Internet Interventions 33 (2023) 100632

Contents lists available at ScienceDirect

Internet Interventions

journal homepage: www.elsevier.com/locate/invent

Mobile applications for cognitive training: Content analysis and quality review

Myeonghwan Bang [a,b,1], Chan Woong Jang [c,1], Hyoung Seop Kim [a], Jung Hyun Park [b,c,d], Han Eol Cho [c,*]

[a] Department of Physical Medicine and Rehabilitation, National Health Insurance Service Ilsan Hospital, Goyang, Republic of Korea
[b] Department of Integrative Medicine, Yonsei University Graduate School, Seoul, Republic of Korea
[c] Department of Rehabilitation Medicine, Gangnam Severance Hospital, Rehabilitation Institute of Neuromuscular Disease, Yonsei University College of Medicine, Seoul, Republic of Korea
[d] Department of Medical Device Engineering and Management, Yonsei University Graduate School, Seoul, Korea

ARTICLE INFO

Keywords:
Cognition
Training
Rehabilitation
mHealth
Mobile apps
Quality

ABSTRACT

Background: As the number of individuals suffering from cognitive diseases continues to rise, dealing with the diminished cognitive function that comes with age has become a serious public health concern. While the use of mobile applications (apps) as digital treatments for cognitive training shows promise, the analysis of their content and quality remains unclear.

Objective: The aim of this study was to systematically search and assess cognitive training apps using the multidimensional mobile app rating scale (MARS) to rate objective quality and identify critical points.

Methods: A search was conducted on the Google Play Store and Apple App Store in February 2022 using the terms "cognitive training" and "cognitive rehabilitation." The cognitive domains provided by each app were analyzed, and the frequency and percentage according to the apps were obtained. The MARS, a mHealth app quality rating tool including multidimensional measures, was used to analyze the quality of the apps. The relationship between the MARS score, the number of reviews, and 5-star ratings were examined.

Results: Of the 53 apps, 52 (98 %) included memory function, 48 (91 %) included attention function, 24 (45 %) included executive function, and 19 (36 %) included visuospatial function. The mean (SD) scores of MARS, 5-star ratings, and reviews of 53 apps were 3.09 (0.61), 4.33 (0.30), and 62,415.43 (121,578.77). From the between-section comparison, engagement (mean 2.97, SD 0.68) obtained lower scores than functionality (mean 3.18, SD 0.62), aesthetics (mean 3.13, SD 0.72), and information (mean 3.11, SD 0.54). The mean quality score and reviews showed a statistically significant association (r = 0.447 and P = .001*). As the number of domains increased, the mean quality score showed a statistically significant increasing trend (P = .002*).

Conclusions: Most apps provided training for the memory and attention domains, but few apps included executive function or visuospatial domains. The quality of the apps improved significantly when more domains were provided, and was positively associated with the number of reviews received. These results could be useful for the future development of mobile apps for cognitive training.

1. Introduction

As society ages, the prevalence of geriatric diseases, such as dementia, mild cognitive impairment, and stroke, is increasing. The number of people with cognitive disorders is steadily rising (Nichols and Vos, 2021; Katan and Luft, 2018; Kalaria et al., 2016). The domain and severity of cognitive impairment can vary depending on the underlying disease. For example, the initial symptom of dementia is typically a decline in memory function. After a stroke, cognitive impairment can affect various domains, including memory, attention, language, and executive function, depending on the location and extent of the brain lesion (Katzman, 1986; Srikanth et al., 2003). Additionally, stroke

* Corresponding author at: Department of Rehabilitation Medicine, Gangnam Severance Hospital and Rehabilitation Institute of Neuromuscular Disease, Yonsei University College of Medicine, 211 Eonju-ro, Gangnam-gu, Seoul 06273, Republic of Korea.
E-mail address: wtzephyr@yuhs.ac (H.E. Cho).

[1] Myeonghwan Bang and Chan Woong Jang contributed equally to this work.

https://doi.org/10.1016/j.invent.2023.100632
Received 10 July 2022; Received in revised form 31 March 2023; Accepted 25 May 2023
Available online 26 May 2023

patients may experience visuospatial neglect, which causes them to ignore the side of their visual field opposite to the damaged hemisphere (Esposito et al., 2021). Previous studies have shown that patients with dementia or stroke who experience persistent cognitive impairment tend to have a high degree of functional dependence and poor long-term survival (Jokinen et al., 2015; Obaid et al., 2020). In addition, cognitive impairment is quite prevalent among elderly individuals, even in the absence of dementia or stroke (Unverzagt et al., 2001). From this perspective, coping with decreased cognitive function due to aging is an important public health challenge. Cognitive training or rehabilitation therapy is actively being performed to improve cognitive function and compensate for cognitive deficits (Bahar-Fuchs et al., 2019; Bahar-Fuchs et al., 2013). Previous studies have demonstrated that cognitive training can enhance cognition through neural plasticity, which involves both structural and functional changes in the brain (Leung et al., 2015). Specifically, cognitive training has been found to increase the volume of grey matter and enhance activity in subcortical regions, thereby improving brain function (Nguyen et al., 2019). Although more research is needed, clinical practice guidelines for treating patients with dementia and stroke in various countries, including the United States (US) and Australia, recommend cognitive training to improve the attention, memory, visuospatial, and executive function because its benefits outweigh the risks (Winstein et al., 2016; Foundation, 2017).

Recently, digital therapeutics using software such as mobile applications (apps), games, virtual reality, and artificial intelligence have become popular (Recchia et al., 2020), including in the field of cognitive training (Abbadessa et al., 2022). Despite the growing popularity of mobile apps for cognitive training, there is still a lack of evidence regarding their effectiveness as digital therapeutics. Previous studies have analyzed the quality of mobile apps for conditions such as low back pain (Machado et al., 2016) and speech disorders (Furlong et al., 2016), however, there is currently a lack of comprehensive analysis regarding the content and quality of cognitive training apps.

