

УДК 159.9.075

## ПАРАМЕТРИЧЕСКАЯ СЛОЖНОСТЬ ЗАДАНИЙ В ТЕСТАХ ИНТЕЛЛЕКТА<sup>1</sup>

© 2020 г. А. Н. Воронин\*, Н. Б. Горюнова\*\*

ФГБУН Институт психологии РАН;  
129366, г. Москва, ул. Ярославская, д. 13, Россия.

\*Доктор психологических наук, профессор, главный научный сотрудник  
лаборатории психологии речи и психолингвистики.

E-mail: voroninan@bk.ru

\*\*Кандидат психологических наук, старший научный сотрудник  
лаборатории психологии речи и психолингвистики.

E-mail: nat-goryunova@yandex.ru

Поступила 07.04.2020

*Аннотация.* Статья посвящена малоизученной проблеме параметров сложности отдельных заданий теста. Многозначность и недостаточная определенность понятия “сложность” задания приводит к смешению разнородных проблем в области разработки и оценки средств психодиагностики. В статье последовательно раскрывается содержание понятия “сложность” в рамках различных теорий тестирования: классической теории тестирования и теории латентных черт. Дается определение понятию “сложность” задания для разных типов тестирования: линейного, многоступенчатого и адаптивного. В статье рассмотрены различные коннотации понятия “сложность”. Показано разнообразие характеристик сложности в рамках науки о сложности (complexity science), в различных когнитивных и лингвистических исследованиях. В статье отмечено, что “сложность” и “трудность” задания являются близкими, но не совпадающими понятиями, при этом сложность задания теста предлагается рассматривать как конструктор, обладающий различными параметрами. В связи с этим сложность отдельного задания теста рассматривается как конструктор, составляющие части которого представляют отдельные параметры. Описаны основные параметры сложности заданий тестов интеллекта: количество однородных и/или разнородных элементов задания, количество нерелевантных элементов, количество операций (правил), производимых с элементами задачи, сочетание разных типов стимульного материала, изменение нескольких признаков элементов задания, использование различных типов ответов, временные ограничения на выполнение задания, динамика вариативности условий, нарушение принципов хорошей формы, “выпуклость” элементов, количество дистракторов, нетранзитивность объектов, игровые компоненты в структуре заданий, использование в задачах контекста разной степени конкретности. Варьирование описанными параметрами сложности позволяет конструировать задания теста с заданной вероятностью правильного ответа.

*Ключевые слова:* познавательные способности, интеллект, психодиагностика, сложность пунктов теста, трудность задания.

DOI: 10.31857/S020595920012585-9

Современные дискуссии в проблемном поле диагностики интеллекта разворачиваются в основном из-за изменений в понимании самих когнитивных способностей [3]: ставится под сомнение культурная независимость невербальных заданий в матрицах Равена [6]; продолжают развиваться представления об интеллекте, включающем общую

способность и дискретные функции специальных способностей с широким полем применения; подвергается критике одна из ведущих концепций, объединяющая кристаллизованный и флюидный интеллект, — модель Кетэлла—Хорна—Кэрролла [37]; развитие теории перекрывающихся процессов (process overlap theory) дало новую интерпретацию фактора “g” как проявление доменно-неспецифичных управляющих процессов в структуре рабочей памяти [25]. При этом, несмотря на продолжающуюся критику g как

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-013-00495 “Эмпирическая верификация структурно-функциональной модели когнитивного ресурса”.

неконстантного и неуловимого конструкта, он остается не только одним из наиболее верифицируемых, но и практически значимых в психологии [12; 13]. Несмотря на то что современные методы диагностики интеллекта базируются на теории отклика (item response theory — IRT), использование этой технологии наталкивается на ряд существенных ограничений: “статичность” оценок и игнорирование фактора изменения уровня проявления способности во времени (например, вследствие утомляемости); игнорирование времени выполнения заданий как одного из важных параметров проявления интеллектуальных способностей; необходимость выполнения значительного числа заданий для получения достоверных оценок [5]. При этом одна из основных проблем адаптивного подбора заданий, опирающегося на оценки IRT, обусловлена приблизительным равенством вероятностей правильного и неправильного выполнения заданий, что делает результаты тестирования зависимыми в основном от посторонних случайных факторов, не связанных с измеряемыми конструктами.

При разработке тестов диагностики интеллекта или отдельных когнитивных способностей его задания (пункты теста) должны соответствовать определенным требованиям: 1) обладать известной трудностью; 2) иметь определенную дифференцирующую способность; 3) положительно коррелировать с баллами по всему тесту и др. [1]. Требования к отдельным заданиям и тесту в целом формулируются в рамках двух теорий тестирования: классической теории тестирования и теории латентных черт [11]. В основе классической теории тестирования (The Classical Theory of Mental Tests), разработанной Ч. Спирменом и обобщенной Л. Гуттманом и Г. Гулликсеном, лежит теория погрешности измерений, предполагающая, что измеряемые качества имеют некоторые “истинные” значения, искажаемые случайными и систематическими погрешностями [20]. В соответствии с постулатами классической теории при разработке теста необходимо подбирать задания примерно равной трудности. Трудность задания оценивается через долю испытуемых, давших на него правильный ответ. Стремление к примерно равной трудности заданий в рамках классической теории тестирования приводит к тому, что проблема сложности заданий не представляется актуальной, а трудность и сложность заданий воспринимаются как тесно связанные, синонимичные понятия. Предполагается, что трудность задания определяется его сложностью [16], при этом сложность задания рассматривается как наиболее общая причина трудности задачи [40], как один из существенных аспектов

трудности [24]: в сложной задаче много элементов, каждый из которых в отдельности не представляет проблемы, но вместе они затрудняют решение.

