VisuoSpats: A Gamified Application to Measure Visuospatial Working Memory Volume
Visuospatial Memory Performance across the Lifespan

Takahiro Miura†1, Ken-ichiro Yabu†1, Kenichi Tanaka†2, Kazutaka Ueda†3, Tohru Ifukube (member)†1,†4

Abstract Because of rapid population aging, it is necessary to design interfaces that can decrease cognitive workload. The design implications and evaluation criteria for creating such senior-friendly or disability-friendly interfaces need to be established. One element related to memory function that can be manipulated in an interface is visuospatial working memory. However, there are few reports regarding the relationship between visuospatial working memory volume and age. In this paper, we aim to clarify how visuospatial working memory volume changes across the lifespan. We implemented a gamified application named VisuoSpats, which is based on the visual pattern span test, to measure visuospatial memory. The introduced gamification elements included points, leaderboards, and feedbacks. Three hundred and sixty-nine individuals aged 2 to 92 years old participated in this study. The results indicate that the median number of cells memorized was 7.0 (interquartile range: 5.0–9.0) across all age groups. Moreover, the number of cells memorized tends to increase as age increases until the age range of 21–25 years, and then decreases gradually with increasing age. Based on the comments by teenagers or seniors, the effective gamification elements of VisuoSpats could be competition elements such as points and ranking, or diagnostic factors, respectively.

Key words: Visuospatial memory, working memory, pattern span test, children and adolescents, elderly people.

1. Introduction

Rapid population aging necessitates the urgent development of adjustable systems and interfaces that take into account the physical and cognitive functions of individuals with either mild or severe impairments. Various multi-functional systems have evolved, and developing and refining these systems in response to user requirements has become increasingly challenging [1–3]. In particular, though some computer-literate elderly people can manipulate current interfaces as well as young adults, some elderly people resist the interfaces due to a decrease in interest in novelty as well as a decrease in cognitive function [1]. To resolve these issues, it is necessary to design an interface based on the concept of universal design or human-centered design [4–6] or to characterize an interface using a ludic design that can induce hedonic effects in the users and decrease the cognitive workload of the interface [7,8].

As well as the interface design for elderly people, an evaluation method is also needed for adjusting interfaces’ parameters to make them senior-friendly and disability-friendly. Many evaluation methods have been developed to briefly measure usability and accessibility of the interface themselves or interface elements that are related to specific subjective impressions of the interface, such as its physical, cognitive, and mental workloads.

One element that relates to memory functions for manipulating interfaces includes working memory (WM) [9]. Purpose-oriented dynamic manipulations rely on several cognitive functions, including WM [10]. WM is an advanced concept of short-term memory, which is defined as a static memory function that is capable only of preserving information. WM also serves, however, as a dynamic repository of memory information that is capable of actively modifying its contents depending on the context [10]. Some studies have evaluated interfaces based on a WM scheme. The conceptual model of WM proposed by Baddeley consists of four components: phonological loop, visuospatial sketchpad, episodic buffer, and central executive [10]. The visu-
ospatial sketchpad stores visual and spatial information, and can relate closely to the ability to use current graphical user interfaces.

The volume of the visuospatial sketchpad can be measured by spatial span tests such as Corsi block-tapping and visual pattern span tests [11, 12]. Some studies of user interface designs have found that WM can be used to improve fluent manipulations of the interface [13, 14]. However, a rough standard of visuospatial memory remains unknown. In particular, there are few reports on the relationship between visuospatial WM volume and age. Knowledge of this characteristic of memory can be used to inform interface design. For example, this knowledge can inform the interface designer about the number of cells to include in the interface. However, few studies focused on the rough quantitative standard of the number of storable buttons on visuospatial sketchpad. For gathering the data of storable buttons effectively, we assumed that gamification elements can decrease the participants’ burden to conduct a task on a measurement system.

In this paper, we aim to clarify the relationship between the visuospatial WM volume and age for guiding interface design using a visual pattern span test. We administered a gamified measurement application based on existing visual pattern span tests to over 300 participants. We employed this approach because Komarov et al. reported that there are no significant differences and similar effect sizes in evaluating user interface performance between crowdsourced-based and laboratory-based experimental methodologies [15]. This finding suggests that similar results can be obtained from small-scale laboratory experiments and large-scale field-based experiments.