Furthermore, patients may find it challenging to identify high-quality apps that can truly benefit them. Relying on 5-star ratings or reviews to evaluate an app's usefulness is often subjective and provides little information about its actual quality (Kuehnhausen and Frost, 2013). Currently, the only way to ensure proper use and recommendation of cognitive training apps is through supervision by professionals. However, even experts may have difficulty determining which apps are truly helpful due to limited information on app quality.

Therefore, this research aims to examine the content of cognitive training or rehabilitation apps currently available in the commercial market, with a focus on cognitive domains, and to assess their quality. In addition, we analyze the relationship between the quality of apps, the number of reviews, and 5-star ratings to examine the criteria when choosing apps. We expected that this study will provide valuable guidance for elderly individuals and healthcare professionals seeking to select effective cognitive training apps.

2. Methods

2.1. Apps search and selection

A search on the Google Play Store and Apple App Store was performed in February 2022 by two independent experts with experience in cognitive rehabilitation (MHB and CWJ). The search terms were "cognitive training" and "cognitive rehabilitation" using US accounts. Cognitive training typically focuses on the isolated cognitive domains, but cognitive rehabilitation aims at groups of cognitive ability required to perform everyday tasks. These terms have been applied somewhat interchangeably in the previous studies, so we used both as search terms (Bahar-Fuchs et al., 2013). Android apps were installed on a Galaxy S21+ and iOS apps on an iPhone 12. Information on the app name, developer, platform (Android or iOS), version, in-app purchases, last update, 5-star ratings, and the number of reviews were collected in Excel. When the apps were duplicated on Android and iOS, the app on the platform with the largest number of reviews was selected and analyzed.

First, we screened out the apps that were unrelated to this study based on the title and information. Next, we evaluated the eligibility of the remaining apps. We excluded those that (1) had a number of reviews <100; (2) were related to puzzles, math, games, and things for fun only; (3) addressed mood disorder or mental health, such as cognitive behavioral therapy; (4) had not been updated since 2020; (5) were related to language or speech therapy; (6) proposed an IQ test; and (7) were not written in English.

2.2. Domain analysis

We analyzed four cognitive domains: memory, attention, executive function, and visuospatial function. Memory is defined as the ability to store, recall, utilize, and manipulate specific information. Attention is defined as the ability to pay attention to external objects instantaneously, maintain attention for a long time, or move attention from one place to another (Powell, 2017; Harvey, 2019; Cramer et al., 2023). The executive function consists of problem planning, implementation, inhibition of behavior, and flexibility of thinking and is defined as the ability to control, regulate, and manage various cognitive processes (Diamond, 2013). The visuospatial function is defined as the ability to interpret and organize visual perception data (Esposito et al., 2021). Because not all apps included training for the four domains of cognition, we analyzed which domains each app provided.

Before the evaluation, we reviewed the definition and test methods for the four domains of cognition and checked which domains of training each app provides (Powell, 2017). The apps were analyzed independently by two reviewers (MHB and CWJ) after each app was used for >30 min; when a disagreement occurred, the reviewers found a consensus through discussions with a third party (HEC).

Data were collected to determine which of the four domains each app provided training on. Subsequently, the frequency and percentage of the apps that provided a certain domain were obtained and were also calculated according to the type of domains provided by each app.

2.3. Quality analysis

When assessing the quality of mobile applications, it's crucial to take into account various criteria such as usability, functionality, security, and reliability. These factors enable us to determine how well-designed and effective an app is, and whether it's appropriate for its intended use. The Mobile Application Rating Scale (MARS) is a widely used mHealth app quality assessment tool that includes a multidimensional measure of engagement, functionality, aesthetics, information, as well as app subjective quality (Stoyanov et al., 2015; Terhorst et al., 2020). In contrast, previous quality evaluation tools had one-dimensional measures, such as the Perceived Useful and Ease of Use questionnaire (PUEU), which only evaluated usability (Price et al., 2016), or Abbott's scale, which only assessed interactivity (Van Singer et al., 2015). As we wanted a multidimensional analysis in our study, we used MARS for our evaluation.

In this analysis, we excluded the app subjective quality score of Section E and used the app quality mean score, which averages the score for the four dimensions (19 questions) from Sections A to D. The exclusion of the subjective quality score from the app quality mean score due to its subjective nature strengthens the objectivity of the MARS as a measure of app quality (Stoyanov et al., 2015). All items were rated on a 5-point scale (5 = *excellent*, 4 = *good*, 3 = *acceptable*, 2 = *poor*, 1 = *inadequate*). Each section and its corresponding items are as follows:

Section A. Engagement (Five items: entertainment, interest, customization, interactivity, and target group).

Section B. Functionality (Four items: performance, ease of use, navigation, and gestural design).
Section C. Aesthetics (Three items: layout, graphics, and visual appeal).
Section D. Information (Seven items: accuracy of app description, goals, quality of information, quantity of information, visual information, credibility, and evidence of base).

In addition, apps were analyzed independently by two reviewers (MHB, CWJ) after using each app for >30 min, and when a disagreement arose, they reached a consensus through discussions with a third party (HEC).

2.4. Correlation analysis of the MARS, 5-star ratings, and reviews

The relationship between app quality mean score, 5-star ratings, and the number of reviews was analyzed, and the number of reviews was applied as a log value to match the scale with the MARS score. The correlation with the MARS, ratings, and reviews was confirmed using a scatter plot, and the degree of correlation was analyzed through correlation analysis.

2.5. Trend analysis of the MARS and the number of domains

The apps were divided into four groups of 1, 2, 3, and 4, according to the number of domains of training provided by each app. We analyzed whether differences and trends in the MARS existed between each group.