В теории тестирования, основанной на теории отклика (IRT), проблема сложности отдельных заданий становится более явной. В рамках IRT, известной также как “теория латентных черт”, трудность задания является одним из параметров модели. Необходимость априорного установления трудности отдельных заданий и их возможности дифференцировать испытуемых по диагностируемому качеству предполагает апелляцию к сложности как некоей объективной характеристике. Лишь незначительное количество психологических тестов было разработано на основе IRT: “Адаптивная интеллектуальная диагностика” (AID); “Игральные кубики” (Free Dimension Cubes); “Венская матрица Форманна” (WMI) [8], “Тест интеллектуального потенциала” (разработан в HR-лаборатории “Гуманитарные технологии”). Часть из упомянутых тестов реализована в рамках компьютерного адаптивного тестирования (Computer Adaptive Testing — CAT), и именно там проблема сложности отдельных заданий представлена в полной мере.

Адаптивное тестирование позволяет алгоритмизировать последовательность предъявления заданий в зависимости от результатов на предыдущем шаге. Алгоритм отбора и предъявления заданий строится по принципу обратной связи: при верном ответе очередное задание выбирается более трудным, при неверном — менее. Условиями реализации такого алгоритма являются: 1) банк упорядоченных по трудности заданий, или “конструктор”, позволяющий генерировать задания нужной сложности, 2) компьютерные средства предъявления заданий и 3) алгоритмы оценки результатов выполнения [4; 5]. Менее строгим вариантом адаптивной диагностики является многоступенчатое тестирование с использованием тестов “уровня”. Алгоритм в этом случае также позволяет подстраиваться под испытуемого, но трудность очередного задания выбирается из блока типовых заданий определенного уровня сложности [9]. Необходимость разработки заданий различной сложности и их оперативная генерация по ходу тестирования делают проблему определения сложности разрабатываемых заданий крайне актуальной. В соответствии с этим целью данного исследования является экспликация и систематизация характеристик сложности, априорно заложенных в тестах познавательных способностей и исследовательских задачах при их разработке и конструировании.

## РАЗНООБРАЗИЕ ФОРМАЛЬНЫХ ХАРАКТЕРИСТИК СЛОЖНОСТИ ТЕСТОВЫХ ЗАДАНИЙ

Конструкт “сложность” имеет множество коннотаций, и большинство из них сосредоточено в рамках относительно новой междисциплинарной науки — complexity science — науке о сложности. Она сфокусирована на разработке моделей для объяснения свойств сложных систем, которые возникают из взаимодействий более простых составляющих. Сложные системы приобретают новые свойства, которые количественно и качественно отличаются от свойств составляющих и не могут быть выведены или предсказаны по ним. При этом указывается на нелинейность, эмерджентность, хаотичность и многие другие свойства систем как проявления их сложности [38]. С основными направлениями развития науки о сложности и персоналиями, внесшими существенный вклад в конкретное направление, можно ознакомиться в интернет-ресурсах по ссылке<sup>2</sup>.

Многообразие характеристик “сложности” ярко демонстрирует неоднородность данного конструкта и определяет необходимость описания ее параметров, используемых в психодиагностике познавательных способностей. Для определения сложности многие исследователи [21; 22] используют понятие теории информации, ключевым из которых является колмогоровская сложность, идея которой заключается в том, что чем сложнее объект, тем длиннее его описание. Другими словами, мерой сложности выступает длина кратчайшего возможного описания объекта. Следует отметить, что в лингвистических исследованиях обычно используется скорее неформальное определение колмогоровской сложности: чем короче описание, тем проще объект. Так, например, в работе [22] сложность рассматривается как степень сжимаемости текстов на разных языках. Использование имплицативных иерархий предполагает, что чем выше языковые структуры на иерархической шкале, тем они сложнее [30].

М. Гелл-Манн предложил использовать меру эффективной сложности, которую определил как минимальную длину описания закономерностей, демонстрируемых системой или объектом [15]. Согласно Мак-Уортер [28], область некоторой грамматики сложнее этой же области другой грамматики в той степени, в которой она содержит больше явно выраженных различий и/или правил. Р. Вуд выделяет три аспекта сложности задач:

- 1) компонентная сложность, представляемая как число отдельных актов для выполнения задачи;
- 2) динамическая сложность, в которой каждый шаг при выполнении задачи ведет к новому набору условий для принятия решений;
- 3) координативная сложность, при которой индивид должен в своих действиях отслеживать и объединять результаты предыдущих шагов, управляя прошлым и текущим состояниями решения задачи [40]. С. Гупта и Е. Караханна выделяют четыре аспекта сложности: 1) количество и разнообразие компонентов; 2) количество и сила взаимодействий; 3) совокупная скорость изменения; 4) трудность понимания причинно-следственных отношений в системе [19].

