Научная статья / Research Article
https://doi.org/10.11621/LPJ-25
УДК/UDC 159.952

Сравнение сгенерированных ИИ стимулов и фото: исследование зрительного поиска

Ф.А. Сапронов, Е.С. Горбунова ✉

Национальный исследовательский университет «Высшая школа экономики», Москва, Российская Федерация

✉ esgorbunova@hse.ru

Резюме

Актуальность. Алгоритмы ИИ быстро вошли в повседневную жизнь за последние несколько лет, что повлекло за собой появление большого количества как текстов, так и изображений, сгенерированных нейросетями. Тем не менее особенности восприятия искусственно сгенерированных изображений человеком еще не были изучены. Актуальность исследования состоит в изучении возможности использования искусственно сгенерированных изображений для планирования экспериментов в области психологии восприятия.

Цель. Изучение различий в характеристиках зрительного поиска между стимулами — фотографиями и сгенерированными ИИ изображениями.

Выборка. В исследовании приняли участие 30 человек (21 женщина) от 18 до 43 лет ($M_{возраст}$ = 21,29, SD = 5,18).

Методы. Исследование выполнено на основе экспериментального дизайна с двумя независимыми и двумя зависимыми переменными. Задача испытуемых состояла в поиске целевого стимула среди дистракторов, в качестве стимульного материала выступали изображения объектов, сгенерированные с помощью нейросети, и фотографии. Также варьировалось количество объектов на экране. Измерялись точность ответов испытуемых и время реакции.

Результаты. Различий в точности обнаружения фотореалистичных и сгенерированных ИИ стимулов обнаружено не было, как и различий во времени реакции.

Выводы. С помощью метода эксперимента в парадигме зрительного поиска было показано, что вид стимула (сгенерированный ИИ или фотореалистичный) не влияет на скорость и точность его нахождения. Перспективы данного исследования связаны с возможностями использования искусственно сгенерированных стимулов для планирования экспериментов в области когнитивной психологии.

© Сапронов, Ф.А., Горбунова, Е.С., 2025

Ключевые слова: зрительный поиск, искусственный интеллект, стимульный материал, зрительное внимание, зрительное восприятие

Финансирование. Работа выполнена при финансовой поддержке Российского научного фонда (РНФ), проект №20-78-10055-П, тема проекта «Категоризация в зрительном восприятии и формировании действий».

Благодарности. Авторы благодарят за помощь с программированием эксперимента стажера-исследователя лаборатории когнитивной психологии пользователя цифровых интерфейсов НИУ ВШЭ Клименкова Никиту и за помощь со стимульным материалом студента школы дизайна НИУ ВШЭ Васильева Кирилла.

Для цитирования: Сапронов, Ф.А., Горбунова, Е.С. (2025). Сравнение сгенерированных ИИ стимулов и фото: исследование зрительного поиска. *Вестник Московского университета. Серия 14. Психология, 48*(2). https://doi.org/10.11621/LPJ-25

Comparing AI-Generated Stimuli and Photos: Visual Search Study

Frol A. Sapronov, Elena S. Gorbunova ✉

National Research University Higher School of Economics, Moscow, Russian Federation

✉ esgorbunova@hse.ru

Abstract

Background. AI algorithms have rapidly entered everyday life in the last few years, which caused the appearance of a large number of both texts and images generated by neural networks. Nevertheless, the perception of artificially generated images has not yet been studied. The relevance of the study is to observe the possibility of using artificially generated images to plan experiments in psychology of perception.

Objective. The study investigated the differences in visual search for various stimuli types: photographs and AI-generated images.

Study Participants. Thirty participants (21 females) between 18 and 43 years old (M = 21.29, SD = 5.18) took part in the study.

Methods. The study had an experimental design with two factors and two dependent variables. The subjects' task was to find the target stimulus among distractors; the stimulus material included images of objects generated using a neural network and photographs. The number of objects on the screen also varied. The accuracy of the subjects' responses and reaction time were measured.

Results. No differences in detection accuracy were found between photorealistic and AI-generated stimuli, nor were there any differences in reaction times.
Conclusions. It was shown that the type of stimulus (AI-generated or photorealistic) does not affect the speed and accuracy of finding them. The perspectives of this study are related to the possibilities of using artificially generated stimuli for planning experiments in cognitive psychology.

Keywords: visual search, Artificial Intelligence, stimulus material, visual attention, visual perception

Funding. The study has been supported by the Russian Science Foundation (RSF), project No. 20-78-10055-P, title of the project "Categorization in visual perception and formation of actions".

Acknowledgments. The authors are grateful to Nikita Klimenkov, research assistant of the Laboratory for Cognitive Psychology of Digital Interface Users, National Research University Higher School of Economics, for the help with programming the experiment and Kirill Vasiliev, student of the HSE Art and Design School, for the help with the stimulus material.

For citation: Sapronov, F.A., Gorbunova, E.S. (2025). Comparing AI-generated stimuli and photos: visual search study. *Lomonosov Psychology Journal, 48*(2). https://doi.org/10.11621/LPJ-25

Введение

Современные и наиболее актуальные тренды в развитии технологий в большинстве своем связаны с системами искусственного интеллекта (ИИ). Например, такие нейронные сети, как ChatGPT и MidJourney, получили практически повсеместное распространение и огромный охват пользователей. Большое количество всевозможных программ доступны всем даже для бесплатного использования. Это, в свою очередь, приводит к возникновению споров, сомнений и опасений касательно взаимодействия человека с новыми технологиями (Brauner et al., 2023). Проникновение ИИ в повседневную жизнь вызывает множество вопросов, в частности — могут ли нейросети заменить некоторых специалистов? Например, иллюстраторов, журналистов, редакторов, переводчиков и других. В одном из выпусков The New York Times описывалось присуждение победы на конкурсе художников картине, созданной нейросетью (Roose, 2022). Очевидно, данный факт вызвал поток недовольства от профессионального со-

© Sapronov, F.A., Gorbunova, E.S., 2025

общества. Не менее громкий скандал возник в области образования: в одной социальной сети пользователь поделился, что использовал большую языковую модель для написания выпускной квалификационной работы (Жадан, 2023). Можно ли считать эту работу его интеллектуальной собственностью? Должны ли подобные тексты оцениваться наравне с работами с большим процентом плагиата?

