



Разработка и валидизация инструмента оценки «Тест организационного планирования»

САДОВА Анастасия Романовна

ORCID: <https://orcid.org/0000-0001-8472-5305>

Автономная некоммерческая организация «Россия — страна возможностей», Москва, Россия

СОБОЛЕВА Ольга Борисовна

ORCID: <https://orcid.org/0000-0002-2348-2925>

Автономная некоммерческая организация «Россия — страна возможностей», Москва, Россия

ИВАНОВА Алина Евгеньевна

ORCID: <https://orcid.org/0000-0003-3340-7651>

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия

СТЕПАШКИНА Екатерина Алексеевна

ORCID: <https://orcid.org/0000-0002-4340-3753>

Автономная некоммерческая организация «Россия — страна возможностей», Москва, Россия

Аннотация. Цель. Исследование посвящено разработке и валидизации инструмента оценки «Тест организационного планирования», направленного на оценку способности планирования у взрослых профессионалов (специалистов и руководителей). Метод. В работе использованы современные подходы психометрического тестирования, включая современную теорию тестирования (*item response theory, IRT*) и классическую теорию тестирования. Процедура разработки включала несколько этапов: (1) разработка теоретической рамки, (2) создание первичного банка заданий, (3) когнитивные интервью, (4) экспертная оценка заданий, (5) апробация ($N = 1094$; 74% женщин, 26% мужчин; средний возраст = 49, $SD = 9.2$), (6) психометрический анализ и сбор свидетельств валидности. Результаты. Психометрический анализ на основе модели частичного оценивания (*partial credit model, PCM*) позволил собрать базовые свидетельства конструкторной валидности — корректное функционирование заданий и теста: одномерную структуру (собственное значение первого контраста = 1.33); согласие заданий с моделью ($\text{Infit MNSQ} = .92 - 1.17$); высокую связь заданий со способностью (.23 – .65); высокую надёжность ($\alpha = .78$, $\omega = .78$). Свидетельства конвергентной валидности получены при помощи корреляционного анализа взаимосвязей со следующими переменными: анализ информации ($r = .43$, $p < .001$), системное мышление ($r = .42$, $p < .001$), самоорганизация ($r = .15$, $p < .001$), тревожность ($r = -.14$, $p < .001$). При помощи ANOVA были получены свидетельства критериальной валидности: различия по полу — выше балл у мужчин ($\eta^2 = .01$, $p = .003$), возрасту — выше балл у более молодых ($\eta^2 = .03$, $p = .000$), должности — выше балл у руководителей ($\eta^2 = .01$, $p = .023$). Выводы. Разработанный тест является надёжным инструментом для оценки способности планирования с подтверждёнными свидетельствами валидности, пригодным для использования в кадровом отборе, оценке персонала, программах профессионального развития и исследованиях когнитивных компетенций. Будущие исследования могут быть направлены на адаптацию теста к различным профессиональным группам и анализ его взаимосвязи с другими когнитивными и личностными характеристиками. Ценность результатов.

Методика может быть использована специалистами по персоналу и организационными психологами для диагностики навыков планирования в профессиональной деятельности, а также в исследованиях когнитивных компетенций в контексте трудовой эффективности.

Ключевые слова: планирование, когнитивные способности, психометрический анализ, современная теория тестирования, организационная психология, психодиагностика.

Введение

Планирование играет одну из ключевых ролей в эффективной организации деятельности, как на индивидуальном, так и на организационном уровнях (Antony, 2000; Roshanai, Janeslätt, White, Lidstroem-Holmqvist, Holmefur, 2021). Способность прогнозировать развитие событий, формулировать цели и управлять ресурсами необходима для достижения высоких профессиональных результатов в условиях неопределённости и высокой сложности задач (Kerzner, 2019). В организационной среде способность планирования связана с принятием стратегических решений, координацией процессов и повышением продуктивности сотрудников (Verulava, 2023; Адаменко, Хорольская, Петров, 2020; Ясин, Бажданова, 2024). Несмотря на значимость данного конструкта, его измерение в практике организационной психологии остается методологически непростой задачей.

