DOI:10.6129/CJP.202209_64(3).0002
中華心理學刊 民 111，64 卷，3 期，279-310
Chinese Journal of Psychology 2022, Vol.64, No.3, 279-310

Development of a Tablet-Based Task Battery for Executive Function Assessment

Yei-Yu Yeh[1], Ting-Yu Lin[1], Tsung-Ren Huang[1], and Chun-Yu Kuo[2]
Department of Psychology, National Taiwan University [1]
Department of Adult and Continuing Education, National Taiwan Normal University [2]

The aim of this study was to develop a theory-driven, tablet-based battery of executive function tasks that is free of charge. According to a theoretical framework, three levels of executive function can be identified. The basic level consists of working memory and inhibitory control, the second level focuses on cognitive flexibility, and the third level includes reasoning, planning, and problem solving. Given that the functions at the third level require effective processing at the first two levels, this battery measured individual abilities on executing the latter functions. Additional functions included executive controls that have been suggested to be important. The battery consisted of 17 tasks assorted into six categories: processing speed, information maintenance, inhibition, updating, switching and resource allocation. Processing speed was measured by the speed in simple and choice reaction times. Information maintenance included maintenance of single items and paired associates. Three types of representations were used for single items: spatial information, symbols that are difficult to name, and objects that can be interfaced with the representations in long-term memory. A delayed matching-to-sample task and a delayed response task were also adopted to assess the information maintenance ability. The former task evaluated maintenance over 5 and 10 seconds compared with perceptual matching; the latter examined the effects of encoding and retrieving one or four items. Inhibition functions were assessed by three tasks: the reverse Stroop task, stop signal task, and anti-saccade task. Updating functions were also evaluated by three tasks: running memory of spatial information, running memory of symbols, and running memory of objects. The trail-making test, a figure task, and a hearts-and-flowers task were used to measure the ability to switch for cognitive flexibility. These three tasks differ in the dimensions for switching: between color and digits, between task rules, and between color and shapes. A rotation span task was adopted to assess the ability to allocate cognitive resources. In the battery, 10 tasks adopted accuracy and reaction time as the measures, and 7 adopted a 1-up-1-down staircase method to measure the span of information maintenance, updating, and resource allocation. Two hundred and eight college students participated in the study across two or three sessions of approximately 50 minutes each, with a rest between tasks. The pattern of results was in agreement with the findings in the literature, supporting the use of a tablet-based battery that employed a staircase method for measuring working memory span. Given the limitations in the current version, a recommendation of the context for using this battery is provided in the discussion. Future studies could consider measuring the abilities of processing linguistic information, evaluating each executive function with a broader range of tasks, and using an accuracy-based battery. More importantly, future studies should examine the reliability and validity of this battery while developing norms for different age levels.

Keywords: *executive function, tablet-based battery*

Extended Abstract

Executive functions (EFs) play essential roles in the performance of goal-directed behaviors and enable people to adapt to novel situations. Different disciplines have investigated this psychological construct using various operational definitions, as no gold standard exists for its measurement. To help naïve learners gain a basic understanding of EFs via hands-on experience, this study developed a battery of laboratory tasks that assess important components of EFs. The task battery, the EFTB, is theory-driven, tablet-based, and free of charge. It consists of a series of tasks targeting processing speed, the retention of stimulus representations in the working memory, inhibition, updating, shifting, and resource allocation between two sequential tasks. This paper describes the conceptualization and development of this battery, our data collection and analyses to validate the battery against past findings, the results of these analyses, the limitations of this study, and future research directions. The EFTB was programmed using Javascript, was generated as a tablet application through Cordova, and can be installed on Apple products with iOS 12 or higher versions. The EFTB is freely available to qualified users upon request at http://mil.psy.ntu.edu.tw/eftb/. The accompanying user manual provides detailed instructions for each task in terms of test administration, the design of conditions, and data construction.