В сложных задачах обычно много элементов, наблюдается несоответствие правил (rule incongruity) анализируемым ситуациям, ситуации характеризуются неопределенностью (situational ambiguity), а время решения (time criticality) обычно ограничено и недостаточно [24]. Специальное исследование компонентов заданий в матрицах Равена показало, что в качестве дескрипторов сложности матричных заданий выступают симметрия/асимметрия, вертикальные, горизонтальные и диагональные оси, прямые и кривые линии, размер атрибутов ячейки, вариативность (например, разная ширина линий, разные формы), пропорциональное и непропорциональное изменение размера, затенение, количество цветов, пересечение и объединение размеров, вращение и отражение элементов, число, ориентация и тип фигур, сложные правила комбинирования элементов и т.д. [17]. Исследования влияния элементов матриц, используемых для измерения флюидного интеллекта (Gf) на вероятность правильного решения, показывают, что она снижается с увеличением числа фигур, увеличением количества правил, касающихся этих фигур, сложностью этих правил и перцептивной сложностью стимула. При этом самые сложные задания нагружены этими характеристиками одновременно, а наиболее важным фактором, определяющим сложность задания, является количество правил, применяемых в матрице [32].

## ВЛИЯНИЕ ПАРАМЕТРОВ СЛОЖНОСТИ НА ТРУДНОСТЬ ВЫПОЛНЕНИЯ ЗАДАНИЙ

Многокомпонентный состав “сложности” ставит проблему парциального влияния параметров сложности на трудность выполнения заданий. Однако предположение о том, что “трудность” выполнения заданий также является многокомпонентной (что не является предметом рассмотрения в нашей

<sup>2</sup> [http://scimaps.org/mapdetail/map\\_of\\_complexity\\_sc\\_154](http://scimaps.org/mapdetail/map_of_complexity_sc_154)

статье), существенно усложняет проблему сопоставления “сложности” и “трудности”. Для ее решения, вероятно, требуется обращение к теории множеств и использования диаграмм Венна для формулирования исходных гипотез, включающих различные отношения (объединение, пересечение, разность, симметрическая разность) нескольких подмножеств. В нашем случае мы рассматриваем вариант отсутствия пересечения между трудностью и различными параметрами сложности. Опишем для такого варианта наиболее часто встречающиеся параметры сложности тестовых заданий, влияющие на динамику когнитивного функционирования и приводящие к изменению трудности их выполнения.

Влияние количества однородных и/или разнородных элементов изучалось в работе [27], где анализировались факторы, опосредующие связь между типом множественных внешних репрезентаций (гомогенных или гетерогенных) и количеством правильно решенных заданий. Оказалось, что субъекты, обладавшие высокой вербальной памятью, лучше справлялись с гомогенными множественными репрезентациями. Напротив, такие переменные, как предметные знания и пространственные способности, не обладали функцией опосредования. Тем не менее предметные знания положительно коррелируют с количеством правильно решенных заданий и отрицательно — с воспринятой трудностью задания. В сбалансированном дизайне были представлены отдельные репрезентации (графика, текст или формула) или разнородные (например, текст + графика) или однородные (текст + формула) комбинации. Продуктивность оказалась выше в условиях множественных гетерогенных репрезентаций по сравнению с условиями отдельных репрезентаций. Кроме того, гетерогенные репрезентации привели к экономии времени по сравнению с гомогенными, графические репрезентации оказались наиболее выгодными по сравнению с текстовыми. По мнению М. Миестамо, сложность в основном сводится к количеству элементов, из которых состоит система [30]. В сложных заданиях анализ отдельных элементов не представляет особой проблемы, но их совокупность явно затрудняет решение [24]. Так, в матричных заданиях (прогрессивные матрицы Равена, “культурно-независимый” тест Кеттела) варьируется количество различных дескрипторов: вертикальные, горизонтальные и диагональные оси, прямые и кривые линии, размер атрибутов ячейки, пропорциональное и непропорциональное изменение размера, затенение, количество цветов, пересечение и объединение размеров, вращение и отражение элементов, число, ориентация и тип фигур и т.д. [17]. Одним

из наиболее важных факторов, определяющих сложность задания, является количество правил, применяемых в матрице [29; 32].

Включение в задание нерелевантных элементов изучалось в исследованиях влияния дизайна дистракторов на поведение испытуемых в процессе решения задачи. Согласно Р. Прими, нерелевантные элементы правил, которые в настоящее время не используются, могут нарушать перцептивную непрерывность обрабатываемого в данный момент правила [32]. Кригер с соавторами [26] предположили, что нерелевантная информация в фигурной матрице приводит к увеличению времени пребывания на определенной ячейке в заданиях с нерелевантной информацией по сравнению с заданиями без нерелевантной информации.