Многие инструменты оценки планирования, данные про которые опубликованы, разработаны с использованием шкалы Ликерта (Cao, Schüz, Xie, Lippke, 2013; Nassi, Hashimoto, Amaral, 2014; Papa-Gusho, Viçaku-Çekrezi, 2015; Вилкова, 2020). Несмотря на свою популярность, данный формат имеет некоторые ограничения (Joshi, Kale, Chandel, Pal, 2015). Например, субъективность интерпретации утверждений: респонденты могут представлять разные жизненные ситуации при выборе варианта ответа в ликертовской шкале (Wakita, Ueshima, Noguchi, 2012). В шкалах Ликерта отмечается также «эффект потолка», связанный с социальной желательностью, который проявляется в том, что респонденты могут склоняться к ответам, которые считаются социально приемлемыми или ожидаемыми (Vita, Coplin, Feiereisel, Garten, Mechaber, Estrada, 2013). Кроме того, самоотчётные методики подвержены влиянию субъективных факторов: например, в некоторых случаях мужчины, по сравнению с женщинами, склонны более высоко оценивать свои способности (Reilly, Neumann, Andrews, 2022; Syzmanowicz, Furnham, 2011). Данные особенности снижают дискриминирующую способность инструмента и ограничивают возможность выявления различий среди высококвалифицированных специалистов.

Кроме того, опросники, оценивающие планирование, могут быть неприменимы в ситуациях с высокими ставками, таких как отбор и развитие персонала в организациях. В этих условиях тесты способностей, требующие от респондента выполнения конкретных задач, позволяют получить более объективные данные о компетенциях, чем самооценочные шкалы. В отличие от самоотчётных методик, тесты способностей не зависят от уровня рефлексии и представлений респондента о своих способностях, что делает их более надёжным инструментом оценки (Bertua, Anderson, Salgado, 2005; Morgeson, Campion, Dipboye, Hollenbeck, Murphy, Schmitt, 2007; Viswesvaran, Ones, 1999). Это особенно актуально для управленческих позиций и позиций, связанных с проектной деятельностью, где ошибки в планировании могут приводить к значительным экономическим и организационным потерям.

Таким образом, актуальность разработки нового инструмента обусловлена необходимостью объективной оценки навыков планирования, которые играют ключевую роль в

успешной реализации проектов и управленческих задач. Тест, позволяющий диагностировать способность к организационному планированию, представленный в данной работе, предлагает надёжный и валидный способ измерения этого конструкта, что делает его ценным инструментом для исследователей и практиков в области организационной психологии.

Настоящая статья посвящена разработке и первичной валидации «Теста организационного планирования» (ТОП). В исследовании были использованы современные подходы к тестированию, включая анализ заданий в рамках современного подхода к тестированию — «современной теории тестирования» (*item response theory, IRT*)¹ — и сбор различных свидетельств валидности. Основной целью исследования является создание надёжного и качественного инструмента, который позволит объективно измерять способность к организационному планированию, преодолевая ограничения самоотчётных методик и обеспечивая высокую дискриминирующую способность. Разработанный тест ориентирован на измерение четырёх ключевых индикаторов планирования: планирование времени, оценка ресурсов, приоритизация задач и корректировка выполнения плана, что позволяет наиболее точно выявлять индивидуальные различия в способности к организации деятельности.

ТОП предназначен для оценки способности планирования у взрослых профессионалов (специалистов и руководителей) старше 18 лет. Планирование является критически важным для эффективной профессиональной деятельности, особенно в сферах, связанных с управлением, аналитикой и стратегическим планированием. Тест может использоваться в различных контекстах: от профессионального отбора и оценки персонала до программ развития управленческих компетенций. Он подходит как для ситуаций «низких ставок» (саморазвитие, корпоративное обучение), так и для ситуаций «высоких ставок» (отбор кандидатов на руководящие должности, кадровые решения).

Таким образом, ценность ТОП заключается в сочетании двух ключевых преимуществ: объективности теста способностей, свободного от влияния социальной желательности и самооценочных искажений, и «горячего» контекста заданий, моделирующих реальные рабочие ситуации. В отличие от существующих инструментов, ТОП позволяет непосредственно измерить когнитивный компонент планирования в условиях, максимально приближённых к организационной реальности, но без субъективности опросников и ограниченности «холодных» нейропсихологических проб.

Психологический конструкт «планирование» в современной научной литературе

Планирование как важный элемент профессиональной деятельности охватывает процесс, связанный с постановкой целей, задач и определением действий, необходимых для их достижения (Morris, Ward, 2004). Этот процесс представляет собой основу эффективного функционирования как отдельных людей, так и организаций в целом. Важно отметить, что организационное планирование включает в себя не только создание иерархий действий, но и оценку ресурсов, управление временем и анализ возможных альтернатив (Pellerin, Perrier, 2019).