We adopted Diamond's (2013) model of EFs as the core theoretical foundation of our battery to focus on working memory and cognitive flexibility. We also included tasks targeting the EFs of updating information, accessing long-term memory, and allocating cognitive resources. These three EFs are considered important by other researchers. Furthermore, considering that processing speed may covary with EFs, we included a task on processing speed. In total, the battery consists of six categories of different laboratory tasks. Category 1 assesses information processing speed and consists of the simple reaction time task (T1) and choice reaction time task (T2). Category 2 assesses the ability to maintain different types of information in the working memory and consists of the spatial memory task (T3), symbol–location associative memory task (T4), object–location associative memory task (T5), delayed matching-to-sample task (T6), and delayed response task (T7). Category 3 assesses inhibition ability and consists of the reverse Stroop task (T8), stop signal task (T9), and antisaccade task (T10). Category 4 measures the ability to update information and consists of the running memory of locations task (T11), running memory of symbols task (T12), and running memory of objects task (T13). Category 5 tests the ability to shift and consists of the color trail making task (T14), hearts and flowers task (T15), and figures switching task (T16). Category 6 measures the ability to allocate cognitive resources between two sequential tasks and consists of the rotation span task (T17). Table 1 describes the cognitive processes underlying each task and their measurement.

Table 1

Tasks, index, and the related processing in each cognitive domain

Task	Index		General Cognitive Processes			Executive Function				
	Accuracy and Reaction Times	Capacity	Maintain One Task Goal	Maintain Two Task Goals	spatial Processing/ Pairs	Short-term Storage	Access to Long-term Memory	Updating	Switching	Conflict Resolution /Inhibition
Simple Reaction (T1)	V		V							
Choice Reaction (T2)	V		V		V					
Spatial Memory (T3)		V	V		V	V				
Symbol-Location Associative Memory (T4)		V	V		V	V				
Object-Location Associative Memory (T5)		V	V		V	V	V			

Table 1
Tasks, index, and the related processing in each cognitive domain (continue)

Task	Index		General Cognitive Processes				Executive Function			
	Accuracy and Reaction Times	Capacity	Maintain One Task Goal	Maintain Two Task Goals	spatial Processing/ Pairs	Short-term Storage	Access to Long-term Memory	Updating	Switching	Conflict Resolution /Inhibition
Delayed Matching-To-Sample (T6)	V		V			V				
Delayed Response (T7)	V		V		V	V				
Reverse Stroop (T8)	V		V							V
Stop Signal (T9)	V		V							V
Antisaccade (T10)	V		V		V					V
Running Memory Of Locations (T11)		V	V		V	V		V		
Running Memory Of Symbols (T12)		V	V			V		V		
Running Memory Of Objects (T13)		V	V			V	V	V		
Color Trail Making (T14) (Nonswitch Block)	V		V		V		V	V		V
Color Trail Making (T14) (Switch Block)	V			V	V	V	V	V	V	V
Hearts And Flowers (T15) (Nonswitch Block)	V		V		V			V		V
Hearts And Flowers (T15) (Switch Block)	V			V	V	V		V	V	V
Figures Switching (T16) (Nonswitch Block)	V		V				V	V		V
Figures Switching (T16) (Switch Block)	V			V		V	V	V	V	V
Rotation Span (T17)		V	V		V	V	V			

Methods

We recruited 208 self-described healthy adults, aged 18–29 years, at two university-based sites. All of the participants had normal or corrected-to-normal vision and were naïve to the purpose of the study. Each participant performed the 17 tasks in the same order (T1, T2, T4, T15, T10, T12, T17, T3, T9, T13, T6, T8, T7, T11, T16, T5). They performed the tasks in two to three sessions, with each session lasting about 50 mins. They were allowed to rest for 2 minutes after completing each task.

