aspects of sudoku and other similar number puzzles, where users employ aspects of problem solving, memory, and executive function to complete tasks, all of which require the use of the information

processing and attentional domains. An interesting additional potential link between number-puzzle use and neurogenesis was identified using the CogTrack PS task. The task measures ability to correctly reject closely similar pictures. Early Functional magnetic resonance imaging (fMRI) work in patients has shown that activity in the hippocampal dentate gyrus occurs when rejecting closely similar stimuli but not when identifying original stimuli.[19-21] The dentate gyrus is one of the two brain areas where neurogenesis is known to occur, and thus PS ability can be considered to be a proxy measure for this neuropathological important process.[22] In this study the effect sizes of the peak differences between the frequency groups for both the accuracy and speed outcome measures were greater for the closely similar stimuli ($d = 0.25$ and $d = 0.26$, respectively) than for the neurogenesis insensitive stimuli ($d = 0.15$ and $d = 0.09$, respectively), suggesting a tentative link between the frequency of puzzle use and neurogenesis.

Overall the pattern for both cognitive test batteries and robustness of the results are highlighted by the fact that the main effects of 13 ANCOVAs had P values of <0.0005 (Table 2). Thus, these findings further support the case that the frequency of number-puzzle use and its relationship to the quality of cognitive function in older adults may offer further foundation to there being value in such maintenance of cognitive function in older adults. There are however a number of limitations of this study, which must be addressed. There were variations in the demographic makeup of the cohort groupings. For example, those who performed number puzzles more than once per day were the oldest group, and those who engaged in number-puzzle use monthly were the youngest group. It has previously been shown that there are age-related declines on the CogTrack tasks of attention and PS in this population,[14] and thus age was used as a covariate in the analyses to address this. The nature of the analysis dictated that multiplicity of comparisons will have occurred, and thus it would be dismissive not to raise the utility of the Bonferroni corrections. However, because of the large number of comparisons reaching significance in this analysis, it is evident that there would not be a benefit from such a correction. Finally, the frequency of use group differences identified can only be classified as an association, and they do not represent evidence that number-puzzle use alone has caused the superior cognitive function. As such, it is essential that these findings be followed up with longitudinal data and comprehensive interventional trials to explore the potential value of different types and usage patterns of number puzzles as a means of maintaining cognitive health.

5 | CONCLUSION

These findings have contributed to the growing body of literature that supports the case for regular use of activities that challenge the brain in order to promote cognitive stability in ageing. In particular, these findings show that regular number-puzzle use is related to the superiority of performance of core aspects of cognitive function.

ORCID

Helen Brooker https://orcid.org/0000-0002-7908-263X
Keith A. Wesnes https://orcid.org/0000-0002-4498-8511
Anne Corbett https://orcid.org/0000-0003-2015-0316

REFERENCES

1. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *Lancet*. 2017;390(10113):2673-2734.

2. Alzheimer's Disease International. World alzheimer report: the global economic impact of dementia. 2010. Accessed 8th September, 2014.

3. Levy R. Aging-associated cognitive decline. Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization. *Int Psychogeriatrics/IPA*. 1994;6(1):63-68.

4. Salthouse TA, Berish DE, Miles JD. The role of cognitive stimulation on the relations between age and cognitive functioning. *Psychol Aging*. 2002;17(4):548-557.

5. Marioni RE, van den Hout A, Valenzuela MJ, Brayne C, Matthews FE. Active cognitive lifestyle associates with cognitive recovery and a reduced risk of cognitive decline. *J Alzheimer's Dis: JAD*. 2012;28(1):223-230.

6. Valenzuela MJ, Sachdev P. Brain reserve and cognitive decline: a non-parametric systematic review. *Psychol Med*. 2006;36(8):1065-1073.

7. Chang HS, Gibson JM. The odd-even effect in Sudoku puzzles: effects of working memory, aging, and experience. *Am J Psychol*. 2011;124(3):313-324.

8. Jin G, Li K, Qin Y, et al. fMRI study in posterior cingulate and adjacent precuneus cortex in healthy elderly adults using problem solving task. *J Neurol Sci*. 2012;318(1-2):135-139.

9. Nombela C, Bustillo PJ, Castell PF, Sanchez L, Medina V, Herrero MT. Cognitive rehabilitation in Parkinson's disease: evidence from neuroimaging. *Front Neurol*. 2011;2:82.

10. Paulraj S, Harley A, Jongon G, et al. Wisconsin Card Sorting Test Performance in healthy older adults who use sudoku or crossword puzzles. *Arch Clin Neuropsychol*. 2016;31(6):586.

11. Ferreira N, Owen A, Mohan A, Corbett A, Ballard C. Asociation between cognitively stimulating leisure activities, cognitive function and age-related cognitive decline. *Int J Geriatr Psychiatry*. 2014;30(4):422-430.

12. Litwin H, Schwartz E, Damri N. Cognitively stimulating leisure activity and subsequent cognitive function: a SHARE-based analysis. *Gerontologist*. 2017;57(5):940-948.

13. Huntley J, Corbett A, Wesnes K, et al. Online assessment of risk factors for dementia and cognitive function in healthy adults. *Int J Geriatr Psychiatry*. 2018;33(2):e286-e293.

14. Wesnes KA, Brooker H, Ballard C, McCambridge L, Stenton R, Corbett A. Utility, reliability, sensitivity and validity of an online test system designed to monitor changes in cognitive function in clinical trials. *Int J Geriatr Psychiatry*. 2017;32(12):e83-e92.

15. Eraydin IE, Mueller C, Corbett A, et al. Investigating the relationship between age of onset of depressive disorder and cognitive function. *Int J Geriatr Psychiatry*. 2019;34(1):38-46.

16. Ballard CG, Aarsland D, McKeith I, et al. Fluctuations in attention: PD dementia vs DLB with parkinsonism. *Neurol*. 2002;59(11):1714-1720.

17. Cohen J. *Statistical Power Analysis forthe Behavioural Sciences*. Second ed. New Jersey, USA: Lawrence Erlbaum Associates; 1988.

18. Rockwood K, Fay S, Song X, MacKnight C, Gorman M. Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. *CMAJ*. 2006;174(8):1099-1105.

19. Kirwan CB, Stark CE. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learn Mem*. 2007;14(9):625-633.

20. Wesnes K. Visual object pattern separation: a paradigm for studying the role of the dentate gyrus in memory disorders. *Alzheimers Dement*. 2010;6(4):e45.

21. Bakker A, Kirwan CB, Miller M, Stark CE. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*. 2008;319(5870):1640-1642.

22. Faghihi F, Moustafa AA. Impaired neurogenesis of the dentate gyrus is associated with pattern separation deficits: a computational study. *J Integr Neurosci*. 2016;15(3):277-293.