¹ Считается, что название «современная теория тестирования» закрепилось в русскоязычной науке (прежде всего благодаря работам В. С. Аванесова) для того, чтобы подчеркнуть принципиальное отличие *«item response theory, IRT»* от «классической теории тестов» (*classical test theory, CTT*), когда в 1950–1960-х гг. оформился математический аппарат IRT. Прямой перевод на русский язык («теория ответа на задание») показался специалистам неуклюжим и неточным. Однако спустя 70 лет эта теория утратила свою символическую «современность», уже став фактически стандартом в области психометрики, в том числе компьютерного адаптивного тестирования. Сегодня всё чаще звучит новый вариант названия «Теория моделирования и параметризации педагогических тестов». — Прим. ред.

Планирование можно рассмотреть, как одну из исполнительных функций (*executive functions*) человека. Исполнительные функции — умственные операции, необходимые для контроля и выполнения целенаправленного поведения, а также сложных, новых и неавтоматических задач (Sorel, Pennequin, 2008). Базовыми исполнительными функциями являются рабочая память, когнитивная гибкость и сдерживающий контроль. На их основе строятся более высокого порядка, такие как рассуждение, решение проблем и планирование (Collins, Koechlin, 2012). Планирование — это исполнительная функция высокого уровня, поддерживаемая исполнительными функциями нижнего уровня и скоростью обработки (Sorel, Pennequin, 2008). Планирование часто необходимо для оптимизации выполнения сложных задач (Gross, Grossman, 2010).

Существующие определения понятия «планирование» похожи. Планирование обычно определяют как процесс формулирования последовательности операций, направленных на достижение некоторой цели (Hayes-Roth, Hayes-Roth, 1979). Планирование, как самостоятельный конструкт в когнитивной психологии, реализуется как общий когнитивный навык, используемый для организации поведения таким образом, чтобы указанная цель могла быть достигнута посредством серии промежуточных шагов, причём мысленное представление этой серии шагов составляет «план» (Owen, 1997). Планирование включает в себя несколько исполнительных подпроцессов. Задачи планирования требуют от участников определения подцелей, необходимых для достижения цели высокого уровня, выстраивания правильной и эффективной последовательности этих подцелей, обнаружения и исправления ошибок по мере их возникновения, одновременного удержания в уме нескольких (иногда противоречивых) фрагментов информации и поддержания устойчивых когнитивных усилий какое-то время (Allan, Sniehotta, Johnston, 2013).

К планированию также можно подойти с функциональной точки зрения, разбив на две взаимодополняющие функции: аккомодативную (деятельность «снизу вверх») для формулирования планов и ассимилятивную (деятельность «сверху вниз») для осуществления планов (Shallice, 1982). Первая функция позволяет человеку разрабатывать свои стратегии исходя из воздействий окружающей среды, а вторая предполагает воздействие на ситуацию путём применения этих планов. Х. Аартс с коллегами продемонстрировали, что эффективное планирование требует подготовки действий и осознания экологического контекста, в котором цель будет достигнута (Aarts, Dijksterhuis, Midden, 1999).

Инструменты, измеряющие исполнительные функции, можно разделить на «горячие» и «холодные». Инструменты описываются как «холодные», когда их задания вырваны из контекста и лишены аффективных или мотивационных компонентов. В отличие от них «горячие» инструменты считаются более подходящими для принятия решений в реальном мире, поскольку их задания подчёркивают социальную, мотивационную и эмоциональную значимость (Zelazo, Cunningham, 2007). Труднее отложить удовлетворение, проявить гибкость и следовать планам, когда речь идёт о социальных ожиданиях или чувствах (Prencipe et al., 2011). Инструменты для измерения производительности обычно оценивают «холодные» функции. Они имеют меньшую значимость в реальных ситуациях, которые вызывают возбуждение или эмоциональную валентность. Поэтому для взрослых актуально разрабатывать «горячие» способы измерить исполнительные функции, чтобы придать эмоциональную значимость задачам.

В организационной психологии конструкт «планирование» представлен в моделях управленческой компетентности, таких как «Руководство к своду знаний по управлению проектами» (РМВОК®), где планирование рассматривается как один из ключевых факторов

успешного проектного управления (Project Management Institute, 2021)². Одним из наиболее цитируемых изданий, описывающих в том числе модели для оценки персонала и его способностей к планированию на базе Руководства, является монография Г. Керцнера, где он говорит об иерархии планов в менеджменте (Kerzner, 2019). Модель Г. Керцнера определяется через следующие элементы:

- непрерывный процесс;
- переоценка;
- система реализации проекта и поддерживающая инфраструктура;
- соответствие своим стратегическим целям и задачам.