Таким образом, вопрос об этичном использовании технологий ИИ встает не только в философских работах, но и в других областях. Психология не является исключением. Несмотря на то, что нейросети уже сейчас могут создавать изображения крайне высокого качества, большинство людей могут отличить сгенерированные нейросетями изображения от фотографий и рисунков: ИИ-изображения кажутся более тревожными, странными и пугающими (Rapp et al., 2024). Также более субъективно привлекательными и эстетичными кажутся те изображения, которые более похожи на сделанные человеком (Ragot et al., 2020). Таким образом, изучение особенностей восприятия ИИ-сгенерированных стимулов оказывается актуальной темой для изучения в рамках когнитивной психологии. Более того, уже сейчас можно сказать о том, что некоторые функции нейросетей могут облегчить и улучшить качество выполняемых задач, в частности — быть полезными при разработке дизайна эксперимента.

На данный момент нельзя дать четкого определения искусственного интеллекта, однако в литературе можно найти представление об ИИ как о системе с возможностью корректной интерпретации данных, обучения на основании этих данных и использования полученной информации для достижения определенных целей, адаптируясь под них (Kaplan, Haenlein, 2020). Несомненным преимуществом ИИ является его возможность манипулировать с разными типами данных: числовыми, текстовыми и даже фото- и видеоформатами. Под манипуляцией в данном случае подразумевается редактирование, дополнение и даже создание (генерация) нового контента. Те нейросети, которые в первую очередь приходят на ум при упоминании ИИ (например, ранее упомянутые ChatGPT, Midjorney или StableDiffusion), являются моделями, обучающимися на больших объемах данных. Так, например, ChatGPT является большой языковой моделью (large language models, LLM), а Midjorney и StableDiffusion — диффузными вероятностными моделями (diffusion probabilistic models). Большие языковые модели знакомы многим в связи с появлением доступных и бесплатных приложений и ботов, которые обладают удобным интерфейсом для введения запросов (промптов) и получения ответа. Эти

модели обучаются на огромном количестве текстовых данных, что позволяет им предсказывать последовательность слов. Таким образом, самые актуальные версии LLM могут генерировать большие тексты хорошего качества. Диффузные вероятностные модели научаются на большом количестве изображений, которые заранее размечаются специалистами (например, фотографии собак подписываются как «собаки», а кошек — как «кошки»). После обучения специальный алгоритм создает изображение по запросу (Ho et al., 2020).

Использование ИИ в исследованиях зрительного внимания

Как уже было упомянуто выше, системы ИИ широко используются в создании визуальных изображений, и эта функция представляет особый интерес для исследователей в области психологии, в частности — когнитивной. Одним из примеров задачи, в рамках которой можно наблюдать особенности восприятия ИИ-сгенерированных стимулов — задача зрительного поиска. Зрительный поиск представляет собой перцептивную задачу, в которой испытуемым необходимо найти целевой объект (один или несколько) среди отвлекающих — дистракторов (Wolfe, 2020).

Зрительный поиск является классической методикой в исследованиях внимания. Помимо этого, методика зрительного поиска широко применяется для изучения аффективных процессов и, например, взаимодействия высокоуровневых и низкоуровневых процессов (Ivanina et al., 2020). Одной из самых влиятельных теорий, объясняющих механизмы работы внимания при решении задачи зрительного поиска, является теория интеграции признаков (ТИП), созданная Энн Трейсман (Treisman, Gelade, 1980). Во множестве исследований Трейсман и коллеги манипулировали различными признаками объектов и просили респондентов искать целевые стимулы среди дистракторов. Предполагалось, что некоторые базовые признаки обрабатываются параллельно. Подобные признаки были названы базовыми, к ним отнесли форму, цвет, размер и пространственную ориентацию. Тем не менее в реальной жизни мы крайне редко сталкиваемся с объектами, обладающими одним или двумя признаками. Согласно ТИП, признаки связываются друг с другом посредством обращения внимания на место в пространстве, где располагается объект. Таким образом, внимание выступает в роли «клея» для соединения множества признаков в образе объекта. Было показано, что, если целевой стимул отличается от остальных по одному из базовых признаков, а все дистракторы будут гомогенны по этому признаку, цель будет най-

дена практически моментально. Этот феномен назвали «эффектом выскакивания», так как цель будто бы выскакивает для респондента при начале поиска. Так, примером эффекта выскакивания может быть поиск красного квадрата среди синих квадратов. Если же цель отличается не по одному базовому признаку, а дистракторы гетерогенны, то поиск будет осуществляться последовательно: испытуемый будет поочередно просматривать каждый стимул на экране, пока не идентифицирует один из них как целевой.

Тем не менее не все результаты эмпирических исследований согласуются с ТИП — например, иногда поиск по сочетанию признаков оказывается более эффективным, чем предсказывает эта теория (Quinlan, Humphreys, 1987). Ученик Э. Трейсман Дж. Вольф развил теорию интеграции признаков, дополнив ее представлениями о нисходящей регуляции зрительного поиска. Так появилась теория управляемого поиска (Guided Search Model) (Wolfe et al., 1989; Wolfe, 1994; 2021). Как и в ТИП, в теории управляемого поиска первые этапы обработки информации направлены на извлечение базовых признаков объектов, что тесно связано с ограничениями и особенностями зрительной системы. На этих первых этапах формируется «карта приоритетов» (priority map), где представлены репрезентации пространственного положения стимулов и вероятность каждого из объектов оказаться целевым. При дальнейшем поиске внимание направляется на объект, наиболее похожий на целевой стимул. В последней версии модели — Guided Search 6.0. — постулируется, что поиск реализуется не только восходящими, но и нисходящими процессами. То есть предшествующий опыт, ценность стимула и другие его характеристики также оказывают влияние на формирование «карты приоритетов». В частности, было показано, что эффект выскакивания может произойти, если мы столкнемся с незнакомым для нас объектом (Wolfe, 2001).