2.6. Statistical analyses

The relationships between the MARS, 5-star ratings, and reviews were analyzed using Pearson's correlation analysis. The Jonckheere-Terpstra test was performed to confirm the difference in the MARS according to the number of domains, and the Mann-Whitney U test was performed as post hoc analysis. Data were analyzed using SPSS (version 21.0, IBM SPSS Statistics), and scatter plots were visualized using R (version 4.1.2). Statistical significance was set at $P < .05$.

2.7. Ethical approval

This study did not involve human subjects.

3. Results

3.1. General characteristics of the apps

A total of 708 apps (Google Play Store, $n = 500$(cognitive training, $n = 250$; cognitive rehabilitation, n = 250); Apple App Store, $n = 208$ (cognitive training, $n = 200$; cognitive rehabilitation, $n = 8$)) were identified via search terms, of which 83 duplicated apps were excluded. Thus, 625 apps were preliminarily screened, and 282 irrelevant apps were excluded. The remaining 343 relevant apps were screened according to the exclusion criteria; finally, 53 apps (Google Play Store, $n = 33$, 63 %; Apple app store, $n = 5$, 9 %; both platforms, $n = 15$, 28 %) were included in the analysis (Fig. 1). Forty-six (87 %) apps had the in-app purchase functionality, and seven (13 %) did not. The distribution of apps in the last update period was as follows: <1 month ($n = 16$, 30 %), 1–6 months (22, 42 %), 7–12 months (9, 17 %), and >12 months (6, 11 %). The distribution of apps regarding the number of reviews was <1000 ($n = 18$, 34 %), 1000 to 10,000 (12, 23 %), and >10,000 (23, 43 %). For ratings, five (9 %) were 3.00 to 4.00, and 48 (91 %) were >4.00 (Table 1).

3.2. Domains of the apps

The analysis of the domains of training provided by each app revealed that 52 (98 %) of the 53 apps included memory function, 48

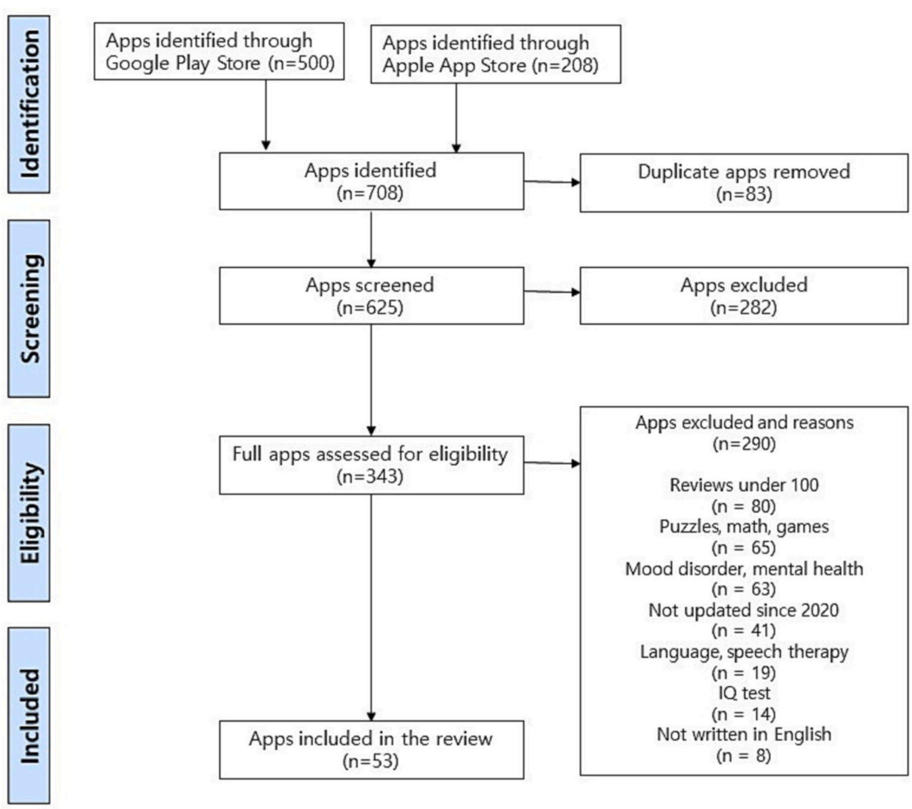

Fig. 1. Study flow diagram.

Table 1
General characteristics of the apps.

Characteristics	Apps, n (%)
Platform	
Google play store	33 (63 %)
Apple App store	5 (9 %)
Both	15 (28 %)
In-app purchases	
Yes	46 (87 %)
No	7 (13 %)
Last update	
<1 month	16 (30 %)
1–6 months	22 (42 %)
7–12 months	9 (17 %)
>12 months	6 (11 %)
Reviews	
<1000	18 (34 %)
1000–10,000	12 (23 %)
>10,000	23 (43 %)
Ratings	
3.00–4.00	5 (9 %)
>4.00	488 (91 %)

(91 %) included attention function, 24 (45 %) included executive function, and 19 (36 %) included visuospatial function.

The apps were classified according to the domains of training included in each app: memory function was referred to as M, attention function as A, executive function as E, and visuospatial function as V. Out of a total of 53 apps, 16 (30 %) included M + A + E, 15 (28 %) M + A, nine (17 %) M + A + V, eight (15 %) M + A + E + V, three (6 %) M, one (2 %) M + V, and one (2 %) V (Fig. 2).

3.3. MARS score, reviews, and 5-star ratings of the apps

Table 2 shows the title (developer), cognitive domains, engagement, functionality, aesthetics, information, mean quality score, 5-star ratings, and reviews of 53 apps.