Our research questions are as follows:

Q1. Is a gamified measurement application effective in measuring memory performance?
Q2. What is the volume of visuospatial memory in young and elderly people?
Q3. How about other performance metrics including correct rates, task and completion time in various age groups?

2. Related work

2.1 Interface evaluation methods

One of the most useful schemes for usability evaluation is the GOMS [16, 17]. This scheme provides a system to evaluate human interfaces qualitatively and quantitatively by sectioning an interaction into goals, operators, methods, and section rules. In particular, the GOMS-KLM (keystroke-level model) quantifies the usability of an interface by evaluating the total interaction time, which includes both user manipulation times and system execution times [18, 19]. The most well-known brief interface evaluation metric is the system usability scale (SUS), which was proposed by Brook [20]. This reliable “quick and dirty” scale is composed of 10 simple questions with a total score of 100 points. This scale can measure the general brief usability and learnability of target interfaces. The specialized scale of user acceptance includes the technology acceptance model (TAM) [21–23] and the unified theory of acceptance and use of technology (UTAUT) [24]. These two models also provide questionnaires that enable target users to be evaluated on a scale of 1 to 5.

2.2 Working memory evaluation methods

Various span tests can be used to evaluate WM performance. The reading span test and operation span test are de-facto standard tests for evaluating phonological loop. The reading span test requires participants to read sets of sentences and then try to recall the last word of each sentence at the end of each set [25]. Likewise, in the operation span test, participants check arithmetic operations and try to remember words presented after each equation [26].

The Corsi block-tapping test, visual pattern span test, and spatial span test are standard methods for measuring the visuospatial sketchpad. [11, 12]. The original Corsi block-tapping test consists of nine cubes mounted on a board. After the experimenter taps a sequence of blocks, the participant repeats the sequence in the correct order. Visuospatial WM volume can be measured by increasing the length of the sequence to be repeated [27]. Unlike the visual pattern span and spatial span tests, the Corsi block-tapping test does not use a grid layout. In these grids (e.g., 4 × 4 or 5 × 5), participants should remember the cell locations within the grid. In the visual pattern test, participants first watch each cell press and then input the cell locations without considering the sequence. In contrast, the spatial span test requires participants to remember a series of cell locations and then input the cells sequentially.

The classification performance of the Corsi block-tapping and spatial span tests is higher than that of the visual pattern span test, but these tests require more time to complete, impose a higher cognitive workload, and include rules that are difficult for children and el-
elderly individuals to understand. Thus, in this study, we employed the visual pattern span test in order to measure memory performance across age groups.

3. Proposed measurement application: VisuoSpats

We assessed the effect of workload on the visuospatial sketchpad by employing the visual pattern span test (PST) [12] (also referred to as visual pattern test (VPT) [28]). We implemented a gamified version of the PST, named VisuoSpats derived from “VisuoSpatial Sketchpad,” on a tablet computer. From the conventional gamification studies [8, 29, 30], we selected included points, and leaderboards, and feedbacks as gamification elements. The reason for employing points and feedbacks is that points and feedbacks encourage all of the participants to concentrate on continuing tasks pleasantly. In addition, points and leaderboards vitalize them, particularly enthusiastic participants, to compete and communicate each other.

The development environment of this application was Xcode 5.1.1~6.3.2 on a Mac OS X 10.10.5. We implemented this application on an Apple iPad (OS: iOS 6.1.6~9.0.2).

The view transitions of our PST application are illustrated in Fig. 1. First, participants enter their initials, age, and sex at the top view if they allow inputting. At this view, the user can check how to play the game and the score ranking of the previous players. Then, after pushing the start button, participants can select their agreement of the letter of intent on the game. Once the user presses the “Agree” button, the user moves on to the game.

When the participant touches the screen to start the memory session (start view (A) as shown in Fig. 1), the participant is immediately brought to the memory view screen (B) and must memorize the positions of the blue cells in an $n \times n$ matrix ($n > 2$) within 5 seconds. At this view, the number of blue cells and the size of the matrix can be adjusted to modify the cognitive workload. Once the blue cells are cleared, the wait view (C) is displayed for 5 seconds. After the transition to the input view (D), the participant must enter the blue cell set presented in the memory view as quickly as possible.