Несмотря на то, что в первых исследованиях зрительного поиска в качестве стимульного материала использовались геометрические фигуры (Treisman, 1964), в более поздних все чаще изучается поиск более сложных объектов (Chen, Zelinsky, 2006). Так, в задачах зрительного поиска варьируются такие свойства объектов, как одушевленность или неодушевленность, знакомость объекта, уровень категории объекта, эмоциональность изображения и другие (Vuilleumier, Huang, 2009). Существующих баз данных стимулов (BOSS, OASIS и др.) не всегда достаточно для подбора и уравнивания стимульного материала (Kudri et al., 2017; Сопов и др., 2019). Более того, при ис-

пользовании стоковых фотографий из Интернета можно столкнуться с вопросом о нарушении авторских прав при публикации. В связи с вышеописанным нейросети выступают удобным инструментом для генерации уникальных и подходящих под исследовательскую задачу стимулов. В свою очередь, это открывает ряд вопросов. С одной стороны, изображения, созданные такими нейросетями, как StableDiffusion и Midjorney, имеют высокое качество и практически не отличаются от фотографий. С другой стороны, основываясь на предыдущих исследованиях когнитивных процессов, очевидно, что восприятие может быть чувствительно к даже на первый взгляд незначительным параметрам стимульного материала (Julesz, 1975). Примерами таких параметров могут быть, например, текстура изображения или же его симметричность, а также эмоции, которые вызывает изображение. Несмотря на быстрое развитие систем искусственного интеллекта, влекущее за собой улучшение качества сгенерированного ими контента, полученные изображения могут восприниматься нереалистично в связи с самыми незначительными деталями: цвет, форма, изгибы и т.д. В реальной жизни мы можем встречаться с «необычными» изображениями — например, в случае сюрреалистичных и абстрактных изображений, созданных художниками, однако предметы повседневного использования, как правило, являются достаточно «стандартными».

В нашем исследовании мы задались вопросом о разнице в процессе поиска ИИ-сгенерированных стимулов и фотографий. Актуальным направлением для данной предметной области является использование экологически валидного стимульного материала — изображений объектов повседневного использования, и применение нейросетей могло бы быть здесь весьма полезным, позволив сократить затраты на создание стимулов. Однако возникает вопрос о том, действительно ли созданные нейросетью стимулы воспринимаются аналогично изображениям реально существующих объектов. В связи с этим была выдвинута гипотеза о том, что время и точность поиска ИИ-сгенерированных стимулов будет отличаться от времени и точности поиска фотографий объектов.

Выборка

В исследовании приняли участие 30 человек (21 женщина) от 18 до 43 лет ($M_{возраст}$= 21,29, SD = 5,18). Все респонденты имели нормальное либо скорректированное до нормального зрение, не имели

психиатрических и неврологических заболеваний, а также не переносили черепно-мозговые травмы в последние 6 месяцев. Русский язык являлся родным для всех испытуемых. Рекрутинг респондентов проводился через социальные сети.

Методы исследования

В качестве методики была выбрана одна из самых популярных для исследований зрительного восприятия и внимания — задача зрительного поиска. В качестве первой независимой переменной использовался тип стимула: сгенерированный ИИ или фотография из базы данных. Второй независимой переменной являлось количество стимулов на экране: 6, 8 или 12 изображений. Точность (количество верных ответов) и скорость поиска целевого стимула являлись зависимыми переменными.

Стимульный материал

Для генерации изображений использовалась нейросеть Stable Diffusion. Эта нейросеть работает по принципу стабильной диффузии, то есть превращая гауссовский шум в изображение, путем обучения на большом количестве данных (фотографий и других видов изображений). Пример запроса (промпта) для генерации изображения: «платье на вешалке» или «керамическая кружка на столе». Все сгенерированные изображения могут быть найдены на Open Science Framework (OSF)[1]. Фотореалистичные стимулы были взяты из нашего предыдущего исследования зрительного поиска, в котором одним из факторов являлся уровень категории цели, что требует уравнивания стимулов по этой характеристике, и базы данных BOSS (Сопов и др., 2019; Ангельгардт и др., 2021).

Процедура

Эксперимент был создан в программе PsychoPy v2023.1.2 и проводился онлайн с помощью сервиса Pavlovia. Респонденты могли проходить эксперимент только с персональных компьютеров (ПК) или ноутбуков. Прохождение эксперимента со смартфона или планшета не было предусмотрено. При регистрации для участия в эксперименте каждому респонденту присваивался собственный индивидуальный номер из нескольких цифр (например, 398016). После регистрации респонденту высылалась ссылка, ведущая на страницу эксперимента

[1] https://osf.io/duzs6/

в Pavlovia. Эксперимент длился примерно 30–40 минут. Все данные были собраны в 2024 г. Любая дополнительная информация, связанная с дизайном эксперимента, может быть запрошена у авторов по электронной почте.

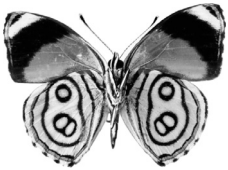
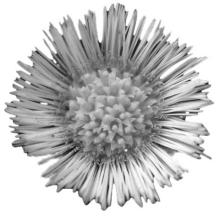
Задача испытуемых состояла в том, чтобы найти вербально заданный стимул среди отвлекающих (дистракторов). Эксперимент состоял из 2 блоков: в одном блоке все стимулы (как целевые, так и дистракторы) являлись изображениями, сгенерированными нейросетью, во втором блоке — фотографиями из баз данных. Порядок предъявления блоков был уравнен между испытуемыми.

