

КОГНИТИВНАЯ НАУКА

В МОСКВЕ



НОВЫЕ ИССЛЕДОВАНИЯ

МАТЕРИАЛЫ
КОНФЕРЕНЦИИ
2019

Под ред. Е.В. Печенковой, М.В. Фаликман

УДК 159.9
ББК 88.25
К57

Когнитивная наука в Москве: новые исследования. Материалы конференции 19 июня 2019 г. Под ред. Е. В. Печенковой, М. В. Фаликман. – М.: ООО «Буки Веди», ИППиП. 2019 г. – 656 стр.

ISBN 978-5-4465-2346-7

УДК 159.9
ББК 88.25

ISBN 978-5-4465-2346-7

©Авторы статей, 2019

RT_TESTS – БАТАРЕЯ КОГНИТИВНЫХ ТЕСТОВ ДЛЯ ОЦЕНКИ ИНДИВИДУАЛЬНЫХ РАЗЛИЧИЙ В СКОРОСТИ ОБРАБОТКИ ИНФОРМАЦИИ

И. А. Воронин (1), И. Д. Воронина (1), Л. А. Мерзон (2), Е. П. Масленникова (3)
ivan.a.voronin@gmail.com

1 – ПИ РАО, Москва; 2 – НИУ ВШЭ, Москва; 3 – ЮУрГУ, Челябинск

Аннотация. В рамках проекта по исследованию индивидуальных различий в скорости обработки информации у подростков была разработана новая батарея когнитивных тестов RT_tests, предназначенная для группового тестирования в школах. Батарея опирается на функционал свободного программного обеспечения (Python, PsychoPy) и объединяет фреймворк, контролирующий последовательность выполнения тестов, абстрактное определение теста и определения отдельных тестов. В первую версию батареи вошли тесты на простое время реакции, тесты на вербальную классификацию и классификацию зрительных стимулов, а также тест Струпа с двумя цветами. В рамках первого сбора данных с применением новой батареи были протестированы 160 школьников в возрасте от 10 до 16 лет. Батарея запускалась на школьных компьютерах с помощью автономной операционной системы, записанной на USB-накопитель. Среднее время выполнения батареи составило 46 минут, среднее время выполнения отдельных тестов варьировалось в пределах 5–9 минут.

Ключевые слова: когнитивное тестирование, батарея когнитивных тестов, Python, PsychoPy, индивидуальные различия, время реакции

Работа выполнена при поддержке РФФИ, грант № 17-36-01135-ОГН.

Введение

Индивидуальные различия когнитивных функций и характеристик – это актуальное направление исследований. Считается, что базовые когнитивные характеристики служат основой для более сложных когнитивных способностей (Van Der Maas et al., 2006). Исследования индивидуальных различий требуют точной оценки соответствующих признаков у множества респондентов. В настоящее время в исследованиях индивидуальных различий когнитивных характеристик, как правило, применяются нейропсихологические тестовые батареи, такие как CANTAB (Neuropsychological Test Automated Battery (CANTABeclipse) manual, 2006). Нейропсихологические батареи позволяют быстро составить когнитивный профиль респондента, однако получаемые оценки индивидуальных различий недостаточно точны. В пользу этого свидетельствует низкая ретестовая надежность тестов CANTAB (Lowe, Rabbitt, 1998). Кроме того, это платное программное обеспечение, которое нельзя модифицировать. В рамках проекта по исследованию индивидуальных различий в скорости обработки информации у подростков мы поставили задачу разработать новую

батарею когнитивных тестов, которая была бы пригодна для группового тестирования в школах.

Методика. В основу новой тестовой батареи *RT_tests* (Voronin, 2018) легло свободное программное обеспечение: Python 2.7 и PsychoPy 1.90.3. Батарея включает следующие составляющие:

1. Фреймворк, управляющий выполнением тестовой батареи в целом. Эта часть программы показывает общую инструкцию, последовательно запускает тесты из списка, записывает статус их выполнения и собирает информацию о конфигурации компьютера.
2. Абстрактное определение теста. В этой части определяются структура и логика выполнения каждого теста. Тест начинается с инструкции и демонстрации правил выполнения, затем следует набор тренировочных проб, после чего — одна или несколько основных серий. В абстрактном определении теста также описываются способ сбора ответов респондента и параметры теста.
3. Определение отдельных тестов. В этой части определяются составляющие, специфические для конкретного теста. Для большинства тестов из первой версии батареи это описание тестового экрана.

Батарея *RT_tests* — это автономное приложение, предназначенное для оценки когнитивных характеристик с минимальным вовлечением тестера. Программа отслеживает действия респондента и реагирует на них. Например, программа прерывает выполнение и демонстрирует инструкцию повторно, когда респондент дает много неверных или невалидных ответов. Для того чтобы снизить утомляемость респондентов, выполнение теста останавливается на 6 секунд через каждые 40 проб. Также действует глобальное ограничение по времени, которое прерывает выполнение батареи при переходе от одного теста к другому, если превышен лимит времени (45 минут в первой версии батареи).