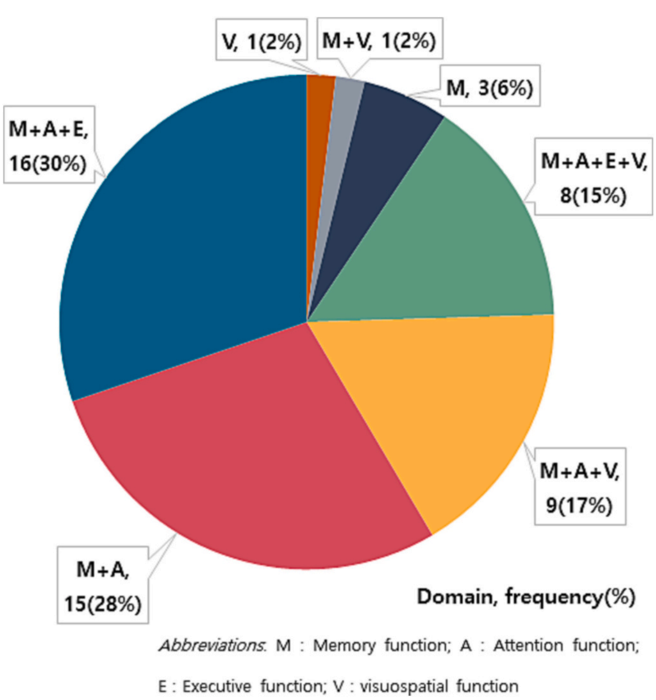

Fig. 2. Classification of the apps by domains.

The mean quality score across all apps had a mean of 3.09 (SD 0.61), 5-star ratings a mean of 4.33(SD 0.30), and reviews a mean of 62,415.43 (SD 121578.77). On average, the highest rated section was functionality (mean = 3.18, SD = 0.62), followed by aesthetics (mean = 3.13, SD = 0.72), and information (mean = 3.11, SD = 0.54), while the lowest rated section was engagement (mean = 2.97, SD = 0.68; Table 3).

3.4. Correlation between the MARS, 5-star ratings, and reviews

Correlation analysis revealed a statistically significant association between the MARS and reviews (r = 0.447 and P = .001*). Ratings and reviews showed r = 0.120, and ratings and the MARS showed r = 0.129; however, neither showed a statistically significant correlation (Fig. 3).

3.5. Trend between the MARS and the number of domains

The apps were divided into four groups according to the number of training domains covered by each app. We also analyzed whether differences and trends in the MARS existed between each group. The groups with one domain were n = 4 (mean 2.57, SD 0.47); those with two domains were n = 16 (mean 2.78, SD 0.46); those with three domains were n = 25 (mean 3.31, SD 0.59); and those with four domains were n = 8 (mean 3.31, SD 0.67; Fig. 4, Table 4). As the number of domains increased, MARS showed a statistically significant increasing trend (P = .002*). However, post hoc analysis showed statistically significant differences only between the groups with one and three domains (P = .025*) and between those with two and three domains (P = .004*).

4. Discussion

In this study, we identified 53 commercial cognitive training or rehabilitation apps and systematically analyzed their content based on four domains. Our analysis revealed that the majority of the apps offered training for memory and attention domains, but few apps included executive function or visuospatial domains. The quality of the apps, evaluated by the MARS score, improved significantly when more domains were provided, and was positively associated with the number of reviews received.

In recent years, there has been a rise in the popularity of digital therapeutics. Some digital therapeutics, such as Pear Therapeutics' reSET for substance use disorder and Akili Interactive's EndeavorRx for attention deficit hyperactivity disorder, have even gained approval from the Food and Drug Administration (FDA) (Kollins et al., 2020; Maricich et al., 2022). With the increasing use of smartphones globally and the transition to telemedicine due to the COVID-19 pandemic, mobile apps are becoming more prominent as a type of digital therapeutics. This has led to the exploration of cognitive training through mobile apps as a promising alternative to traditional face-to-face training with occupational therapists (Abbadessa et al., 2022). However, despite the growing interest in these digital therapeutics, there is currently a lack of systematic analysis on cognitive apps, which prompted the development of this study to examine the current status of cognitive training apps utilized in the commercial market.

Our results showed that most apps contained memory (98 %) and attention (91 %) domains. The decline in memory function is the most prominent feature of Alzheimer's dementia, and memory loss problems are frequently observed in vascular dementia (Katzman, 1986; Braaten et al., 2006). Disruption of attention is another common symptom in patients with brain disorders, such as hypoxic brain injury (Anderson and Arciniegas, 2010). For these reasons, most developers develop and design apps that focus on aspects of memory and attention training; in contrast, apps that contained training for executive or visuospatial function accounted for less than half of the total apps, and only eight (15 %) included training for all four domains. Patients with abnormal

Table 2
MARS scores, 5-star ratings, and reviews of the evaluated apps.