Перед прохождением эксперимента испытуемым предъявлялась инструкция, в которой описывалась последовательность действий при выполнении задачи. Целевой стимул обозначался словом (например, «бабочка») в начале каждой пробы. Иными словами, поиск был категориальным: если респонденту предъявлялось слово «бабочка», то в дальнейшем необходимо было искать любую бабочку на экране. Всего было 4 категории целевых стимулов: «бабочка», «одежда», «кружка» и «цветок» (Таблица 1). После предъявления слова, обозначающего целевой стимул, на экране появлялись изображения (6, 8 или 12), среди которых нужно было искать соответствующий категории целевой стимул. Всего было по 3 вида целевого стимула на каждую категорию: например, для целевого стимула «бабочка» было создано 3 разных изображения бабочки, которые различались по ряду характеристик (цвет, форма и положение в пространстве). Целевой стимул мог присутствовать на экране либо отсутствовать. Если целевой стимул присутствовал, то он мог быть только один в каждой пробе. В каждом блоке было по 300 проб (всего 600 проб). На каждую категорию приходилось по 75 проб, 45 из которых содержали целевой стимул. Порядок предъявления слов, обозначающих целевой стимул, и порядок предъявления проб с разным количеством стимулов был рандомизирован. Ответ респондентов давался с помощью клавиш «→» и «←» на компьютере, где «→» — цель присутствует на экране, а «←»— цель отсутствует. Фиксировалась точность ответов испытуемого (количество верных ответов), а также время реакции. Визуализация процедуры исследования представлена на Рисунке 1.

Результаты исследования

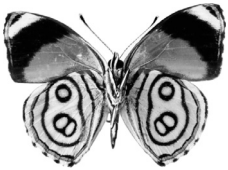
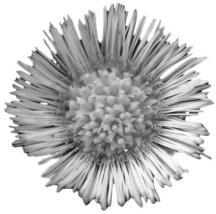
Анализу подлежала точность (процент верных ответов) и время реакции. Процент верных ответов был подсчитан по всем пробам

Таблица 1
Примеры целевых стимулов

Слово	Изображение ИИ	Изображение фото
Бабочка		
Кружка		
Одежда		
Цветок		

суммарно, без разделения на разные типы ответов (правильные попадания и правильные отрицания считались верными ответами, а ложные тревоги и пропуски — неверными). Для анализа использовался метод дисперсионного анализа (repeated measures ANOVA) с повторными измерениями, где в качестве внутригрупповых факторов выступали тип стимула и количество стимулов на экране. Так как ANOVA не чувствительна к нормальности распределения, тесты на нормальность не проводились (см., например, Наследов, 2004). Тест гомогенности дисперсий Левена показал равенство дисперсий в сравниваемых условиях. Попарно с помощью t-критерия Стьюдента сравнивались результаты для разного количества стимулов на экране,

Table 1
Examples of targets

Word	AI generated image	Photo image
Butterfly		
Mug		
Clothes		
Flower		

для контроля ложноположительных результатов для множественных сравнений использовалась поправка Бонферрони. Поправка Гринхауса — Гейсера применялась, если результаты теста Моучли на сферичность были статистически значимыми. При анализе точности (процента верных ответов) анализировались только пробы, в которых присутствовал целевой стимул. При анализе времени реакции анализировались только пробы, в которых испытуемый находил целевой стимул, то есть давал правильный ответ. Пробы, в которых время реакции отличалось от среднего сильнее, чем на ±2SD, были исключены из анализа. Предобработка данных производилась

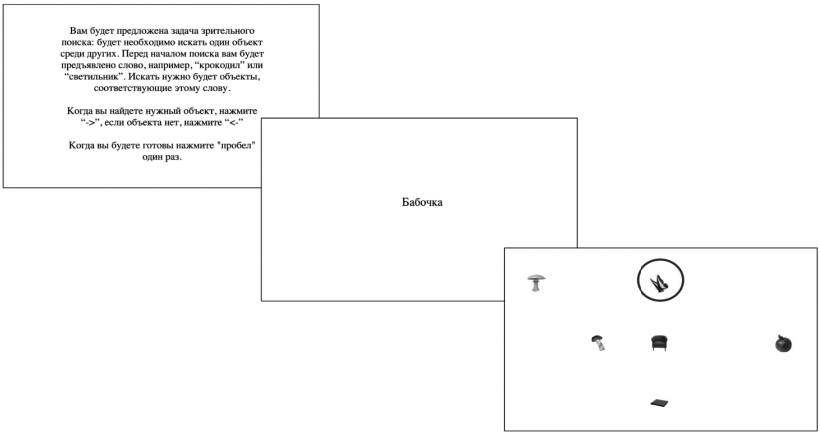


Рисунок 1

Визуализация процедуры исследования. Кругом на рисунке обозначен целевой стимул в пробе (в самой экспериментальной процедуре на экране обозначения целевого стимула кругом не было)

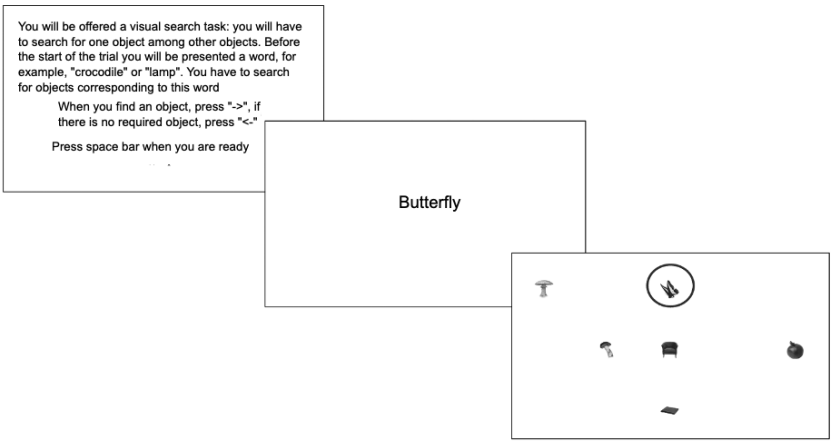


Figure 1

Visualization of the research procedure. The circle in the figure indicates the target stimulus in the trial (in the experimental procedure itself, there was no circle indicating the target stimulus on the screen)

на языке программирования Python с использованием библиотеки Pandas (McKinney, 2011). Для анализа данных использовалась программа JASP 0.16.4 (Love et al., 2019).