Каждый тест начинается с инструкции и тренировочной серии. Инструкция сопровождается демонстрацией тестового экрана и правил выполнения теста (в том числе положение рук на клавиатуре). Тренировочная серия сопровождается обратной связью и подсказками. Для перехода к основной серии респондент должен дать достаточное количество правильных ответов.

В данный момент тестовая батарея может использоваться в двух формах. Первая форма предполагает выполнение питоновского кода в операционной системе, установленной на рабочем компьютере. В этом случае на компьютере также должны быть установлены совместимые Python и PsychoPy. Вторая форма батареи представляет собой USB-накопитель (флешку) с портативной операционной системой (Lubuntu 18.10) и установленными на нее Python, PsychoPy и файлами тестовой батареи. В этом случае используются только вычислительные ресурсы компьютера, но не его «родная» операционная система или файлы.

В состав первой версии батареи вошли следующие тесты:

1. *SRT, CRT2, CRT4* — тесты на простое время реакции. Респондент должен ответить на появление стимула в одном, двух или четырех возможных положениях.

2. *stroop2_1, stroop2_2* – тест Струпа с двумя цветами. Респондент должен дать ответ о цвете стимула. В первой серии (*stroop2_1*) предъявляется нейтральный стимул (XXXXXXX), во второй (*stroop2_2*) – слово, обозначающее цвет (красный или зеленый). Во второй серии конгруэнтные и неконгруэнтные стимулы предъявляются в равной пропорции.
3. *VerbCRT* – тест на категоризацию слов. Респондент должен дать ответ, описывает ли предъявляемое на экране слово животное или растение. Стимулы составлены из слов длиной 5, 6 и 7 букв. Количество слов сбалансировано по категориям и длине.
4. *VisCRT* – тест на категоризацию зрительных стимулов. На экране предъявляются две фигуры, которые могут различаться по форме (круг или треугольник), цвету (красная или зеленая) и заливке (залитая или контур). Также на экране предъявляется слово-ключ, описывающее признак (цвет или форма), по которому нужно сравнить фигуры. Респондент должен ответить, совпадают ли фигуры по признаку, на который указывает ключ.

Следующие тесты были реализованы, но не вошли в состав первой версии батареи:

5. *stroop3_1, stroop3_2* – тест Струпа с тремя цветами. Респондент должен дать ответ о цвете стимула (красный, зеленый или синий). В первой серии (*stroop3_1*) предъявляется нейтральный стимул (такой же, как в *stroop2_1*), во второй (*stroop3_2*) – слово, обозначающее цвет.
6. *SRT_rel, CRT2_rel* – тесты на простое время реакции в парадигме Дженсена (Jensen, Munro, 1979). В начале каждой пробы респондент должен удерживать кнопку и затем отпустить, когда стимул появится на экране.

Пробное тестирование проводилось в одной из школ города Южноуральска Челябинской области. Тестирование прошли 160 подростков в возрасте от 10 до 16 лет. Школьники выполняли тесты на школьных компьютерах. Поскольку система, установленная на компьютерах в школе (Альт Линукс 7.0.4), оказалась несовместима с PsychoPy, тестирование проводилось с помощью автономной операционной системы, записанной на флешки (объем 16Gb, в школу было направлено 10 таких флешек). Школьные компьютеры имели одноплатные процессоры Intel Celeron D 3.066 Гц, 512 Мб оперативной памяти, экраны с разрешением 1280 × 1024 и частотой обновления 60 Гц.

Результаты

В табл. 1 представлены основные характеристики тестов и данные, полученные в ходе первого тестирования. Общее количество проб, продолжительность тренировочной серии и время неответа относятся к настройкам, установленным для каждого теста перед началом тестирования. Количество ответов (*n*), время выполнения, точность ответа и среднее время ответа по пробам были вычислены на основе данных, полученных в ходе пробного тестирования.

Среднее время тестирования составило 46 минут, разброс в пределах 33–78 минут. Среднее время выполнения отдельных тестов варьировалось

в пределах 5–9 минут. Широкий разброс времени выполнения позволяет предположить, что в отдельных случаях могли происходить сбои выполнения батареи. В частности, низкое минимальное время выполнения тестов SRT, CRT2 и CRT4 (1–2 минуты) свидетельствует о том, что выполнение этих тестов было прервано раньше времени. Возможно, это связано с низким объемом оперативной памяти на школьных компьютерах: программа определила, что объем доступной оперативной памяти составил 71–164 Мб.

В свою очередь, различие в точности и среднем времени ответа соответствует различию в сложности заданий. Наибольшая точность и скорость ответа была получена для тестов на простое время реакции (SRT, CRT2, CRT4), наименьшая — для теста классификации зрительных стимулов (VisCRT).