В работе ближневосточных авторов представлена интересная модель, в которой планирование описывается как часть процесса управления проектами, которое обеспечивает эффективное течение проекта на протяжении всего его существования (Barati, Asadi, Kalantari, Azadi, Mamoogian, 2014). Проект будет успешным, если любой элемент из процессов, перечисленных в модели, выполняет свои задачи или роли должным образом. Группы процессов, выделяемые авторами:

- иницилирующая группа процессов;
- планирующая группа процессов;
- исполняющая группа процессов;
- контролирующая и мониторинговая группа процессов;
- закрывающая группа процессов;
- успех проекта.

Таким образом, организационное планирование — важный элемент профессиональной деятельности, обеспечивающий постановку целей, разработку стратегий и эффективное распределение ресурсов. Оно связано с исполнительными функциями, такими как рабочая память, когнитивная гибкость и контроль, которые помогают управлять сложными задачами.

В управлении проектами планирование рассматривается как непрерывный процесс, включающий формирование последовательности действий и переоценку стратегий. Различные модели подчёркивают его структурированность и важность для успешной реализации проектов. В целом, планирование — это ключевой инструмент для достижения целей и эффективного управления ресурсами.

Обзор существующих инструментов оценки планирования

В этом разделе будет рассмотрено несколько примеров, демонстрирующих возможности оценки планирования при помощи различных способов: тестов способностей, тестов знаний, методик самоотчёта. Методики оценки планирования разрабатываются для самых разных областей: от нейропсихологии до самодиагностики. В ходе литературного обзора фактически не было обнаружено опубликованных данных о валидизированных инструментах на русском языке, направленных на измерение планирования³. В связи с этим для

² Руководство РМВОК® (*Project Management Body of Knowledge*) — это стандарт, разработанный Институтом проектного управления (*Project Management Institute, PMI*), который описывает лучшие практики, методы и процессы управления проектами, охватывая такие области, как инициация, планирование, исполнение, мониторинг, контроль и завершение проектов.

³ В период активного увлечения руководителей российских организаций технологией «Ассесмент-Центр» для формирования резерва на выдвижение получил широкое распространение так называемый «Оргтест», направленный на измерение, в числе прочих, компетенции «Планирование». Пример стимульного материала этого теста можно найти здесь: (Штроо, 2025, с. 102–105). В те времена Оргтест фактически представлял собой русифицированный и модифицированный на скорую руку вариант зарубежного теста «Разбор деловых бумаг» (*The business in-basket test*), но не прошедший стандартную процедуру валидизации (Frederiksen et al., 1957; Ward, 1959). Сегодня на рынке услуг для кадровых служб можно найти коммерческую электронную версию «Оргтеста», при этом разработчики «гарантируют» её валидность и надёжность (Смарт Бизнес Солюшнс). — *Прим. ред.*

анализа существующих подходов к оценке планирования были рассмотрены инструменты, описанные в основном в англоязычной литературе.

Примером нейропсихологического инструмента оценки планирования выступает тест, который позволяет оценить способность планирования в разрезе времени, описываемый авторами Р. Моррисом и Дж. Вардом (Morris, Ward, 2004). Процедура тестирования описывается следующим образом. Перед участниками расположены компьютерные экраны, содержащие некое место назначения (обозначенное как «дом») и от двух до шести компьютерных кнопок (обозначенных числами от 1 до 6), расположенных на случайных расстояниях от целевого места. Задачей участников было спланировать последовательность нажатий кнопок таким образом, чтобы кнопки прибыли к целевому месту одновременно (то есть, чтобы их прибытие было синхронизировано). Время нажатия первой кнопки составляло время планирования, а разница во времени между прибытием первой и последней кнопки составляла время синхронизации. Таким образом, задача в этом тесте касается организации последовательности компонентов действий испытуемым (Morris, Ward, 2004).

Другим примером является тест «Лондонская башня» (*Tower of London, ToL*), нейропсихологический тест, изначально разработанный на базе теста «Ханойская башня» в 1980-х годах Т. Шаллисом для оценки планирования и исполнительных функций (Karat, 1982; Shallice, 1982). В стандартной версии теста «Лондонская башня» респонденту предлагаются три стержня и набор дисков. Задача состоит в том, чтобы переместить диски с одной конфигурации на другую, следуя определённым правилам, таким как ограничение по количеству дисков, которое может быть помещено на стержень, или последовательности перемещения дисков. Чтобы успешно выполнить задание, участник должен спланировать серию движений (Morris, Ward, 2004).