Увеличение количества операций с элементами задачи (умозаключения, формирование концепций, классификация, генерация и проверка гипотез, выявление отношений, понимание последствий, решение проблем, экстраполяция и преобразование информации и т.д.), необходимых для ее выполнения. Использование более широкого и разнообразного набора “когнитивных процессов” при решении разнообразных задач приводит к увеличению количества ошибок и времени отклика. Снижение продуктивности связывается с увеличением нагрузки на рабочую память, вызванной необходимостью отслеживать дополнительные элементы и преобразования [26].

Сочетание разных типов стимульного материала: символов, букв, цифр, линий, фигур и т.д. [10; 35]. В нейрокогнитивных исследованиях внимание акцентируется на механизмах кодирования разных символов. Некоторые вычислительные модели предполагают, что кодирование разных элементов реагирует на общие гибкие механизмы визуальной системы, другие предполагают, что кодирование запускает специфические процессы при обработке разных типов визуальных объектов. В работе [10] исследователи манипулировали положением и идентичностью символов посредством транспонирования или замены двух из четырех внутренних элементов в строках. Результаты показали, что некоторые области мозга реагировали больше на буквы, чем на цифры, что подтверждает предположение о специфичности нейронных путей.

Изменение нескольких параметров (цвет, форма, ориентация и т.д.) эффективно используется в парадигме обнаружения изменений для оценки объема кратковременной памяти [23; 34]. При генерации задания теста в него могут включаться различные дополнительные элементы, тогда сложность задания может рассчитываться как сумма

сложностей, вносимых в задание теста путем модификации дополнительных элементов задачи, их добавления или устранения [2]. Изменение расположения или иных характеристик каждого из них вносит свою долю в возрастание/снижение сложности задания.

Сложность задания зависит не только от содержания задачи, но и от вариантов ответа. Показано, что люди могут угадывать правильные ответы в психометрических тестах с несколькими вариантами ответов лучше, чем случайным образом, исходя из вариантов ответов, не видя самих заданий. Увеличение количества ответов и использование ответов “открытого” типа повышают сложность любого задания [39].

Еще одним параметром сложности является ограничение времени на выполнение задания. Анализ влияния фактора скорости на взаимосвязь между интеллектом и рабочей памятью показал, что фактор скорости необходим для достижения хорошего соответствия данных при тестировании с ограничением времени [33]. Конвергентная достоверность фактора скорости была подтверждена данными о скорости обработки. Кроме того, оказалось, что учет латентного фактора скорости привел к снижению корреляции между интеллектом и рабочей памятью.

Вариативность условий выполнения заданий. Люди часто сталкиваются со сложными и многоаспектными проблемами в повседневной жизни. Подмножество этих проблем может быть описано как взаимодействие взаимосвязанных переменных между человеком и динамической системой [14]. Управляя некоторыми из этих переменных, человек может попытаться перевести систему из ее нынешнего состояния в целевое состояние или оставить определенные критические переменные в допустимых пределах. При компьютерном моделировании такого рода проблем оценивают несколько возможных параметров ситуации:

- 1) сложность: многие аспекты ситуации должны приниматься во внимание одновременно;
- 2) взаимосвязанность: различные аспекты ситуации не зависят друг от друга и поэтому не могут контролироваться отдельно;
- 3) непрозрачность: только часть соответствующей информации предоставляется для решения проблем;
- 4) динамика: изменения в системе происходят без вмешательства агента;

5) политика: решатель проблемы должен иногда одновременно преследовать несколько целей, которые могут противоречить друг другу;

6) неопределенность: цели сформулированы лишь смутно и должны быть более точно определены решателем задач.

Используемый при моделировании подход составных минимально сложных систем (Minimally Complex Systems) позволил операционализировать сложность оценки ситуации и заменить экологическую достоверность на высоконадежные инструменты оценки, моделирующие отношения “искусственных микромиров” [18; 36].

Нарушение принципов “хорошей формы” [32]. Исследование сложности элементов геометрической индуктивной матрицы показывает, что сложность задания тесно связана с перцептивной сложностью элементов, из которых оно состоит.

Выпуклость элементов, используемых в фигурных матрицах, — еще один источник сложности [29]. Усложнение задания происходит, когда элементы матрицы трудно идентифицировать.

Использование в условии задания принципа нетранзитивности объектов [7]. При нетранзитивных отношениях превосходство системы  $A$  над  $B$ , а  $B$  над  $C$  не распространяется и на пару  $A-C$ , т.е. система  $A$  не превосходит систему  $C$  или даже уступает ей. Использование принципа нетранзитивности применительно к тестовым задачам повышает их сложность и позволяет расширить диапазон применения модели когнитивного ресурса.

Использование в задачах контекста разной степени конкретности (от предельно абстрактного до социально значимого). Показано, что контекстная версия при ее конкретизации и повышении социальной значимости уменьшает сложность задания — эффект контекстного облегчения [31].