If the cells input by the user correctly match the cells displayed at the memory view, then the user is returned to the start view, and the number of cells presented on the next memory view are increased. If the cells input by the user do not match those displayed at the memory view, then the participant is returned to the start view, and the number of cells presented on the next memory view are the same as the previous session. Whether the participant answers correctly or incorrectly is fed back by with a sound as soon as the transition from the input view to the other views. Simultaneously, the additional points are calculated by the equation of $\sum_{k=1}^{n} 10k$ where $n$ means the number of correct cells.

When the participants make mistakes three times cumulatively, their results are displayed with a rough index of the memorability performance provided by Miller’s work [31]. Moreover, when the cumulative score that the participant has acquired exceeds the 10th score of the scoreboard shown in the top view, the board displays this score with the registered initial, age, and sex on the corresponding order.

In this study, based on a preliminary experiment, we employed the following parameters: initial number of blue cells = 2 and matrix size = $5 \times 5$. In the actual game, when the number of blue cells reaches or exceeds half the number of matrix cells, the row and column size of the matrix increases by one.
4. Experiment

We conducted our experiments to measure memory performance using our gamified system. Approval for our experiment was obtained from the Office of Life Science Research Ethics and Safety at our institution.

4.1 Participants

Three hundred and sixty-nine individuals participated in this experiment. Ages of the participants ranged from 2 to 92 years (mean age: 33.7, S.D.: 23.4 years, median age: 31.0 years, interquartile range: 11.0–48.0 years) as shown in Fig. 2. Participants included 154 males (mean age: 34.5 years, S.D.: 25.2 years), 151 females (mean age: 34.0 years, S.D.: 22.5 years), and 64 individuals who did not report their sex (mean age: 31.1 years, S.D.: 20.1 years). Participants included mainly students under 20 years old because many of the participants were visitors of the open campuses taken place in a university. All of the participants were almost healthy enough to come to the open campus by themselves.

4.2 Procedure

First, participants were asked to play a memory game. At that time, they can select whether their played results can be used for the academic use. Participants who did not consent to the use of their results were asked to input their age as “999,” whereas those who did consent to the use of their results were asked to input their initials (arbitrary), age, and sex, which were displayed on the score board as shown in the top picture in Fig. 1. Then, according to the procedure mentioned in Section 3, participants played the gamified PST. Pictures of participants playing the gamified PST are provided in Fig. 3.

4.3 Statistical analysis

Performance metrics, including visuospatial WM volume, number of correct responses, and task completion time across different conditions, were summarized by aggregating participant responses and then applying analysis of variance (ANOVA) to test for significant differences across age groups. We used a two-way ANOVA to analyze the effects of age range and the number of cells to correct responses, and a three-way ANOVA to examine the effects of the age range, the number of cells, and whether a participant correctly answered to task completion time.

5. Results and discussion

5.1 Visuospatial memory performance

Fig. 4 reports the number of cells memorized across age groups. The median and mean number of memorized cells were 7.0 (interquartile range: 5.0–9.0) and 6.97 (S.D.: 3.1), respectively, across age groups. Surprisingly, the median and its interquartile range is almost the same as the magical number seven (plus or minus two, [31]), independent of the difference in memory targets and experimental conditions.

According to the trend illustrated in Fig. 4, the number of cells memorized increases as age increases until the age range of 21–25 years, and then decreases gradually with increasing age. This tendency of age-memory relationship is consistent with the findings of previous work [32–34]. The peak median in this age group is approximately 10 cells. It can also be observed that participants under the age of 5 and over the age of 66 exhibit a similar median with an approximate median of 4 cells. However, a subset of participants over 66 years old were superior to those under 5 years old and exhibited WM performance similar to participants in their 20s and 30s. This result may be due to our study population, which included individuals completing an open campus tour; participants who attend these types of events tend to maintain cognitive and physical performance by partaking in social activities.