Title (developer)	Cognitive domains	Engagement	Functionality	Aesthetics	Information	Mean quality score	5-Star ratings	Reviews
Peak (PopReach Incorporated)	M, A, E	4.20	4.25	4.33	3.86	4.16	4.00	496,366
Skills (App Holdings)	M, A, V	3.00	3.25	3.00	3.17	3.10	4.30	446,885
Elevate (Elevate Labs)	M, A	3.80	4.00	3.67	3.67	3.78	4.60	399,406
Neuronation (NeuroNation)	M, A, E	4.40	4.25	4.33	4.14	4.28	4.50	364,616
Lumosity (Lumos Labs, Inc.)	M, A, E	4.20	4.25	4.00	3.83	4.07	4.50	268,485
Brain Training (App Holdings)	M, A	3.20	2.75	3.00	3.33	3.07	4.30	238,088
Mind Games (Mindware Consulting, lnc)	M, A, E, V	3.40	3.50	3.33	3.33	3.39	4.40	188,050
Impulse - Brain Training (GMRD Apps Limited)	M, A, E	4.20	4.00	4.33	3.83	4.09	4.70	143,541
Reaction training (nixGames)	M, A, V	2.80	3.00	2.33	2.50	2.66	4.70	124,977
Left vs Right (MochiBits)	M, A, E	3.80	4.00	4.33	3.83	3.99	4.50	109,457
Brain Games (MagiqLab)	M, A	2.20	2.50	2.00	2.33	2.26	4.50	74,808
Memory Games (Maple Media)	M, A	2.80	3.00	3.00	2.83	2.91	4.20	74,940
Brain Games (Godline Studios)	M, A, V	2.60	2.50	2.33	2.67	2.53	4.20	62,515
Train your Brain (Senior Games)	M, A, E, V	3.80	4.00	4.33	3.50	3.91	4.50	57,391
MindPal (Elektron Labs Inc.)	M, A, E	3.80	4.00	4.00	3.83	3.91	4.50	37,819
Super Brain Training (Godline Studios)	M, A, E	2.40	2.75	2.67	2.83	2.66	4.10	35,424
Train your brain (Grove FX)	M, A, V	2.20	2.75	2.33	2.33	2.40	4.30	27,798
Neurobics (Peoresnada.com)	M, A, E	3.40	3.75	4.00	3.83	3.75	4.00	24,886
Smarter (Laurentiu Popa)	M, A, V	3.20	3.25	3.33	3.33	3.28	4.10	20,163
Brain Yoga Brain Training Game (Megafauna software)	M, V	2.40	2.75	3.00	2.50	2.66	4.60	17,977
Brain Training Game (LogicLike)	M, A, E, V	3.40	3.75	3.67	3.17	3.50	4.50	18,088
Big Brain (PocketLand)	M, A, V	2.80	3.25	3.33	3.33	3.18	4.10	15,492
Brain Puzzle Games for Adults (Dmitriy Yudin)	M, A, E	3.80	4.00	4.00	3.67	3.87	4.30	14,325
Memory Games (AppQuiz)	M, A	2.40	2.50	2.67	2.33	2.48	4.20	9941
CogniFit (CogniFit Inc)	M, A, E, V	2.60	2.75	2.67	2.83	2.71	4.30	7483
Brain Games (XGAME9X)	M, A	2.20	2.25	2.33	2.33	2.28	3.90	3609
Brainwell Mind & Brain Trainer (Monclarity, LLC)	M, A, E, V	4.20	4.00	4.33	3.67	4.05	4.60	3327
Concentration training (nixGames)	M, A	2.60	3.25	2.67	3.17	2.92	4.60	3685
Brain Training (JamJam Game Studio)	M, A	2.80	2.75	2.67	2.33	2.64	4.20	2386
Abrain (Oleksandr Balias)	M, A	3.40	3.50	3.67	3.33	3.48	4.70	2962
Brainpower (Branded Brothers)	M, A	2.80	2.50	2.67	2.50	2.62	4.50	1522
Neuriva Brain Gym (Reckitt Benckiser Group plc)	M, A, V	2.80	2.75	2.67	2.83	2.76	4.40	1279
KettleMind (Happy Adda Studios Pvt. Ltd)	M, A, E, V	4.00	4.00	4.00	3.83	3.96	4.10	1186
Excercises for the brain (ABC Programming)	M, A	3.20	3.75	3.33	3.33	3.40	4.10	1182
Brain Games:IQ test & puzzle (Content Arcade Ltd.)	M, A, E	3.20	3.25	3.67	3.50	3.40	4.60	1097
N-Back memory training (E.A.L)	M	2.00	2.50	2.67	2.86	2.51	4.60	788
Constant Therapy (Constant Therapy Health, Inc.)	M, A, E	3.20	3.75	3.33	4.00	3.57	3.80	755
Brain Games - Logic puzzles (WL Pixign Games)	M, A, E	3.00	3.00	3.33	3.33	3.17	4.10	622
Brain Trainer: Tune your brain (Anton Vinokurov)	M, A, E	2.60	3.00	2.33	2.67	2.65	4.60	588
BrainHQ (Posit Science)	M, A, E	2.40	2.50	3.00	2.83	2.68	3.10	499
BrainZ (Brain Train Games Studio)	M, A	2.40	2.75	2.00	2.50	2.41	4.20	463
Cerebrum: Brain Training Game (Wil Corp. Software)	M, A	2.00	2.50	2.00	2.33	2.21	4.30	383
Brain Training Games (iq test)	M, A, E	2.60	3.00	3.00	2.67	2.82	4.10	396
CleverMe (Wowmaking)	M, A, E	3.60	3.75	4.00	3.50	3.71	4.20	369
Memory Games-memory training (Softser)	M	2.00	2.75	2.67	2.83	2.56	5.00	295
Visual Attention Therapy (Tactus Therapy Solutions Ltd)	V	2.60	3.25	3.00	3.83	3.17	4.70	308
Brainschool (Brain Academy)	M, A, E, V	2.40	2.50	2.67	2.83	2.60	4.10	259
Brain test (Almaz Kamaletdinov)	M	1.80	2.00	2.00	2.33	2.03	4.20	251
Brain Games (Train your brain)	M, A	2.60	2.75	2.67	2.67	2.67	4.60	222
Find in Mind Brain Training (Weez Beez)	M, A, V	2.80	3.00	3.00	3.00	2.95	4.40	189
BrainGym:Impulse brain games (Saharapixels)	M, A, V	2.80	3.50	3.00	3.00	3.08	4.60	192
Memory Games (Flutterayu)	M, A	2.40	3.00	3.00	2.67	2.77	4.30	160
RecoverBrain (ImagiRation LLC)	M, A, E, V	2.00	2.25	2.00	3.17	2.35	4.10	123

Abbreviations: M: Memory function; A: Attention function; E: Executive function; V: visuospatial function.

executive function have difficulties with planning, initiation, organization, inhibition, problem-solving, self-monitoring, and error correction (Chung et al., 2013), which means that they may have problems with functional independence. Visuospatial processing is a fundamental aspect of human cognition that has a significant impact on an individual's identity and function. In other words, since various domains affect cognition, it is necessary to accurately identify the area where the problem arises and choose an app that includes training on the patient's needs. Using inappropriate apps is not only a waste of time, but it may also harm patients. However, it is difficult for patients and caregivers to know what kind of treatment is needed, and healthcare professionals must actively intervene to help patients choose which apps to use.