Примером комплексного теста на планирование является игровой тест, которые используется в том числе для оценки снижения когнитивных функций в процессе старения. Этот тест содержит различные вариации инструмента «Виртуальная задача планирования» (*virtual planning task*). Тест представлен в виде игры, в которой участники должны были спланировать коктейльную вечеринку. Но среди перечня действий есть разные отвлекающие вещи. Пациенты с повреждениями лобной доли справлялись с задачей хуже, чем контрольная группа. В частности, они склонны были включать в свой план неактуальные действия (Phillips, MacLeod, Kliegel, 2005).

Можно привести следующий пример теста оценки знаниевого компонента планирования. В монографии Г. Керцнера описываются инструменты оценки для последовательного прохождения по пяти уровням зрелости организации в управлении проектами (Kerzner, 2019). Например, оценочные критерии для уровня 1 можно охарактеризовать как чисто «знаниевые». Они предполагают хорошее понимание «Руководства к своду знаний по управлению проектами». На каждый вопрос даётся пять возможных ответов, из которых верным является один.

Примером использования формата опросников для оценки планирования может служить методика «Пересмотренный личностный опросник NEO» (*Revised NEO Personality Inventory, NEO PI-R*), которая разработана на основе модели «Большая пятёрка» (*“Big five”*). В шкалу «Добросовестность» этого инструмента входят аспекты «Самодисциплина (*self-discipline*)» и «Организованность (*organizing or order(less)*)», соотносящиеся с умением планировать. Самодисциплина отражает склонность доводить задачи до завершения, организованность показывает склонность к планированию дел (Costa, 1996; Costa, McCrae, 2014). Русскоязычная адаптация осуществлена Т. А. Мартином с коллегами в 2002 г. (Martin et al., 2002).

В другой самоотчётной методике — опроснике «Стиль саморегуляции поведения, ССПМ» — шкала «Планирование» применяется для измерения склонности к выдвиганию и удержанию целей, сформированности осознанного планирования. В целом методика предназначена для диагностики развития осознанной саморегуляции и её индивидуальных профилей (Моросанова, Кондратюк, 2020). Теоретическим основанием опросника являются представления об осознанной саморегуляции человека как динамичной и многоуровневой метасистеме регуляторных процессов и свойств.

В опросниках «Стратегии обучения и подготовки к занятиям» (*The Learning and Study Strategies Inventory, LASSI*) и «Саморегулируемое онлайн-обучение» (*The Online Self-Regulated Learning Questionnaire, OSLQ*) шкала «Планирование рабочего времени» (*Planning working hours*) выступает как одна из субшкал в рамках общего процесса саморегулирования (Вилкова, 2020).

Ещё одним перспективным направлением в оценке компетенции планирования является применение методологии современных ситуационно-поведенческих тестов и мультимедийных оценочных шкал. Однако анализ существующих разработок в русскоязычной и англоязычной научной литературе показал, что полностью релевантных нашей задаче инструментов в открытых публикациях не представлено. Тем не менее можно обратить внимание на инструмент «Ситуационно-поведенческий тест для выпускников вузов» (*Graduate Situational Judgement Test, GSJT*) от компании «Услуги профессиональной психологии» (*The OPC*), который включает, как сообщается на официальном сайте разработчика, компетенцию «Планирование» (*Planning*) (*The OPC*, 2025). Однако он разработан как комплексный ситуационно-поведенческий тест для выпускников вузов, не выделяет отдельной шкалы «Планирование» и не раскрывает её психометрических характеристик.

Можно сделать следующий вывод. Нейропсихологические методики ориентированы на планирование в клинических или экспериментальных задачах, но ограничены узким контекстом и не всегда применимы для оценки способности планирования у взрослых в профессиональной или повседневной деятельности. Самооценочные шкалы удобны для быстрой оценки состояния, но их субъективность снижает точность. Тесты знаний оценивают освоение материала, но ограничены только конкретными знаниями и не учитывают когнитивные особенности, важные для общего планирования.

Таким образом, представляемый в данной статье «Тест организационного планирования» расширяет контекст измерения. Он учитывает прикладное значение планирования в профессиональной среде.

Методика

Разработка оригинального инструмента оценки способности планирования

Этапы разработки теста

Разработка ТОП проводилась в соответствии с современными стандартами психометрического тестирования (AERA, APA, NCME, 2014; Батурин и др., 2015) и включала следующие базовые этапы.

1. Разработка плана теста, проведение литературного обзора и формирование теоретической рамки, основанной на анализе существующих подходов к измерению планирования.