Результаты: точность (количество верных ответов)

Фактор типа стимула (сгенерированный ИИ или фотография из базы данных) оказался статистически незначимым, $F(1, 58) = 2{,}70$, $p = 0{,}10$, $\eta_p^2 = 0{,}04$. Фактор количества стимулов на экране оказался статистически незначимым, $F(2, 116) = 0{,}50$, $p = 0{,}60$, $\eta_p^2 < 0{,}01$. Взаимодействие факторов типа стимула и количества стимулов на экране также оказалось статистически незначимым, $F(2, 116) = 0{,}47$, $p = 0{,}62$, $\eta_p^2 < 0{,}01$. Графическое представление результатов представлено на Рисунке 2. Попарные сравнения с поправкой Холма — Бонферронни не выявили статистически значимых различий для разного количества стимулов на экране.

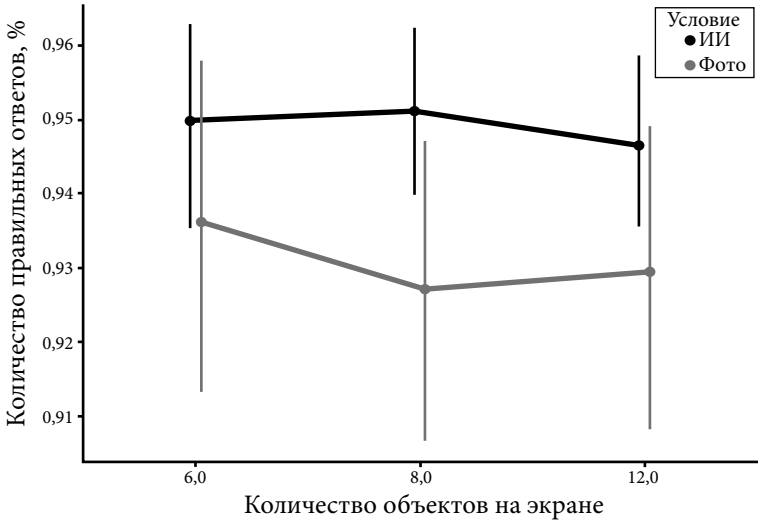


Рисунок 2

Графическое представление результатов по точности (проценту верных ответов) респондентов. «Усы» на графике отображают 95% доверительный интервал

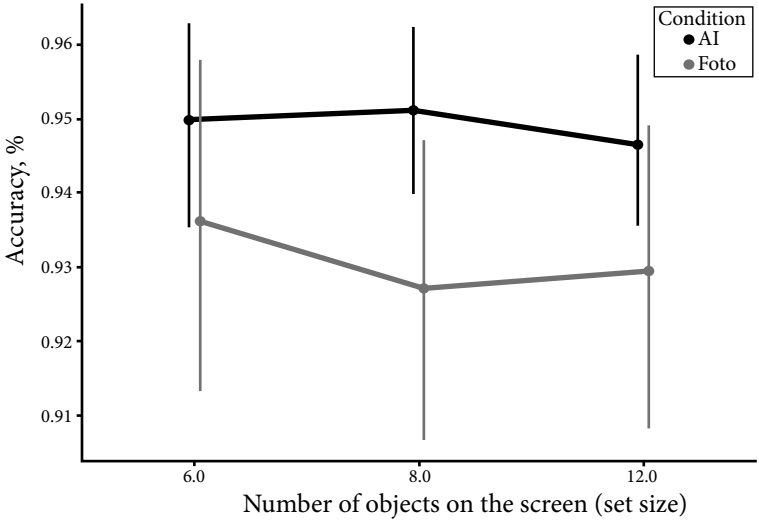


Figure 2

Graphical representation of the results by accuracy (percentage of correct responses) given by respondents. Error bars represent 95% confidence interval

Результаты: время реакции

Фактор количества стимулов на экране оказался статистически значимым $F(2, 116) = 12{,}19$, $p < 0{,}01$, $\eta_p^2 = 0{,}1$. Фактор типа стимула (сгенерированный ИИ или фотография из базы данных) оказался статистически незначимым, $F(1, 58) = 0{,}11$, $p = 0{,}74$, $\eta_p^2 < 0{,}01$. Взаимодействие факторов типа стимула и количества стимулов на экране оказалось статистически незначимым, $F(2, 116) = 1{,}82$, $p = 0{,}16$, $\eta_p^2 = 0{,}03$. Графическое представление результатов представлено на Рисунке 3. Попарные сравнения для фактора количества стимулов на экране представлены в Таблице 2.

Обсуждение результатов

Был получен типичный паттерн для задачи зрительного поиска: возрастание времени реакции при увеличении количества стимулов на экране (Wolfe, 2010). Согласно современным теориям зрительного поиска, например модели управляемого поиска 6.0, испытуе-

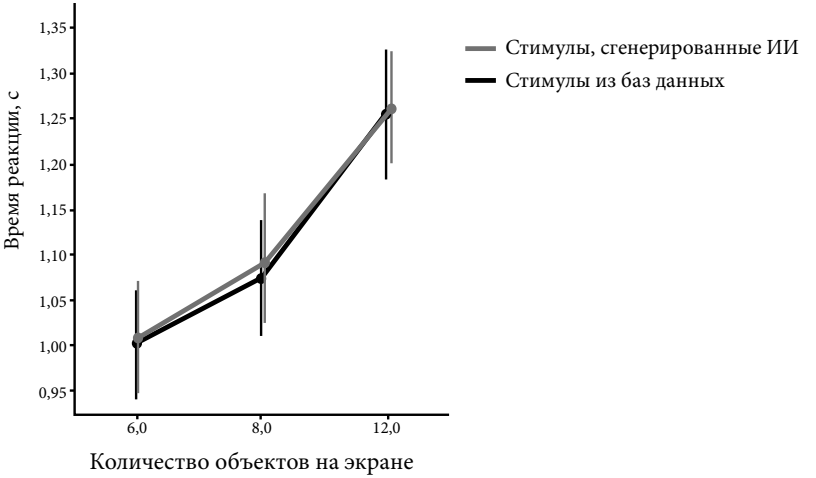


Рисунок 3

Графическое представление результатов по времени реакции. «Усы» на графике отображают 95% доверительный интервал