The average MARS quality score was 3.09. This result was similar to the MARS scores indicated by other researchers to evaluate health apps for other conditions; the average MARS of 3.17 regarded apps evaluated for pain management (Salazar et al., 2018), 3.03 for apps evaluated for hematologic conditions (Narrillos-Moraza et al., 2022), and 3.08 for apps evaluated for the care partners of people with dementia (Werner

Table 3
Descriptive statistics of the MARS scores, 5-star ratings, and reviews.

Category	Mean (SD)	Median	Minimum-maximum
Engagement	2.97 (0.68)	2.80	1.80–4.40
Functionality	3.18 (0.62)	3.00	2.00–4.25
Aesthetics	3.13 (0.72)	3.00	2.00–4.33
Information	3.11 (0.54)	3.17	2.33–4.14
Mean quality score	3.09 (0.61)	2.95	2.03–4.28
5-Star ratings	4.33 (0.30)	4.30	3.10–5.00
Reviews	62,415.43 (121,578.77)	3609	123–496,366

et al., 2022). Considering a score of 3 as acceptable Only five (8.5 %) apps exceeded four points in the overall score, and more than half (27/53, 50.9 %) scored <3. This suggests that many of these apps did not meet acceptable criteria (Stoyanov et al., 2015) for quality. Therefore, it is important for health professionals to provide accurate evaluation and feedback when selecting cognitive training apps for their patients.

The section with the lowest score was engagement, which evaluated entertainment, interest, customization, interactivity, and target group. The average score for the engagement section was 2.97, which was lower than the acceptable score. A low engagement score can negatively affect the frequency and duration of using an app, consequently leading users to cease using it (Short et al., 2018). In particular, engagement in cognitive training or rehabilitation apps is important because they should be used for a long time to determine their efficacy (Bright et al., 2015). Attempts to increase engagement with the app, such as real-time feedback technology or game-style design, are required in the future (Edney et al., 2019).

We attempted to determine which factors were related to the quality of the apps. The MARS quality score significantly increased along with the number of domains. As previously stated, Alzheimer's and vascular dementia imply a problem not only with general cognition or memory function but also with various cognitive domains (Fahlander et al., 2002; Stopford et al., 2012). Developers who made apps that covered several cognitive domains were likely to have in-depth considerations, which would have enabled them to produce high-quality apps. However, even if apps offer training in various domains, it cannot be said that their quality is unconditionally good. For example, "Recover brain" contained all four domains, but its MARS score was only 2.23, and "Visual attention therapy" contained one domain of the visuospatial function and could be useful in visuospatial neglect patients, but its MARS score was 3.17.

The number of reviews showed a statistically significant positive correlation with the MARS quality score. Consumers prefer and use apps with high user convenience; they might write a review only if they are satisfied with the app, and because a degree of effort was put into writing a review, this indicates a certain degree of reliability. However, the 5-star ratings did not show a significant correlation with the MARS. Previous studies also showed a difference between quality evaluation using an objective tool such as the MARS and the real-world user evaluation through star ratings (Machado et al., 2016; Domnich et al., 2016). One reason was that the user ratings of app stores were sometimes derived from pilot reviews or paid autonomous programs deployed by the developer (Zhu et al., 2014). In addition, this might be because it is too easy for users to give 5-star ratings as they can provide star ratings without even trying the app.

In summary, this review suggests that patients looking to use cognitive training apps for self-management must be aware that many apps focus only on memory and attention and do not cover many different areas. It is difficult for patients to know which apps are suitable for use because the content and quality of mobile health apps are

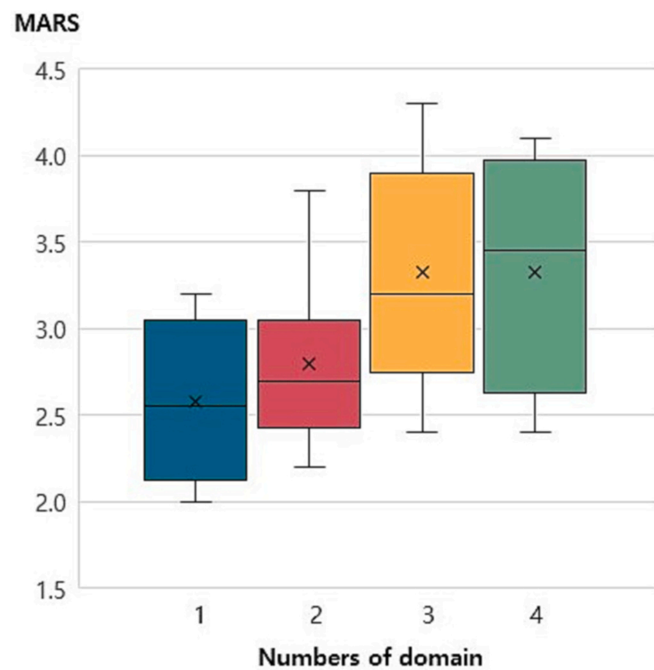

Fig. 4. Trend between the MARS and the number of domains.

Table 4
Descriptive statistics of the MARS score by the number of domains.