Fig. 5 illustrates the number of cells memorized among the 140 participants in our cohort who were under 16 years old. The number of memorable buttons increased as the participants’ age increased, according to this figure. This tendency offers further corroboration
of the findings of Hitch’s and Alloway’s studies [35,36]. This consistency indicated that our gamification scheme enabled the successful measurement of visuospatial WM performance in children, providing quantitative insight into memory development in childhood. According to this result, memory development appears to increase monotonically across childhood. Participants aged 9–11 years old performed similarly to the average across all participants. This fact suggests that visuospatial memory in humans matures to sufficient levels by 9–11 years old, and continues to develop to peak levels until 21–25 years old. However, we were unable to obtain sufficient data among participants in the 12–17-year-old age group. In a future study, we will target this age group to obtain more robust findings in late adolescence.

Based on these findings, it appears that our gamified application could correctly measure the age-related tendency of visuospatial WM volume despite the unregulated experimental conditions employed here.

5.2 Reached and correct responses

We conducted a two-way ANOVA to test for significant differences in correct responses by age range and number of cells presented. A significant main effect of both age group \( F(15, 5975) = 9.738, p < 0.01 \) and number of cells \( F(1, 5975) = 436.8, p < 0.01 \) presented
as well as a significant interaction effect of the two \((F(15,5975) = 3.698, p < 0.01)\) was observed. These results indicate that the number of correct responses is significantly affected by the number of cells presented, the participant’s age, and the interaction of age and experimental condition.

Fig. 6 shows reached and correct responses across age groups. For all age groups, the number of completed and correct responses decreases almost monotonically (e.g., logistic or sigmoid curve) as the number of cells presented increases. This decrease is steeper as the participant’s age falls below the age of 20 and above the age of 40. The curves of the participants aged \(\leq 5\) or \(76 \leq\) are most precipitous. These results reveal both the development and degradation of visuospatial WM performance. In addition, these results are consistent with the age-related performance reported in the previous studies [32–36].

5.3 Task completion time

We conducted a three-way ANOVA to test for significant differences in task completion time by age range, number of presented cells, and number of correct responses. ANOVA yielded significant main effects of age range \((F(15,5943) = 2.883, p < 0.01)\), number of presented cells \((F(1,5943) = 25.227, p < 0.01)\), and number of correct responses \((F(1,5943) = 40.999, p < 0.01)\) and significant interactions between age range and presented cells \((F(15,5943) = 4.390, p < 0.01)\) as well as presented cells and correct responses \((F(1,5943) = 7.860, p < 0.01)\).

Fig. 7 shows the task completion time as a function of the number of cells for both correct and incorrect responses. Incorrect responses were associated with a longer task completion time and more variability in the task completion time. Fig. 8 shows the task completion time as a function of age. Task completion time is observed to increase both for participants under the age of 20 and over the age of 30.

5.4 Comments from the participants

Players aged 14–16 years old reported that they could joyfully concentrate on the application because they could compete for their results among their friends or fight against their results. Most participants in this age group attended the campus event with a friend or family members, and they interacted with each other around the application. Teenaged participants tended to complete more sessions than participants in the other age groups. This result indicates that our gamification application is an effective and entertaining method by which to measure memory performance, especially in younger populations.

On the other hand, participants aged over 65 reported that they could concentrate on our game because the game diagnosed their memory performance. According to Ijsselsteijn et al., games of mental activities such as puzzles and quizzes, may be beneficial for stimulating memory and attentional abilities [7]. In addition, Ijsselsteijn et al. stated that sense of accomplishment and self-efficacy after completing a game can provide a significant boost to seniors’ self-esteem. Taken to-
Fig. 7 Task completion time as a function of the number of cells presented. The left and right panels show cases in which the participants answered correctly or incorrectly, respectively.

Fig. 8 Task completion time as a function of age group. Lines represent the number of cells presented. Both panels present data from correct responses.

gather, the diagnostic elements can encourage seniors to concentrate on conducting a task.

Moreover, most of the participants aged over 65 did not report any difficulty and could play the game after the instruction of the game, regardless of being a beginner or an expert in using a touchscreen computer: some seniors reported that they had no experience using touchscreen computers. This fact could be because our game on a touchscreen computer was simple enough to easily remember the rules and how to manipulate. Also, manipulation per se of a touchscreen computer may be enjoyable for some elderly people, according to Kobayashi et al. [37]. However, several limitations require consideration concerning the participants because our evaluation was carried out at an open campus tour. The participants who attend these types of events tend to maintain their ability to continue their daily activities. Yet, prior research suggested that the framework of gamification can be useful for elderly people with mild cognitive impairments to conduct a task [38]. Therefore, our measurement system can be effective for not only seniors without impairments but also elderly individuals with mild cognitive impairments.