Перечень выделенных параметров сложности тестовых заданий вряд ли является исчерпывающим, однако он дает достаточно полное представление о многообразии направлений варьирования различными характеристиками стимульного материала и условиями его предъявления для создания заданий различной трудности, сохраняя при этом предмет исследования и/или валидность разрабатываемой диагностической методики.

## ЗАКЛЮЧЕНИЕ

В статье предпринята попытка раскрыть содержание понятия “сложность задания” в тестах интеллекта. Показано, что такие характеристики

отдельных заданий теста, как “трудность” и “сложность”, воспринимаются схожими по своему содержанию, так как за ними в теории тестов изначально подразумевалась вероятность правильного решения задания. Однако по мере создания все более сложных тестов, при переходе от классической теории тестирования (КТТ) к теории латентных черт (IRT) и от линейного тестирования к адаптивному такое понимание осталось присущим только “трудности”. Понимание “сложности” изменилось и наполнилось новым содержанием: сложность отдельного задания понимается как объективная характеристика, зависящая от особенностей стимульного материала и обуславливающая вероятность правильного решения. Сложность заданий при таком понимании можно рассматривать как независимую переменную при проведении различных психологических исследований, направленных на изучение познавательных способностей. Варьирование параметров сложности позволяет экспериментатору оставаться в выбранном “поле” структуры интеллекта и элиминировать фактор взаимодействия отдельных познавательных способностей, сохранив конструктивную валидность. В психодиагностике (при адаптивном и/или многоступенчатом тестировании) варьирование сложностью позволяет создавать задания с требуемой вероятностью их правильного решения. В ходе исследования выделены следующие параметры сложности заданий тестов интеллекта: количество однородных и/или разнородных элементов задания, количество нерелевантных элементов, количество операций (правил), производимых с элементами задачи, сочетание разных типов стимульного материала, изменение нескольких признаков элементов задания, использование различных типов ответов, временные ограничения на выполнение задания, вариативность условий выполнения заданий, нарушение принципов хорошей формы, “выпуклость” элементов, нетранзитивность объектов, использование в задачах контекста разной степени конкретности.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Аванесов В.С.* Форма тестовых заданий: Учебное пособие. 2-е изд. М.: Центр тестирования, 2005. 156 с.
2. *Войтов В.К.* Расчет значений сложностей заданий для адаптивного теста интеллекта // Экспериментальная психология. 2013. Т. 6. № 2. С. 120–128.
3. *Корнилова Т.В., Корнилов С.А., Зиренко М.С., Чумакова М.А.* Психометрические свойства модифицированной батареи Интернационального Ресурса Когнитивных Способностей (ICAR) // Национальный психологический журнал. 2019. № 3 (35). С. 32–45.
4. *Куравский Л.С., Артеменков С.Л., Юрьев Г.А., Григоренко Е.Л.* Новый подход к компьютеризированному адаптивному тестированию // Экспериментальная психология. 2017. Т. 10. № 3. С. 33–45.
5. *Куравский Л.С., Юрьев Г.А., Ушаков Д.В., Юрьева Н.Е., Валуева Е.А., Лантева Е.М.* Диагностика по тестовым траекториям: метод паттернов // Экспериментальная психология. 2018. Т. 11. № 2. С. 77–94.
6. *Низбетт Р.* Что такое интеллект и как его развивать. Москва: Альпина нонфикшн, 2013.
7. *Поддьяков А.Н.* Опыт разработки объектов, находящихся в нетранзитивных отношениях превосходства // Когнитивная наука в Москве: новые исследования: Материалы конференции 15 июня 2017 г. / Под ред. Е.В. Печенковой, М.В. Фаликман. М.: ООО “Буки Веди”: ИППиП, 2017. 596 с.
8. *Шеронов Е.А.* Разработка психологического инструментария для определения пригодности к обучению на медицинских специальностях // Вестник КГУ им. Н.А. Некрасова. 2013. Т. 19. С. 51–54.
9. *Юсупов Ф.М.* О соотношении вербальной и невербальной составляющих в структуре интеллекта // Психологический журнал. 1995. № 1. С. 102–106.
10. *Carreiras M., Quiñones I., Hernández-Cabrera J.A., Duñabeitia J.A.* Orthographic Coding: Brain Activation for Letters, Symbols, and Digits. *Cerebral Cortex*, 2015. № 25. P. 4748–4760.
11. *Crocker L.* Introduction to Classical and Modern Test Theory / L. Crocker, J. Algina. N.Y.: Holt, Rinehart and Wilson, 1986.
12. *Cucina J., Byle K.* The bifactor model fits better than the higher-order model in more than 90% of comparisons for mental abilities test batteries // *Journ. of Intelligence*. 2017. № 5 (27). 1–21.
13. *Dubois J., Galdi P., Paul L.K., Adolphs R.* A distributed brain network predicts general intelligence from resting-state human neuroimaging data // *Philosophical Transaction of the Royal Society B: Biological Sciences*. 2018. № 373 (1756).
14. *Fahrenberg J.* Ökologische Validität [ecological validity] // *Dorsch-Lexikon der Psychologie* / ed. H. Wirz. Bern: Huber, 2017.
15. *Gell-Mann M.* The quark and the jaguar: adventures in the simple and the complex. N.Y., 1994.
16. *Gottfredson L.S.* Why g Matters: The Complexity of Everyday Life // *Intelligence*. 1997. № 24 (1). P. 79–132.
17. *Green K.E., Kluever R.C.* Components of Item Difficulty of Raven’s Matrices // *The Journal of General Psychology*. 1992. V. 119. № 2. P. 189–199.