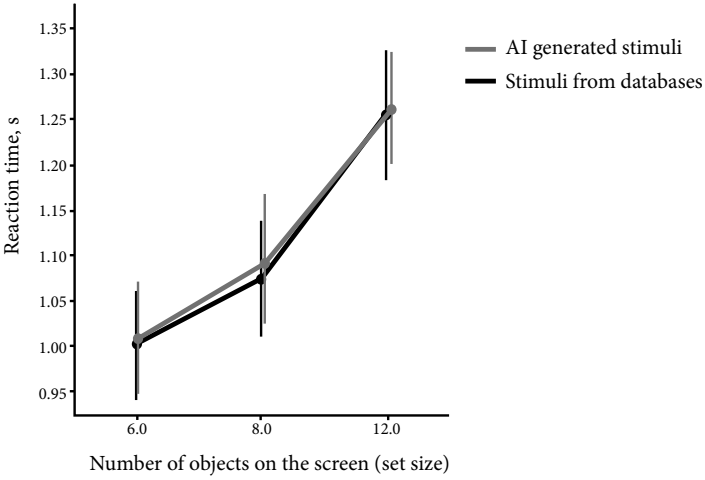


Figure 3

Graphical representation of the results by reaction time. Error bars represent 95% confidence interval

Таблица 2
Попарные сравнения для различных условий по количеству стимулов на экране (6, 8, 12 стимулов) с помощью t-критерия Стьюдента (введена поправка Бонфферрони на множественные сравнения)

Количество стимулов на экране		SE	t	pbonf
6	8	0,05	–1,48	0,16
	12	0,05	–4,80	<0,01
8	12	0,05	–3,39	<0,01

Table 2
Pairwise comparisons for different conditions by set size (6, 8, 12 stimuli) using Student's t-test (Bonfferroni correction for multiple comparisons used)

Set size		SE	t	pbonf
6	8	0.05	–1.48	0.16
	12	0.05	–4.80	<0.01
8	12	0.05	–3.39	<0.01

мые последовательно обрабатывают стимулы на экране в поисках целевого (Wolfe, 2021). Таким образом, при увеличении количества дистракторов увеличивается количество потенциальных целевых стимулов, на обработку которых тратится время. Точность ответов была достаточно высокой и не зависела ни от одного из факторов, что также является стандартным для исследований зрительного поиска результатом и связано с низкой сложностью задачи (Wolfe, 2010).

Главной задачей нашего исследования был поиск различий между поиском стимулов, сгенерированных нейросетью Stable Diffusion, и поиском фотореалистичных стимулов, взятых из баз данных и предыдущих исследований. Ни в точности (проценте правильных ответов), ни во времени реакции различий при поиске разных типов стимулов обнаружено не было, что говорит о сходных поведенческих паттернах при восприятии реально существующих изображений и стимулов, сгенерированных с помощью нейросетей. Мы полагаем, что, несмотря на некоторые особенности сгенерированных ИИ изображений (например, не вполне стандартная форма или изогнутость), они обладают достаточным количеством необходимых базовых признаков для корректной категоризации и идентификации.

При проведении экспериментов в области когнитивной психологии наблюдается тенденция к использованию реалистичных

стимулов. Более того, некоторые исследователи приближают задачу зрительного поиска к реальным жизненным задачам (Морозов, 2017; Ануфриева, Горбунова, 2022), примером которых является поиск объектов внутри сцены (Le-Hoa Võ, Wolfe, 2015). Это многократно увеличивает экологическую валидность исследований. Так, респондентам предлагается искать не геометрические фигуры, а фотореалистичные изображения предметов обихода или же изучать целые сцены (кухни, поля и другие). Однако при дизайне подобных экспериментов встает вопрос о стимульном материале: он должен быть сбалансированным, подходящим под задачи исследования, и его использование не должно нарушать авторские права его создателей. Например, при изучении эффекта категориальности в зрительном поиске исследователям зачастую приходится создавать изображения нескольких представителей одной категории (несколько видов цветов, несколько видов платьев и т.д.) (Сапронов и др., 2023). Не во всех стандартизированных базах данных может быть представлено нужное количество изображений, более того, не все из них доступны для свободного использования. Мы полагаем, что использование инструментов ИИ, таких как диффузные вероятностные модели, может выступать в качестве удобной альтернативы при создании стимулов для экспериментов. Нейросеть принимает запрос (промпт) и по нему формирует изображения. Так, при запросе «красное платье на вешалке» нейросеть предложит несколько изображений, подходящих под запрос (Ho et al., 2020). Это позволяет исследователю адаптировать запрос под свои цели и получить желаемое изображение. Полученное изображение при этом не нарушает авторских прав, а также достаточно легко поддается дальнейшей обработке: выравнивание по центру, изменение яркости, удаление фона и иные манипуляции. Согласно результатам нашего исследования, такие изображения не отличаются по паттернам их поиска в экспериментах на зрительный поиск. Мы полагаем, что они подходят и для других экспериментальных парадигм.

Тем не менее важно понимать, что нейросети не всегда генерируют идеальные изображения. Некоторые из полученных изображений могут получаться нереалистичными и даже пугающими, что также было продемонстрировано в предыдущих исследованиях (Rapp et al., 2024). Более того, иногда запросы (промпты) могут «восприниматься» нейросетью слишком буквально. Например, при запросе «торт Наполеон» нейросеть может сгенерировать изображение с тортом, на котором изображен французский император Наполеон I Бонапарт (Рисунок 4).