2. Создание первичного банка заданий. Графический стимульный материал для заданий был создан в среде R, ver. 4.4.2. Основные пакеты: *dplyr*, *flextable*, *rsvg*, *grid*, *gridExtra*, *grDevices*, *magick*, *ggplot2*.

3. Внутреннее рецензирование и доработка заданий двумя экспертами по разработке тестов.

4. Когнитивные интервью⁴ с семью респондентами для уточнения формулировок и исключения неверно функционирующих заданий.

5. Экспертная оценка теоретической рамки и заданий тремя внешними экспертами, специалистами по организационному планированию, для подтверждения содержательной валидности теста.

6. Апробация теста. Тестирование прошли 1094 респондента. Описание выборки приведено в разделе 4.2.

7. Психометрический анализ, включавший проверку надёжности, анализ функционирования заданий, анализ размерности теста, сбор свидетельств валидности. На основе проведённых работ банк заданий был доработан, сформирована итоговая версия теста.

8. Нормирование результатов и разработка формы обратной связи. В Приложении 1 представлены примеры заданий теста.

Теоретическая рамка

В рамках разработки теста *планирование* понимается как *целенаправленный когнитивный процесс разработки и координации действий, направленный на поиск и выбор оптимального способа распределения доступных ресурсов (временных, материальных, человеческих) для достижения поставленных целей*. Этот процесс включает оценку ресурсов, необходимых для выполнения плана, приоритизацию задач, распределение времени и корректировку стратегии в соответствии с изменяющимися обстоятельствами. Планирование обеспечивает эффективность и адаптивность деятельности, минимизируя риски и способствуя рациональному использованию ресурсов.

- С опорой на литературный обзор, а также на Руководство к своду знаний по управлению проектами выделены следующие индикаторы планирования.
- Планирование времени. Корректный расчёт временного ресурса для достижения поставленных целей и задач (три задания).
- Оценка ресурсов для выполнения плана. Определение доступных ресурсов (люди, материалы, за исключением времени) и их ограничений (четыре задания).
- Ранжирование задач. Распределение задач в порядке их важности и срочности для достижения планов (четыре задания).

Корректировка выполнения плана. Принятие решений на основе анализа данных о прогрессе проекта (четыре задания).

Предполагается, что тест измеряет общий конструкт, а приведённые индикаторы являются содержательными компонентами, важными для операционализации и очевидной валидности, а не отдельными размерностями-субфакторами.

В связи с особенностями целевой аудитории был выбран «горячий» контекст заданий, связанный с рабочими задачами, направленными на поиск наилучшего способа распределения ресурсов для достижения целей.

Выборка

В исследовании приняли участие 1094 респондента; средний возраст участника — 49 лет (со стандартным отклонением, равным 9,2). Апробация проведена на профильной группе специалистов, постоянно использующих навыки планирования в работе (руководители финансово-экономической службы, бухгалтеры). Данные были собраны в ходе проведения оценки надпрофессиональных компетенций руководителей и сотрудников финансовых служб ряда

⁴ Когнитивные интервью — глубокие преапробационные интервью с респондентами, выполняющими тестовые задания вслух с проговариванием того, что и как они делают, насколько хорошо понимают задание, инструкции, стимульные материалы и прочее.

образовательных организаций. Учитывая профессиональную компетентность и опыт участников тестирования, данные, собранные на этой группе, являются полезными для анализа.

Средние значения и стандартные отклонения баллов групп представлены в *T*-баллах (шкала со средним значением, равным 500 баллам, стандартным отклонением, равным 100 баллам, и минимумом и максимумом, равным 200 и 800 баллам соответственно). В таблице 1 представлены социально-демографические параметры выборки.

Таблица 1. Социально-демографические характеристики респондентов

Переменная	Категория	Частота	%	Ср. знач.	Станд. откл.
Должность	Руководитель финансово-экономической службы	586	54.06	513	101
	Бухгалтер	498	45.94	484	96
Пол	Женский	323	76.36	492	98
	Мужской	100	23.64	526	103
Федеральный округ	Центральный	405	37.74	495	101
	Сибирский	182	16.96	505	97
	Северо-Западный	112	10.44	516	97
	Приволжский	93	8.67	503	99
	Южный	72	6.71	500	105
	Уральский	71	6.62	500	106
	Дальневосточный	70	6.52	502	101
	Северо-Кавказский	37	3.45	479	98

Используемые инструменты

Для сбора свидетельств конвергентной валидности разрабатываемого теста через связь с другими переменными был использован ряд методик (Таблица 2). Все инструменты, использованные в исследовании, прошли психометрическую оценку, подтверждающую их валидность и надёжность, согласно внутренним исследованиям организации (Россия — страна возможностей. Департамент оценки и методологии, 2024).