18. *Greiff S., Wüstenberg S., Funke J.* Dynamic problem solving — A new assessment perspective // *Applied Psychological Measurement*. 2012. № 36. P. 189–213.
19. *Gupta S., Karahanna E.* Technology adoption in complex systems // 7th Annual Conference of the Southern Association for Information Systems. Savannah GA: University of Georgia, 2004.
20. *Guttman L.* A basis for analyzing test-retest reliability // *Psychometrika*. 1945. № 10. P. 255–282.
21. *Hammarström H.* Complexity in numeral systems with an investigation into pidgins and creoles // *Language complexity: typology, contact, change*. Amsterdam, 2008.
22. *Juola P.* Assessing linguistic complexity // *Language complexity: typology, contact, change*. Amsterdam, 2008.
23. *Kessler Y., Rac-Lubashevsky R., Lichtstein C., Markus H., Simchon A., Moscovitch M.* Updating visual working memory in the change detection paradigm // *Journal of Vision*. 2015. № 15 (9). P. 1–12.
24. *Kirschenbaum S.S.* What Makes Decision Tasks Difficult? // Newport. RI: Naval Undersea Warfare Center, 1999.
25. *Kovacs K., Conway A.R.A.* Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*. 2006. № 27 (3). P. 151–177.
26. *Krieger F., Zimmer H.D., Greiff S., Spinath F.M., Becker N.* Why are difficult figural matrices hard to solve? The role of selective encoding and working memory capacity // *Intelligence*. 2019. № 72. P. 35–48.
27. *Malone S., Altmeyer K., Vogel M., Brünken R.* Homogeneous and heterogeneous multiple representations in equation-solving problems: An eye-tracking study // *J Comput Assist Learn*. 2020. P. 1–18.
28. *McWhorter J.* Linguistic simplicity and complexity: why do languages undress? Berlin, 2011.
29. *Meo M., Roberts M.J., Marucci F.S.* Element salience as a predictor of item difficulty for Raven's Progressive Matrices // *Intelligence*. 2007. № 35. P. 359–368.
30. *Miestamo M.* Implicational hierarchies and grammatical complexity // *Language complexity as an evolving variable*. Oxford, 2009.
31. *Pérez-Salas C.P., Streiner D.L., Roberts M.J.* A comparison between element salience versus context as item difficulty factors in Raven's Matrices // *Intelligence*. 2012. № 40. P. 325–332.
32. *Primi R.* Complexity of geometric inductive reasoning tasks: Contribution to the understanding of fluid intelligence // *Intelligence*. 2002. № 30 (1). P. 41–70.
33. *Ren X., Wang T., Sun S., Deng M., Schweizer K.* Speeded testing in the assessment of intelligence gives rise to a speed factor // *Intelligence*. 2018. № 66. P. 64–71.
34. *Rouder J.N., Morey R.D., Morey C.C., Cowan N.* How to measure working memory capacity in the change detection paradigm // *Psychon Bull Rev*. 2011. № 18 (2). P. 324–330.
35. *Rüsseler J., Scholz J., Jordan K., Quaiser-Pohl C.M.* Mental rotation of letters, pictures, and three-dimensional objects in German dyslexic children // *Child Neuropsychology*. 2005. № 11. P. 497–512.
36. *Sonnleitner P., Brunner M., Greiff S., Funke J., Keller U., Martin R., Latour T. et al.* The Genetics Lab: Acceptance and psychometric characteristics of a computer-based microworld assessing complex problem solving // *Psychological Test and Assessment Modeling*. 2012. № 54. P. 54–72.
37. *The Evolution of Cognitive Ability Tests and Score Interpretation* // *Applied Measurement in Education*. V. 32. № 3. 2019.
38. *Waldrop M.M.* Complexity: The emerging science at the edge of order and chaos. Touchstone, N.Y., 1993.
39. *White A.P., Zammarelli J.E.* Convergence principles: Information in the answer sets of some multiple choice tests // *Applied Psychological Measurement*. 1981. № 5. P. 21–27.
40. *Wood R.E.* Task complexity: Definition of the construct // *Organizational Behavior and Human Decision Processes*. 1986. № 37. P. 60–82.