Results

Only the mean reaction times (RTs) of correct responses were included in our analyses. RTs exceeding 2.5 standard deviations (*SD*) were also excluded for each participant. Table 2 presents the descriptive statistics. The focal aim of our analyses was to validate whether the results were consistent with findings in the literature. For

Table 2
Performance in each task

Task	n	Accuracy (%)		Reaction Times (ms)		Capacity	
		M	SD	M	SD	M	SD
Simple Reaction (T1)	203	99	2	471	62	-	-
Choice Reaction (T2)	203	98	3	509	51	-	-
Spatial Memory (T3)	202	-	-	-	-	6.43	0.99
Symbol-Location Associative Memory (T4)	206	-	-	-	-	4.10	1.25
Object-Location Associative Memory (T5)	200	-	-	-	-	5.30	1.44
Delayed Matching-To-Sample (T6) (Delay 0 Second)	200	87	13	1777	324	-	-
Delayed Matching-To-Sample (T6) (Delay 5 Second)	200	75	15	2017	358	-	-
Delayed Matching-To-Sample (T6) (Delay 10 Second)	200	69	18	2238	405	-	-
Delayed Response (T7) (1-1a)	204	99	3	706	110	-	-
Delayed Response (T7) (1-4 A)	204	98	4	753	118	-	-
Delayed Response (T7) (4-1 A)	204	80	9	881	128	2.91	0.50
Delayed Response (T7) (4-4 A)	204	90	9	888	151	3.32	0.56
Reverse Stroop (T8) (Neutral)	205	99	2	639	77	-	-
Reverse Stroop (T8) (Incompatible)	205	95	6	728	137	-	-
Stop Signal (T9) (Stop Trials)	153	63	11	180	99	-	-
Antisaccade (T10)	199	94	8	484	84	-	-
Running Memory Of Locations (T11)	203	-	-	-	-	3.86	0.77
Running Memory Of Symbols (T12)	204	-	-	-	-	2.90	1.16
Running Memory Of Objects (T13)	204	-	-	-	-	3.89	1.25
Color Trail Making (T14) (Nonswitch Block)	202	97	6	8874	1994	-	-
Color Trail Making (T14) (Switch Block)	202	96	7	18780	3828	-	-
Hearts And Flowers (T15) (Nonswitch Block)	200	99	2	482	59	-	-
Hearts And Flowers (T15) (Switch Block)	200	98	2	650	103	-	-
Figures Switching (T16) (Nonswitch Block)	189	98	3	562	63	-	-
Figures Switching (T16) (Switch Block - Repetition)	189	94	6	661	124	-	-
Figures Switching (T16) (Switch Block - Switching)	189	91	8	681	148		
Rotation Span (T17)	192	-	-	-	-	3.06	1.15

the analyses conducted across tasks, a listwise method was used to include the participants who had valid data for all tasks.

Simple Reaction Time Task (T1) and Choice Reaction Time Task (T2)

The participants' performance was more accurate and faster in the simple reaction time task (0.99 ± 0.02; 471 ± 62 ms) than in the choice reaction time task (0.98 ± 0.03; 509 ± 51 ms), $t(202) = 2.76$, $p < .01$, $d = 0.19$ for accuracy, $t(202) = 11.87$, $p < .001$, $d = 0.83$ for RT.

Spatial Memory Task (T3), Symbol–Location Associative Memory Task (T4), and Object–Location Associative Memory Task (T5).

The main effect of stimulus type was significant, $F(2, 388) = 248.07$, $p < .001$, $\eta_p^2 = .56$. The Tukey's post hoc test showed that memory capacity was higher in the spatial memory task (6.44 ± 0.99) than in the object-location associative memory task (5.32 ± 1.43), which in turn was higher than that in the symbol-location associative memory task (4.12 ± 1.25).

Delayed Matching-to-Sample Task (T6).

Memory performance deteriorated with increasing delay in both accuracy, $F(2, 398) = 89.05$, $p < .001$, $\eta_p^2 = .31$, and RT, $F(2, 398) = 197.87$, $p < .001$, $\eta_p^2 = .50$.

Delayed Response Task (T7).