Рисунок 4

Примеры неудачных изображений, сгенерированных ИИ: слева — женщина со смазанным лицом и руками по запросу (промпту) «розовое платье», справа — торт с портретом Наполеона I Бонапарта по запросу (промпту) «торт Наполеон»

Figure 4

Examples of unsuccessful images generated by AI: on the left — a woman with a blurred face and hands for the query (prompt) "pink dress", on the right — a cake with a portrait of Napoleon I Bonaparte for the query (prompt) "Napoleon cake"

Все вышеописанное означает, что роль экспериментатора крайне велика при использовании технологий ИИ, так как после генерации изображения необходима его оценка на реалистичность и релевантность использования, а также такие изображения нуждаются в последующей обработке.

Выводы

Таким образом, использование систем искусственного интеллекта при создании экспериментов в когнитивной психологии представляет собой перспективное направление для развития. Алгоритмы машинного обучения и всевозможные инструменты ИИ могут значительно упростить и местами улучшить качество проведения исследований. В нашем исследовании было продемонстрировано, что стимулы, сгенерированные ИИ, и фотореалистичные стимулы из баз данных дают сходные поведенческие показатели при решении задачи зрительного поиска. Это может быть аргументом в пользу

дальнейшего использования ИИ при планировании исследований в этой области.

Список литературы

Ангельгардт, А.Н, Макаров, И.М., Горбунова, Е.С. (2021). Роль уровня категории при решении задачи гибридного поиска. *Вопросы психологии*, (2), 148–158.

Ануфриева, А.А., Горбунова, Е.С. (2022). Аффордансы как часть процесса идентификации объекта в зрительном поиске. *Российский психологический журнал, 19*(2), 188–200.

Жадан, А. (2023). Как я написал диплом с помощью ChatGPT и оказался в центре спора о нейросетях в образовании. Т — Ж, 3 марта 2023 г. URL: https://t-j.ru/neuro-diploma/ (дата обращения: 07.03.2025).

Морозов, М.И. (2017). Влияние категориальных названий на зрительный поиск. *Шаги*, *3*(1), 87–97.

Наследов, А.Д. (2004). Математические методы психологического исследования. Санкт-Петербург: Изд-во «Речь».

Сапронов, Ф.А., Макаров, И.М., Горбунова, Е.С. (2023). Категоризация в гибридном поиске: исследование с использованием регистрации движений глаз. *Экспериментальная психология*, *16*(3), 121–138.

Сопов, М.С., Стародубцев, А.С., Мирошник, К.Г., Шиндриков, Р.Ю. (2019). База стандартизированных изображений BOSS: адаптация для использования на русскоязычной выборке. *Психология. Журнал Высшей школы экономики*, *16*(4), 690–704.

Brauner, P., Hick, A., Philipsen, R., Ziefle, M. (2023). What does the public think about artificial intelligence? A criticality map to understand bias in the public perception of AI. *Frontiers in Computer Science*, (5), 1113903. https://doi.org/10.3389/fcomp.2023.1113903

Chen, X., Zelinsky, G.J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research*, *46*(24), 4118–4133.

Ho, J., Jain, A., Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, *33*, 6840–6851.

Ivanina, E.O., Gracheva, A.S., Markov, Y.A., Gorbunova, E.S. (2020). Search for familiar and dangerous: not seeing a gopnik in the crowd. *Психологические исследования, 13*(71), 1–25.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American*, *232*(4), 34–43.

Kaplan, A., Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Business Horizons*, *63*(1), 37–50.

Kurdi, B., Lozano, S., Banaji, M.R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods*, *49*, 457–470.

Le-Hoa Võ, M., Wolfe, J.M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, (1), 72–81.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q.F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J.N., Morey, R.D., Wagenmakers, E.J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*(2), 1–17.

McKinney, W. (2011). Pandas: a foundational Python library for data analysis and statistics. *Python for High Performance and Scientific Computing*, *14*(9), 1–9.

Quinlan, P.T., Humphreys, G.W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception & Psychophysics*, *41*(5), 455–472.

Ragot, M., Martin, N., Cojean, S. (2020). Ai-generated vs. human artworks. a perception bias towards artificial intelligence? In: R. Bernhaupt, F.F. Mueller, D. Verweij, J. Andres, (eds.). CHI EA '20: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (April 25-30, 2020). (pp. 1–10). New York: Association for Computing Machinery Publ.

Rapp, A., Di Lodovico, C., Torrielli, F., Di Caro, L. (2024). How do People Experience the Images Created by Generative Artificial Intelligence? An Exploration of People's Perceptions, Appraisals, and Emotions Related to a Gen-AI Text-to-image Model and its Creations. *International Journal of Human-Computer Studies*, *193*, 103375. https://doi.org/10.1016/j.ijhcs.2024.103375

Roose, K. (2022). An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy. The New York Times, September 2, 2022. URL: https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html?clckid=aa359dbd (accessed: 03.07.2025).

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, *3*(6), 449–459.

Treisman, A.M., Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Vuilleumier, P., Huang, Y.-M. (2009). Emotional attention: Uncovering the mechanisms of affective biases in perception. *Current Directions in Psychological Science*, *18*(3), 148–152.

Wolfe, J.M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.

Wolfe, J.M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics*, *63*, 381–389.

Wolfe, J.M. (2010). Visual search. *Current Biology*, *20*(8), 346–349.

Wolfe, J.M. (2020). Visual search: How do we find what we are looking for? *Annual Review of Vision Science*, *6*, 539–562.

Wolfe, J.M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092.

Wolfe, J.M., Cave, K.R., Franzel, S.L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

References

Angelgardt, A.N., Makarov, I.M., Gorbunova, E.S. (2021). The role of category level in solving the hybrid search task. *Voprosy Psikhologii*, (2), 148–158. (In Russ.)