Number of domains	Frequency (%)	Mean (SD)	Median	Minimum-maximum
1	4 (8)	2.57 (0.47)	2.53	2.03–3.17
2	16 (30)	2.78 (0.46)	2.67	2.21–3.78
3	25 (47)	3.31 (0.59)	3.18	2.40–4.28
4	8 (15)	3.31 (0.67)	3.44	2.35–4.05

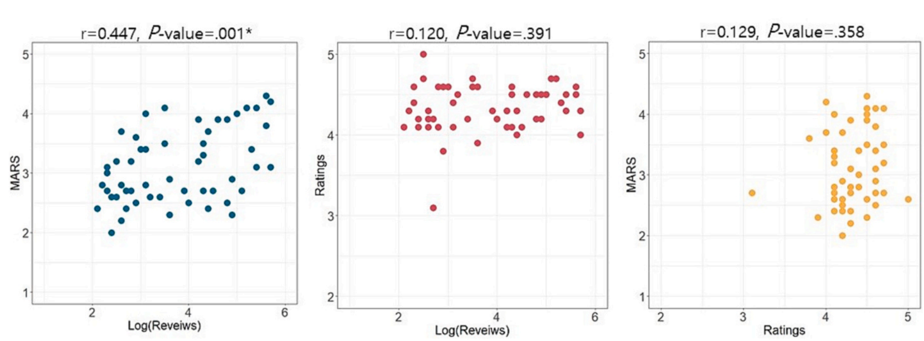

Fig. 3. Correlation between the MARS, 5-star ratings, and reviews.

currently poorly regulated (McCartney, 2013).

The strengths of this study lie in the detailed analysis of cognitive training app content and quality. The results enable users to identify the domains that an app covers and assess its quality, which can guide patients and healthcare providers in selecting appropriate apps. Additionally, these findings may be useful for developers looking to create more effective cognitive training apps.

5. Limitations

This study has some inherent limitations. First, only apps available in the Android Play Store and Apple App Store, with contents only in English and accessed from a US IP address, were included. Thus, we assume the possibility of having missed other apps dedicated to cognitive training or rehabilitation. Second, our study is that we exclusively relied on the Mobile Application Rating Scale (MARS) to assess app quality. It is important to note that there are other assessment tools available, and depending on which tool is used, different results may be obtained. However, given that MARS is currently the most commonly used tool for evaluating app quality, we believe that our results are likely to be widely applicable and informative.

Another limitation of our study is that only four domains were included and analyzed. Domains related to language, speech, and social cognition were excluded because they were usually treated as independent, separate apps. However, because these are included in cognition, it might be a good idea to consider including them in future evaluations. In addition, each domain could be divided into subdomains, such as memory into working, episodic, and procedural, and attention into subdomains, such as selective and sustained; however, we did not analyze these subdomains.

Last, it should be noted that the apps were only evaluated after using them for just over 30 min. While this may have been enough time to evaluate the domains and quality of the apps, it is possible that more extensive use could reveal additional strengths or weaknesses. Additionally, it should be noted that our evaluation did not assess the therapeutic efficacy of the apps, which is an important consideration for their potential use in clinical settings. Mobile apps as digital therapeutics have various advantages, such as accessibility and cost-effectiveness, and many physicians expect them to replace some of the traditional rehabilitation programs (Abbadessa et al., 2022). At the same time, however, there are concerns about whether cognitive therapy using mobile apps is indeed effective (Ge et al., 2018). Several studies have been conducted on this topic. Bonnechère showed that an app named “Peak Brain Training,” which was designed to train cognitive function, could improve cognitive performance in adults (Bonnechère et al., 2021). Another study showed that an app named “Constant Therapy” could help improve language and cognition in patients with traumatic brain injury (Des Roches et al., 2015). However, most apps analyzed in this study are designed to improve certain domains in people with normal cognitive function, it is doubtful whether they will be indeed beneficial for patient with cognitive impairment.

6. Conclusion

Most apps for cognitive training available in commercial markets provided training for the memory and attention domains, and their MARS score significantly increased according to the number of domains. Developers could increase quality by incorporating several cognitive domains, but the number of domains alone cannot guarantee quality; thus, further efforts to improve quality and encourage patient participation are required. The key findings of this systematic research could be useful for the future development of mobile apps for cognitive training.