Таблица 1. Характеристики тестовой батареи *RT_tests*

Тест	Общее кол-во проб	Продолж. трен. серии (кол-во проб)	Время неответа (сек.)	<i>n</i>	Время выполнения (мин.)				Точность ответа	Среднее время ответа (сек.)	
					<i>M</i>	<i>SD</i>	<i>min</i>	<i>max</i>		<i>M</i>	<i>SD</i>
SRT	80	5–20	3	160	5.53	0.77	1.59	10.73	0.97	0.332	0.066
CRT2	80	7–20	3	160	5.17	0.68	0.72	6.71	0.98	0.396	0.070
CRT4	80	7–20	3	160	5.31	0.48	2.14	6.65	0.95	0.513	0.112
VerbCRT	120	12–30	5	160	8.78	1.16	7.38	20.34	0.92	1.001	0.282
stroop2_1	80	12–30	5	160	5.78	0.50	4.82	7.62	0.92	0.659	0.217
stroop2_2	80	12–30	5	154	6.11	0.65	4.05	8.98	0.92	0.775	0.273
VisCRT	96	12–30	7	138	8.37	1.25	5.18	14.85	0.88	1.278	0.423

Примечание: тесты представлены в том порядке, в котором они предъявлялись в ходе тестирования. Продолжительность тренировочной серии — минимальное и максимальное количество проб в тренировочной серии, время неответа — максимальная продолжительность ожидания ответа, *n* — количество респондентов, выполнивших тест, *M* — среднее, *SD* — стандартное отклонение, *min* — минимальное значение, *max* — максимальное значение, точность ответа — доля правильных ответов.

Обсуждение и выводы

Предварительный анализ данных новой тестовой батареи *RT_tests* показал, что батарея действительно может применяться для группового тестирования в школе. Подробный анализ данных первого тестирования позволит оценить надежность тестовой батареи как средства оценки индивидуальных различий когнитивных характеристик. Помимо этого, мы планируем совершенствовать батарею, а именно:

1. Разработать систему электронного отчета (на основе R shiny).
2. Адаптировать батарею к последним версиям Python и PsychoPy.
3. Усилить безопасность данных.
4. Разработать систему автоматической пересылки данных на сервер.
5. Адаптировать батарею для использования на планшетных компьютерах.

6. Усовершенствовать отдельные тесты (например, проконтролировать встречаемость слов в VerbCRT).
7. Сформировать исполняемый (.exe) файл или установочный файл.

Потенциальные сферы приложения батареи *RT_tests* не ограничиваются исследованиями индивидуальных различий в когнитивных характеристиках. Батарея может быть адаптирована для решения самых разных задач, например для когнитивной диагностики. Интерфейс приложения позволяет изменять параметры уже существующих тестов и добавлять в батарею новые тесты.

Литература

Jensen A.R., Munro E. Reaction time, movement time, and intelligence // *Intelligence*. 1979. Vol. 3. No. 2. P. 121 – 126. doi:10.1016/0160-2896(79)90010-2

Lowe C., Rabbitt P. Test-re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues // *Neuropsychologia*. 1998. Vol. 36. No. 9. P. 915 – 923. doi:10.1016/s0028-3932(98)00036-0

van der Maas H.L.J., Dolan C.V., Grasman R.P.P.P., Wicherts J.M., Huizenga H.M., Raijmakers M.E.J. A dynamical model of general intelligence: The positive manifold of intelligence by mutualism // *Psychological Review*. 2006. Vol. 113. No. 4. P. 842 – 861. doi:10.1037/0033-295x.113.4.842

Neuropsychological Test Automated Battery (CANTABeclipse) manual. Cambridge: Cambridge Cognition Limited, 2006.

Voronin I. *RT_tests – the test battery for reaction time assessment at schools* [Электронный ресурс]. https://github.com/IvanVoronin/RT_tests

RT_TESTS: A COGNITIVE TEST BATTERY FOR THE ASSESSMENT OF INDIVIDUAL DIFFERENCES IN MENTAL SPEED

I. Voronin* (1), I. Voronina (1), L. Merzon (2), E. Maslennikova (3)

ivan.a.voronin@gmail.com

1 – Psychological Institute of Russian Academy of Education, Moscow; 2 – Higher School of Economics, Moscow; 3 – South Ural State University, Chelyabinsk

Abstract. As a part of a research project on individual differences of mental speed in adolescents, we developed a new cognitive test battery for group assessment in schools. The battery *RT_tests* relies on open-source software (Python, PsychoPy) and combines the framework for execution control, abstract definition of test logic and definitions of individual tests. The first version of the battery included simple reaction time tests, verbal and visual classification tests and a Stroop task (two colors). The first application of the battery delivered data from 160 school students aged between 10 and 16 years. The battery was run on school computers from an autonomous OS recorded on a USB-stick. The mean execution time was 45 minutes, and the average execution time of individual tests ranged between 5 and 9 minutes.

Keywords: cognitive assessment, cognitive test battery, Python, PsychoPy, individual differences, reaction time