6. Conclusion

To clarify the relationship between age and visuospatial memory volume, we implemented a gamified measurement application based on the visual pattern span test. We collected data from this application from over 300 participants. The contributions of this paper can be summarized as follows:

- The median number of cells memorized was 7.0 (interquartile range: 5.0–9.0). Despite differences in memory targets and experimental conditions, our result was similar to the magical number seven plus or minus two reported by Miller et al. [31].
- Visuospatial WM performance peaks around 21–25 years old. At this age range, humans can memorize a median number of approximately 10 cells. In contrast, elderly participants were able to learn and recall a median of 4–6 cells.
- The number of completed and correct responses as well as task completion time also were measured in our study. The results indicated that performance on these measures also tends to peak in participants in their 20s.
- Our application is an effective and entertaining
method by which to measure memory performance, especially in children, by incorporating gamification concepts.

Our future work includes further measurement of subjective workloads and relationships with other usable interface conditions. In addition, we also plan to implement and evaluate other gamified measurement applications with working memory components.

References

Takahiro Miura  He received his Bachelor’s degree in mechanical engineering from Tohoku University, Sendai, Japan in 2005. Later, in March 2011, he received his Ph.D. degree in information physics and computing from the University of Tokyo, Tokyo, Japan. In the same year, he worked at the National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan. He worked as a project assistant professor at the Graduate School of Information Science and Technology, the University of Tokyo in Japan in 2012, and has been a project assistant professor at the Institute of Gerontology, the University of Tokyo, Japan since 2014. His research interests include assistive technology and information accessibility for seniors and people with disabilities.

Ken-ichiro Yabu  He received his M.E. degree in mechanical engineering from Tokyo Metropolitan University, Tokyo, Japan in March 2006. He received his Ph.D. degree in system design from Tokyo Metropolitan University in March 2010. He was a project researcher at the Research Center for Advanced Science and Technology, the University of Tokyo, Tokyo, Japan from 2010 to 2014. He has been a project researcher at the Institute of Gerontology, the University of Tokyo since 2014. His research focuses on assistive device technology based on speech signal processing and information technologies and mechatronics systems for people with disabilities or the elderly.

Kenichi Tanaka  He earned his Ph.D. degree in engineering from Hokkaido University, Sapporo, Japan in 1982. He then joined Nissan Motor Company Limited. He has engaged in developing sensors and human-machine interfaces for automobiles. His research interests also include cognition in seniors and universal design for the elderly.

Kazutaka Ueda  He received his Ph.D. degree from Hiroshima University, Hiroshima, Japan in 2004. He was an assistant professor at the Training and Research Center for Clinical Psychology at Hiroshima University from 2005 to 2007. He was an assistant professor at the Research Center for Advanced Science and Technology at the University of Tokyo, Tokyo, Japan in 2007, and has been an assistant professor in the Department of Mechanical Engineering at the University of Tokyo since 2012. His research areas of interest comprise cognitive neuroscience and design engineering. His current research interests include brain mechanisms of aesthetic sensitivity, and creativity.

Tohru Ifukube  He received his M.S. and Dr. Eng. degrees in electronics from Hokkaido University in 1971 and 1977, respectively. He joined the Research Institute of Applied Electricity, Hokkaido University, in 1971. He was a Visiting Associate Professor at Stanford University, CA in 1984. In addition, he was a professor at Hokkaido University from 1989 to 2002. He was a professor at the Research Center for Advanced Science and Technology, the University of Tokyo in 2002, and has been a visiting researcher at the Institute of Gerontology, the University of Tokyo and a professor at Hokkaido University of Science since 2016. He is a Fellow of IEICE, and a professor emeritus of Hokkaido University and the University of Tokyo. His research interests include the analysis of the hearing, speaking, and vision, and the design of aid devices for people with disabilities. Some of the aid devices have been used for the blind, the deaf, and the speech-disordered, and also have been applied to virtual reality and robotic systems. He has published several books such as “Design of Voice Typewriter” (1981), “Sound-based Assistive Technology” (1997), “Evaluation of Virtual Reality” (Editor, 2000), and “Challenges of Assistive Technology” (2004).