Таблица 2. Характеристики инструментов⁵, использованных для проверки конвергентной валидности

Название методики ⁶	Кол-во заданий	Формат заданий	Кол-во шкал	Используемые шкалы	Надёжность
«Личностный опросник»	14	Ипсативный; блоки по четыре утверждения	7	Организованность; Тревожность	$\alpha = .61 - .73$
«Тест на оценку мотиваторов и демотиваторов»	96	Шкала Ликерта	16	Управление; Вызов	$\alpha = .60 - .78$
«Комплексный тест на анализ информации (сокращённая версия)»	21	Множественный выбор с одним верным вариантом ответа	1	Анализ информации	.92 (EAP)
«Тест на коммуникативную грамотность»	17	Множественный выбор с одним или несколькими верными ответами	1	Коммуникативная грамотность	.87 (EAP)
«Тест на оценку абстрактного мышления и системный анализ»	12	Множественный выбор с одним верным вариантом ответа	1	Системный анализ	.84 (EAP)

Методы анализа

Общим методологическим руководством разработки и валидации ТОП для нас выступала модель Европейской федерации ассоциаций психологов (*test-review model, EFPA*)

⁵ Выходные данные публикаций представлены в каталоге инструментов оценки и диагностики [Россия — страна возможностей. Департамент оценки и методологии. Каталог инструментов оценки и диагностики (2024)].

⁶ Названия инструментов в таблице, взятыми на учёт в Департаменте оценки и методологии АНО «Россия — страна возможностей», приводятся в соответствии с их техническими паспортами.

(Evers, Muñiz, Hagemester, Høstmælingen, Lindley, Sjöberg, Bartram, 2013). Для установления того, насколько корректно функционируют задания в рамках теории ответа на задание применялась модель Раша, в частности модель частичного оценивания (*partial credit model, PCM*) и дополнительно — анализ на основе классической теории тестирования (Andrich, Marais, 2019; Masters, 2016).

Современная теория тестирования используется для анализа взаимосвязи между латентными характеристиками респондента (например, способностями) и его ответами на задания. Современная теория тестирования основана на идее, что вероятность верного ответа является математической функцией параметров респондента (способности) и задания (например, трудность, дискриминативность, угадывание, параметры порогов и так далее). Анализ в современной теории тестирования позволяет оценить функционирование каждого задания и всего теста в целом. Такой анализ позволяет собрать базовые свидетельства конструктивной валидности. Модель частичного оценивания предназначена для оценки способностей респондентов на основе их ответов на тестовые задания с градуированной шкалой. Модель частичного оценивания позволяет учесть неполные или частично правильные ответы (например, два балла за все верно выбранные ответы среди нескольких предложенных, один балл — за часть верно выбранных ответов).

Предшественница современной теории тестирования, классическая теория тестирования (КТТ) основывается на предположении, что результаты тестирования зависят от истинного уровня способностей респондента и случайной ошибки. КТТ не учитывает сложные взаимодействия между заданиями и способностями респондентов, что может ограничивать точность оценок. Однако анализ в рамках этой концепции может дополнить анализ в современной теории тестирования — в частности, дать базовую полезную информацию об ответных опциях и их функционировании.

В качестве свидетельства конвергентной валидности использовался коэффициент корреляции Пирсона результатов выполнения ТОП и тестов, измеряющих конструкты, которые предположительно должны быть значимо и положительно связаны со способностью планирования. Цель корреляционного анализа: определить, существует ли линейная связь между двумя переменными, и понять, насколько она сильна. В качестве свидетельства критериальной валидности использовался однофакторный дисперсионный анализ (*ANalysis Of VAriance, ANOVA*) для анализа различий по переменным, выступающим внешним критерием для планирования. Цель дисперсионного анализа: определить, есть ли статистически значимые различия между средними значениями групп респондентов по нескольким независимым переменным. Однофакторная ANOVA сравнивает средние значения по одному фактору (зависимой переменной).

Оба вида анализа проводились в программной среде R, ver. 4.4.2. Основные пакеты: TAM, psych, ctt, mirt, stats.

Процедура

Апробация инструмента ТОП проводилась в ноябре и декабре 2024 года. Участники проходили тест в рамках общей оценочной кампании вместе с рядом других инструментов. Для прохождения тестовых методик респонденту необходимо было зарегистрироваться на платформе АНО «Россия — страна возможностей». Тестирование было доступно для прохождения с персонального компьютера или ноутбука. Тест был доступен для однократного прохождения.