PARAMETRIC COMPLEXITY OF TASKS IN INTELLIGENCE TESTS<sup>3</sup>

A. N. Voronin\*, N. B. Goryunova\*\*

*Institute of Psychology, RAS;  
129366, Moscow, Yaroslavskaya str., 13, Russia.**\*Doctor of Psychology, Professor, Chief Researcher at the Institute of Psychology, RAS.**E-mail: voroninan@bk.ru**\*\*PhD in Psychology, Senior Researcher at the Institute of Psychology, RAS.**E-mail: nat-goryunova@yandex.ru*

Received 07.04.2020

*Abstract.* The article is devoted to the poorly studied problem of the component composition of the complexity of individual test items. The ambiguity and lack of certainty of the concept of “complexity” of the task leads to a confusion of problems in the development and evaluation of psychodiagnostic tools. The article consistently reveals the content of the concept of “complexity” in the framework of various testing theories: The Classical Theory of Mental Tests and The Item Response Theory. The meaning of the concept “complexity” of a task is determined for different types of testing: linear, multistage and adaptive. The article discusses various connotations of the concept of “complexity”. The diversity of complexity characteristics in the framework of complexity science, in various linguistic and cognitive studies is shown. “Complexity” and “difficulty” of the task are close, but not coinciding concepts, and the article attempts to dilute and reformulate them. In this connection, the “complexity” of an individual test item is considered as a construct, the constituent parts of which represent individual parameters. The main parameters of the complexity of the tasks of intelligence tests are described: the number of homogeneous and/or heterogeneous elements of the task, the number of irrelevant elements, the number of rules performed with the elements of the task, the combination of different types of stimulus material, the change in several signs of the elements of the task, the use of different types of answers, time restrictions on task performance, dynamics of variability of conditions, violation of the principles of good shape, “convexity” of elements, number of distractors, non-transitivity of objects, game components in the structure of tasks, use of contexts of varying degrees of concreteness in tasks. Varying the described complexity parameters makes it possible to construct test items with a given probability of a correct answer.

*Keywords:* cognitive ability, intelligence, psychodiagnostics, the complexity of test items, the difficulty of the task.

## REFERENCES

1. *Avanesov V.S.* Forma testovyh zadaniy. Uchebnoe posobie. 2 izd. Moscow: Centr testirovaniya, 2005. 156 p. (In Russian)
2. *Voitov V.K.* Raschet znachenii slozhnostej zadaniy dlya adaptivnogo testa intellekta. *Ekspierimental'naya psihologiya*. 2013. V. 6. № 2. P. 120–128. (In Russian)
3. *Kornilova T.V., Kornilov S.A., Zirenko M.S., Chumakova M.A.* Psihometricheskie svoïstva modifitsirovannoi batarei Internacional'nogo Resursa Kognitivnyh Sposobnostei (ICAR). *Nacional'nyi psihologicheskii zhurn.* 2019. № 3 (35). P. 32–45. (In Russian)
4. *Kuravskij L.S., Artemenkov S.L., Yur'ev G.A., Grigorenko E.L.* Novyj podhod k komp'yuterizirovannomu adaptivnomu testirovaniyu. *Ekspierimental'naya psihologiya*. 2017. V. 10. № 3. P. 33–45. (In Russian)
5. *Kuravskij L.S., Yur'ev G.A., Ushakov D.V., Yur'eva N.E., Valueva E.A., Lapteva E.M.* Diagnostika po testovym traektoriyam: metod patternov. *Ekspierimental'naya psihologiya*. 2018. V. 11. № 2. P. 77–94. (In Russian)
6. *Nizbett R.* Chto takoe intellekt i kak ego razvivat'. Moscow: Al'pina nonfikshn, 2013. (In Russian)
7. *Podd'yakov A.N.* Opyt razrabotki ob'ektov, nahodyashchihsvya v netranzitivnyh otnosheniyah prevoskhodstva. *Kognitivnaya nauka v Moskve: novye issledovaniya. Materialy konferencii 15 iyunya 2017 g.* Ed. E.V. Pechenkova, M.V. Falikman. Moscow: Buki Vedi, IPPiP, 2017. 596 p. (In Russian)
8. *Sheronov E.A.* Razrabotka psihologicheskogo instrumentariya dlya opredeleniya prigodnosti k obucheniyu na medicinskih special'nostyah. *Vestnik KGU im N.A. Nekrasova*. 2013. V. 19. P. 51–54. (In Russian)
9. *Yusupov F.M.* O sootnoshenii verbal'noj i neverbal'noj sostavlyayushchih v strukture intellekta. *Psihologicheskij zhurnal*. 1995. № 1. P. 102–106. (In Russian)

<sup>3</sup> The research is supported by a grant of the Russian Foundation for Basic Research (project № 20-013-00495 “Empirical verification of the structural-functional model of the cognitive resource”).