A 2 (memory load: 1, 4)×2 (the number of probes: 1, 4) repeated measures ANOVA was conducted for accuracy and RT, respectively. For accuracy, the main effects of memory load and the number of probes were both significant, memory load: $F(1, 203) = 641.2$, $p < .001$, $\eta_p^2 = .76$, the number of probes: $F(1, 203) = 110.6$, $p < .001$, $\eta_p^2 = .35$. The interaction also reached significance, $F(1, 203) = 179.3$, $p < .001$, $\eta_p^2 = .47$. The interaction arose as the accuracy was higher when memory load and the number of probes were compatible. For RT, the main effects of memory load and the number of probes were both significant, memory load: $F(1, 203) = 1157.51$, $p < .001$, $\eta_p^2 = .85$, the number of probes: $F(1, 203) = 76.51$, $p < .001$, $\eta_p^2 = .27$. The interaction also reached significance, $F(1, 203) = 51.08$, $p < .001$, $\eta_p^2 = .20$. The interaction arose as RT was shorter with one probe than with four probes when the memory load was one, but the simple main effect of the number of probes was not significant when the memory load was four.

Reverse Stroop Task (T8).

Accuracy was higher in the incompatible condition (0.95 ± 0.06) than in the neutral condition (0.99 ± 0.02), $t(204) = 10.35$, $p < .001$, $d = 0.72$. RT was slower in the incompatible condition (728 ± 137 ms) than in the neutral condition (639 ± 77 ms), $t(204) = 14.18$, $p < .001$, $d = 0.99$.

Stop Signal Task (T9).

RT on go trials was slower when stop trials were intermixed across trials (586 ± 95 ms) than when there were no stop trials (509 ± 68 ms), $t(152) = 15.18$, $p < .001$, $d = 1.23$. The correlation between accuracy and stop delay was significant ($r = .18$, $p < .001$).

Running Memory of Locations Task (T11), Running Memory of Symbols Task (T12), and Running Memory of Objects Task (T13).

The one-way repeated measures ANOVA showed that the main effect of stimulus type was significant, $F(2, 392) = 72.82$, $p < .001$, $\eta^2 = .27$. Memory capacity was comparable for updating location information (T11, 3.85 ± 0.72) and objects (T13, 3.91 ± 1.24), and significantly smaller for updating symbolic representations (T12, 2.92 ± 1.14).

Color Trail Making Task (T14).

RT was faster in the nonswitch block (8874 ± 1994 ms) than in the switch block (18780 ± 3828 ms), $t(201) = 41.70$, $p < .001$, $d = 2.92$. Similarly, the accuracy was higher in the non-switch block than in the switch block.

Hearts and Flowers Task (T15).

RT was faster in the nonswitch block (482 ± 59 ms) than in the switch block (650 ± 103 ms), $t(199) = 33.99$, $p < .001$, $d = 2.40$.

Figures Switching Task (T16).

The one-way repeated measures ANOVA and Tukey post hoc test showed that RT was the slowest in the switch trials of the switch block, second slowest in the nonswitch block, and fastest in the repeated trials of the switch block, $F(2, 376) = 191.0$, $p < .001$, $\eta_p^2 = .50$. The accuracy data showed the same pattern.

Rotation Span Task (T17).

Performance in the rotation span task was negatively correlated with that in the switch–switch condition of the figures switching task ($r = -.22$). It was also positively correlated with memory capacity in the spatial memory task ($r = .40$) and that in the running memory of locations task ($r = .28$).

The findings for the 17 tasks were consistent with those observed in previous studies and with our theory-based predictions. We suggest that naïve learners can use the EFTB to examine various cognitive abilities related to EFs. However, our work on the EFTB is still at a very early stage of development. Future research should investigate the reliability of the EFTB and cross-validate it with other batteries measuring EFs, such as the DKEFS (Delis, Kaplan, & Kramer, 2001) and CANTAB (Cambridge-Cognition, 1996). Moreover, further investigations are necessary to validate whether an integrated framework, such as the EFTB, can unify the many suggested subsets of EFs.