Abbreviations

MARS	mobile app rating scale
FDA	Food and Drug Administration
US	United States

CRediT authorship contribution statement

MHB, CWJ, and HEC were responsible for the study conception and design of the work. MHB, CWJ, and HEC collected data and MHB, CWJ performed data analyses. MHB and CWJ drafted the manuscript, HSK, JHP and HEC contribute ed. to the interpretation of the results and critical revision of the manuscript for important intellectual content and approved the final version of the manuscript. MHB and CWJ contributed equally to this work. All authors read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abbadessa, G., et al., 2022. Digital therapeutics in neurology. J. Neurol. 269 (3), 1209–1224.
Anderson, C.A., Arciniegas, D.B., 2010. Cognitive sequelae of hypoxic-ischemic brain injury: a review. NeuroRehabilitation 26 (1), 47–63.
Bahar-Fuchs, A., Clare, L., Woods, B., 2013. Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer’s disease and vascular dementia. Cochrane Database Syst. Rev. 6, CD003260.
Bahar-Fuchs, A., et al., 2019. Cognitive training for people with mild to moderate dementia. Cochrane Database Syst. Rev. 3, CD013069.
Bonnechère, B., et al., 2021. Brain training using cognitive apps can improve cognitive performance and processing speed in older adults. Sci. Rep. 11 (1), 1–11.
Braaten, A.J., et al., 2006. Neurocognitive differential diagnosis of dementing diseases: Alzheimer’s dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. Int. J. Neurosci. 116 (11), 1271–1293.
Bright, F.A., et al., 2015. A conceptual review of engagement in healthcare and rehabilitation. Disabil. Rehabil. 37 (8), 643–654.
Chung, C.S., et al., 2013. Cognitive rehabilitation for executive dysfunction in adults with stroke or other adult non-progressive acquired brain damage. Cochrane Database Syst. Rev. 4, CD008391.
Cramer, S.C., et al., 2023. Cognitive deficits after stroke. Stroke 54 (1), 5–9.
Des Roches, C.A., et al., 2015. Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. Front. Hum. Neurosci. 8, 1015.
Diamond, A., 2013. Executive functions. Annu. Rev. Psychol. 64, 135–168.
Domnich, A., et al., 2016. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. BMC Med. Inform. Decis. Mak. 16, 83.
Edney, S., et al., 2019. User engagement and attrition in an app-based physical activity intervention: secondary analysis of a randomized controlled trial. J. Med. Internet Res. 21 (11), e14645.
Esposito, E., Shekhtman, G., Chen, P., 2021. Prevalence of spatial neglect post-stroke: a systematic review. Ann. Phys. Rehabil. Med. 64 (5), 101459.
Fahlander, K., et al., 2002. Cognitive functioning in Alzheimer’s disease and vascular dementia: further evidence for similar patterns of deficits. J. Clin. Exp. Neuropsychol. 24 (6), 734–744.
Foundation, S., 2017. Clinical Guidelines for Stroke Management 2017.
Furlong, L.M., et al., 2016. Quality of mobile phone and tablet mobile apps for speech sound disorders: protocol for an evidence-based appraisal. JMIR Res. Protoc. 5 (4), e233.
Ge, S., et al., 2018. Technology-based cognitive training and rehabilitation interventions for individuals with mild cognitive impairment: a systematic review. BMC Geriatr. 18 (1), 213.
Harvey, P.D., 2019. Domains of cognition and their assessment. Dialogues Clin. Neurosci. 21 (3), 227–237.
Jokinen, H., et al., 2015. Post-stroke cognitive impairment is common even after successful clinical recovery. Eur. J. Neurol. 22 (9), 1288–1294.
Kalaria, R.N., Akinyemi, R., Ihara, M., 2016. Stroke injury, cognitive impairment and vascular dementia. Biochim. Biophys. Acta (BBA) - Mol. Basis Dis. 1862 (5), 915–925.
Katan, M., Luft, A., 2018. Global burden of stroke. In: Seminars in Neurology. Thieme Medical Publishers.
Katzman, R., 1986. Alzheimer’s disease. N. Engl. J. Med. 314 (15), 964–973.
Kollins, S.H., et al., 2020. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial. Lancet Digit. Health 2 (4), e168–e178.

Kuehnhausen, M., Frost, V.S., 2013. Trusting smartphone apps? To install or not to install, that is the question. In: 2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). IEEE.

Leung, N.T., et al., 2015. Neural plastic effects of cognitive training on aging brain. Neural Plast. 2015.

Machado, G.C., et al., 2016. Smartphone apps for the self-management of low back pain: a systematic review. Best Pract. Res. Clin. Rheumatol. 30 (6), 1098–1109.

Maricich, Y.A., et al., 2022. Safety and efficacy of a digital therapeutic for substance use disorder: secondary analysis of data from a NIDA clinical trials network study. Subst. Abus. 43 (1), 937–942.

McCartney, M., 2013. How do we know whether medical apps work? Bmj 346.

Narrillos-Moraza, Á., et al., 2022. Mobile apps for hematological conditions: review and content analysis using the mobile app rating scale. JMIR mHealth and uHealth 10 (2), e32826.

Nguyen, L., Murphy, K., Andrews, G., 2019. Cognitive and neural plasticity in old age: a systematic review of evidence from executive functions cognitive training. Ageing Res. Rev. 53, 100912.

Nichols, E., Vos, T., 2021. The estimation of the global prevalence of dementia from 1990-2019 and forecasted prevalence through 2050: an analysis for the Global Burden of Disease (GBD) study 2019. Alzheimers Dement. 17, e051496.

Obaid, M., et al., 2020. Long-term outcomes in stroke patients with cognitive impairment: a population-based study. Geriatrics 5 (2), 32.

Powell, T., 2017. The brain injury workbook: exercises for cognitive rehabilitation. In: A Speechmark Practical Therapy Manual, Second edition. Routledge, London.

Price, M., et al., 2016. Usability evaluation of a mobile monitoring system to assess symptoms after a traumatic injury: a mixed-methods study. JMIR Mental Health 3 (1), e5023.

Recchia, G., et al., 2020. Digital therapeutics-what they are, what they will be. Acta Sci. Med. Sci. 4, 1–9.

Salazar, A., et al., 2018. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR mHealth and uHealth 6 (10), e10718.

Short, C.E., et al., 2018. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J. Med. Internet Res. 20 (11), e292.

Srikanth, V.K., et al., 2003. Increased risk of cognitive impairment 3 months after mild to moderate first-ever stroke: a community-based prospective study of nonaphasic English-speaking survivors. Stroke 34 (5), 1136–1143.

Stopford, C.L., et al., 2012. Working memory, attention, and executive function in Alzheimer's disease and frontotemporal dementia. Cortex 48 (4), 429–446.

Stoyanov, S.R., et al., 2015. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 3 (1), e27.

Terhorst, Y., et al., 2020. Validation of the Mobile Application Rating Scale (MARS). PLoS One 15 (11), e0241480.

Unverzagt, F.W., et al., 2001. Prevalence of cognitive impairment: data from the Indianapolis Study of Health and Aging. Neurology 57 (9), 1655–1662.

Van Singer, M., Chatton, A., Khazaal, Y., 2015. Quality of smartphone apps related to panic disorder. Front. Psychiatry 6, 96.

Werner, N.E., et al., 2022. Quality of mobile apps for care partners of people with alzheimer disease and related dementias: mobile app rating scale evaluation. JMIR Mhealth Uhealth 10 (3), e33863.

Winstein, C.J., et al., 2016. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/ American Stroke Association. Stroke 47 (6), e98–e169.

Zhu, H., et al., 2014. Discovery of ranking fraud for mobile apps. IEEE Trans. Knowl. Data Eng. 27 (1), 74–87.