Anufrieva, A.A., Gorbunova, E.S. (2022). Affordances as part of the object identification process in visual search. *Rossiiskii psikhologicheskii zhurnal = Russian Psychological Journal, 19*(2), 188–200. (In Russ.)

Brauner, P., Hick, A., Philipsen, R., Ziefle, M. (2023). What does the public think about artificial intelligence? A criticality map to understand bias in the public perception of AI. *Frontiers in Computer Science*, (5), 1113903. https://doi.org/10.3389/fcomp.2023.1113903

Chen, X., Zelinsky, G.J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research, 46*(24), 4118–4133.

Ho, J., Jain, A., Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, *33*, 6840–6851.

Ivanina, E.O., Gracheva, A.S., Markov, Y.A., Gorbunova, E.S. (2020). Search for familiar and dangerous: not seeing gopnik in the crowd. *Psikhologicheskie issledovaniya = Psychological Research, 13*(71), 1–25. https://doi.org/10.54359/ps.v13i71.189

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American, 232*(4), 34–43.

Kaplan, A., Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Business Horizons, 63*(1), 37–50.

Kurdi, B., Lozano, S., Banaji, M.R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods*, *49*, 457–470.

Le-Hoa Võ, M., Wolfe, J.M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences,* (1), 72–81.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q.F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J.N., Morey, R.D., Wagenmakers, E.J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*(2), 1–17.

McKinney, W. (2011). Pandas: a foundational Python library for data analysis and statistics. *Python for High Performance and Scientific Computing, 14*(9), 1–9.

Morozov, M.I. (2017). The influence of category names on visual search. *Shagi = Steps*, *3*(1), 87–97. (In Russ.)

Nasledov, A.D. (2004). Mathematical methods of psychological research. St. Petersburg: Rech' Publ. (In Russ.)

Quinlan, P.T., Humphreys, G.W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception & Psychophysics, 41*(5), 455–472.

Ragot, M., Martin, N., Cojean, S. (2020). Ai-generated vs. human artworks. a perception bias towards artificial intelligence? In: R. Bernhaupt, F.F. Mueller, D. Verweij, J. Andres, (eds.). CHI EA '20: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (April 25-30, 2020). (pp. 1–10). New York: Association for Computing Machinery Publ.

Rapp, A., Di Lodovico, C., Torrielli, F., Di Caro, L. (2024). How do People Experience the Images Created by Generative Artificial Intelligence? An Exploration of People's Perceptions, Appraisals, and Emotions Related to a Gen-AI Text-to-image Model and its Creations. *International Journal of Human-Computer Studies*, *193*, 103375. https://doi.org/10.1016/j.ijhcs.2024.103375

Roose, K. (2022). An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy. The New York Times, September 2, 2022. URL: https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html?clckid=aa359dbd (accessed: 03.07.2025).

Sapronov, F.A., Makarov, I.M., Gorbunova, E.S. (2023). Categorization in hybrid search: A study using eye movement registration. *Eksperimental'naya psikhologiya = Experimental Psychology, 16*(3), 121–138. (In Russ.)

Sopov, M.S., Starodubtsev, A.S., Miroshnik, K.G., Shindrikov, R.Yu. (2019). BOSS database of standardized stimuli: Adaptation for use on a Russian-speaking sample. *Psikhologiya. Zhurnal Vysshei shkoly ekonomiki = Psychology. Journal of the Higher School of Economics, 16*(4), 690–704. (In Russ.)

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior, 3*(6), 449–459.

Treisman, A.M., Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Vuilleumier, P., Huang, Y.-M. (2009). Emotional attention: Uncovering the mechanisms of affective biases in perception. *Current Directions in Psychological Science*, *18*(3), 148–152. https://doi.org/10.1111/j.1467-8721.2009.01626.x

Wolfe, J.M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.

Wolfe, J.M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics*, *63*, 381–389.

Wolfe, J.M. (2010). Visual search. *Current Biology, 20*(8), 346–349.

Wolfe, J.M. (2020). Visual search: How do we find what we are looking for? *Annual Review of Vision Science*, *6*, 539–562.

Wolfe, J.M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review, 28*(4), 1060–1092.

Wolfe, J.M., Cave, K.R., Franzel, S.L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 419–433.

Zhadan, A. (2023). How I wrote a thesis using ChatGPT and found myself at the center of a debate about neural networks in education. T — J, March 3, 2023. URL: https://t-j.ru/neuro-diploma/ (accessed: 03.07.2025).

ИНФОРМАЦИЯ ОБ АВТОРАХ

Фрол Алексеевич Сапронов, стажер-исследователь научно-учебной лаборатории когнитивной психологии пользователя цифровых интерфейсов Национального исследовательского университета «Высшая школа экономики»,

Москва, Российская Федерация, fa.sapronov@gmail.com, https://orcid.org/0000-0003-2384-8234

Елена Сергеевна Горбунова, кандидат психологических наук, заведующий научно-учебной лабораторией когнитивной психологии пользователя цифровых интерфейсов Национального исследовательского университета «Высшая школа экономики», Москва, Российская Федерация, gorbunovaes@gmail.com, https://orcid.org/0000-0003-3646-2605

ABOUT THE AUTHORS

Frol A. Sapronov, Research Assistant at the Laboratory for Cognitive Psychology of Digital Interfaces User, National Research University Higher School of Economics, Moscow, Russian Federation, fa.sapronov@gmail.com, https://orcid.org/0000-0003-2384-8234

Elena S. Gorbunova, Cand. Sci. (Psychol.), Head of the Laboratory for Cognitive Psychology of Digital Interface User, National Research University Higher School of Economics, Moscow, Russian Federation, gorbunovaes@gmail.com, https://orcid.org/0000-0003-3646-2605

Поступила 08.10.2024. Получена после доработки 27.02.2025. Принята в печать 17.03.2025.

Received 08.10.2024. Revised 27.02.2025. Accepted 17.03.2025.