10. *Carreiras M., Quiñones I., Hernández-Cabrera J.A., Duñabeitia J.A.* Orthographic Coding: Brain Activation for Letters, Symbols, and Digits. *Cerebral Cortex*, 2015. № 25. P. 4748–4760.
11. *Crocker L.* Introduction to Classical and Modern Test Theory. L. Crocker, J. Algina. N.Y.: Holt, Rinehart and Wilson, 1986.
12. *Cucina J., Byle K.* The bifactor model fits better than the higher-order model in more than 90% of comparisons for mental abilities test batteries. *Journ. of Intelligence*. 2017. № 5 (27). 1–21.
13. *Dubois J., Galdi P., Paul L.K., Adolphs R.* A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transaction of the Royal Society B: Biological Sciences*. 2018. № 373 (1756).
14. *Fahrenberg J.* Ökologische Validität [ecological validity]. *Dorsch-Lexikon der Psychologie*. Ed. H. Wirz. Bern: Huber, 2017.
15. *Gell-Mann M.* The quark and the jaguar: adventures in the simple and the complex. N.Y., 1994.
16. *Gottfredson L.S.* Why g Matters: The Complexity of Everyday Life. *Intelligence*. 1997. № 24 (1). P. 79–132.
17. *Green K.E., Kluever R.C.* Components of Item Difficulty of Raven's Matrices. *The Journal of General Psychology*. 1992. № 119 (2). P. 189–199.
18. *Greiff S., Wüstenberg S., Funke J.* Dynamic problem solving — A new assessment perspective. *Applied Psychological Measurement*. 2012. № 36. P. 189–213.
19. *Gupta S., Karahanna E.* Technology adoption in complex systems. 7th Annual Conference of the Southern Association for Information Systems. Savannah GA: University of Georgia, 2004.
20. *Guttman L.* A basis for analyzing test-retest reliability. *Psychometrika*. 1945. № 10. P. 255–282.
21. *Hammarström H.* Complexity in numeral systems with an investigation into pidgins and creoles. *Language complexity: typology, contact, change*. Amsterdam, 2008.
22. *Juola P.* Assessing linguistic complexity. *Language complexity: typology, contact, change*. Amsterdam, 2008.
23. *Kessler Y., Rac-Lubashevsky R., Lichtstein C., Markus H., Simchon A., Moscovitch M.* Updating visual working memory in the change detection paradigm. *Journal of Vision*. 2015. № 15 (9). 18. P. 1–12.
24. *Kirschenbaum S.S.* What Makes Decision Tasks Difficult? Newport. RI: Naval Undersea Warfare Center, 1999.
25. *Kovacs K., Conway A.R.A.* Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*. 2016. № 27 (3). P. 151–177.
26. *Krieger F., Zimmer H.D., Greiff S., Spinath F.M., Becker N.* Why are difficult figural matrices hard to solve? The role of selective encoding and working memory capacity. *Intelligence*. 2019. № 72. P. 35–48.
27. *Malone S., Altmeyer K., Vogel M., Brünken R.* Homogeneous and heterogeneous multiple representations in equation-solving problems: An eye-tracking study. *J Comput Assist Learn*. 2020. P. 1–18.
28. *McWhorter J.* Linguistic simplicity and complexity: why do languages undress? Berlin, 2011.
29. *Meo M., Roberts M.J., Marucci F.S.* Element salience as a predictor of item difficulty for Raven's Progressive Matrices. *Intelligence*. 2007. № 35. P. 359–368.
30. *Miestamo M.* Implicational hierarchies and grammatical complexity. *Language complexity as an evolving variable*. Oxford, 2009.
31. *Pérez-Salas C.P., Streiner D.L., Roberts M.J.* A comparison between element salience versus context as item difficulty factors in Raven's Matrices. *Intelligence*. 2012. № 40. P. 325–332.
32. *Primi R.* Complexity of geometric inductive reasoning tasks: Contribution to the understanding of fluid intelligence. *Intelligence*. 2002. № 30 (1). P. 41–70.
33. *Ren X., Wang T., Sun S., Deng M., Schweizer K.* Speeded testing in the assessment of intelligence gives rise to a speed factor. *Intelligence*. 2018. № 66. P. 64–71.
34. *Rouder J.N., Morey R.D., Morey C.C., Cowan N.* How to measure working memory capacity in the change detection paradigm. *Psychon Bull Rev*. 2011. № 18 (2). P. 324–330.
35. *Rüsseler J., Scholz J., Jordan K., Quaiser-Pohl C.M.* Mental rotation of letters, pictures, and three-dimensional objects in German dyslexic children. *Child Neuropsychology*. 2005. № 11. P. 497–512.
36. *Sonnleitner P., Brunner M., Greiff S., Funke J., Keller U., Martin R., Latour T. et al.* The Genetics Lab: Acceptance and psychometric characteristics of a computer-based microworld assessing complex problem solving. *Psychological Test and Assessment Modeling*. 2012. № 54. P. 54–72.
37. *The Evolution of Cognitive Ability Tests and Score Interpretation // Applied Measurement in Education*. V. 32. № 3. 2019.
38. *Waldrop M.M.* Complexity: The emerging science at the edge of order and chaos. Touchstone, N.Y., 1993.
39. *White A.P., Zammarelli J.E.* Convergence principles: Information in the answer sets of some multiple choice tests. *Applied Psychological Measurement*. 1981. № 5. P. 21–27.
40. *Wood R.E.* Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*. 1986. № 37. P. 60–82.