До прохождения тестирования участникам автоматизированно предъявлялась общая инструкция к тесту. В процессе заполнения методики было возможно использование респондентами вспомогательных средств (например, ручки и бумаги для записей и калькулятора). После

прохождения тестирования участники получали отчёт с обратной связью об уровне выраженности измеряемых компетенций (за исключением “Планирования”, поскольку тест находился на апробации). В рамках апробации время прохождения ТОП не было ограничено по времени.

Результаты

Психометрический анализ

Анализ на основе теории ответа на задание

Для установления особенностей функционирования заданий и подтверждения единой размерности инструмента был проведён психометрический анализ с использованием модели частичного зачёта. Параметры заданий получены при помощи метода максимального правдоподобия (*marginal maximum likelihood, MML*), параметры респондентов — при помощи метода оценки апостериорного математического ожидания (*expected a posteriori, EAP*) (Bock, Aitkin, 1981; Bock, Mislevy, 1982).

В таблице 3 представлены основные психометрические параметры теста. В Приложении 2 приведены параметры отдельных заданий, полученные в рамках анализа при помощи модели частичного оценивания.

Таблица 3. Общие параметры инструмента (современная теория тестирования)

Параметр	Среднее	Минимум	Максимум	Рекоменд. значения
Взвеш. стат. согласия (Infit MNSQ)	1.00	.92	1.17	[.5; 1.5] ⁷
Корр. балла за задание и способности	.47	.23	.65	≥ .23 ^{8 9 10}
Трудность	-.10	-1.33	1.00	-

Параметр	Значение	Рекомендуемое значение
Надёжность (EAP)	.78	≥ .66 ¹¹
Собственное значение одного контраста остатков	1.33	≤ 2 (для одномерности) ¹²

Шкала обладает хорошим показателем надёжности. Средняя корреляция балла за задание и способности высокая: с заданиями чаще соглашаются участники с большим уровнем способности; нет инверсии. Средняя трудность заданий говорит о том, что в целом трудности заданий соответствуют способностям респондентов. Значение первого контраста в остатках после извлечения основной Раш-размерности свидетельствует о том, что остатки формируют случайный шум, то есть тест преимущественно одномерен. Значения статистик согласия RMSD и Infit MNSQ свидетельствует о том, что респонденты в своих ответах на задания дают мало непредсказуемых, не согласующихся с выбранной моделью измерения ответов.

В таблице 4 представлены значения порогов политомических заданий. В целом пороги упорядочены: каждая ответная категория имеет свою область представленности на шкале; более высокий балл набрать труднее, чем более низкий. У одного задания — *plan_15* — пороги 1 и 2 практически тождественны: это может быть связано с тем, что задание достаточно лёгкое, и набрать 1 и 2 балла, в связи с этим, можно с сопоставимой вероятностью.

7 Linacre, J. M. (2025). A User’s Guide to WINSTEPS® MINISTEP Rasch-Model Computer Programs. Program Manual 5.9.1. URL: <https://www.winsteps.com/a/Winsteps-Manual.pdf> (дата обращения: 27.03.2025)

8 Thorndike, R. M., Thorndike-Christ, T. M. (2013). Measurement and evaluation in psychology and education. Pearson Education Limited.

9 Matlock-Hetzl, S. (1997). Basic Concepts in Item and Test Analysis.

10 Crocker, L., Algina, J. (1986). Introduction to classical and modern test theory. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando

11 Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñoz, J., Sjöberg, A. (2013). EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. Version 4.2, 6.

12 Linacre, J. M. (2025). A User’s Guide to WINSTEPS® MINISTEP Rasch-Model Computer Programs. Program Manual 5.9.1. URL: <https://www.winsteps.com/a/Winsteps-Manual.pdf> (дата обращения: 27.03.2025)

Таблица 4. Значения порогов

Задание	Порог 1	Порог 2	Порог 3
plan_15_PR	-.084	-.118	.203
plan_16_PR	-.444	.444	
plan_17_K	-.225	.225	
plan_18_K	-.573	.573	

На рисунке 1 представлена карта переменных. В левой части рисунка представлено распределение способностей; в правой — трудности категорий заданий. Трудности и способности лежат на одной шкале логитов. Можно сделать вывод, что распределение трудностей категорий достаточно хорошо покрывает распределение способностей, хотя стоило бы добавить в тест несколько очень трудных заданий для охвата верхней части распределения.

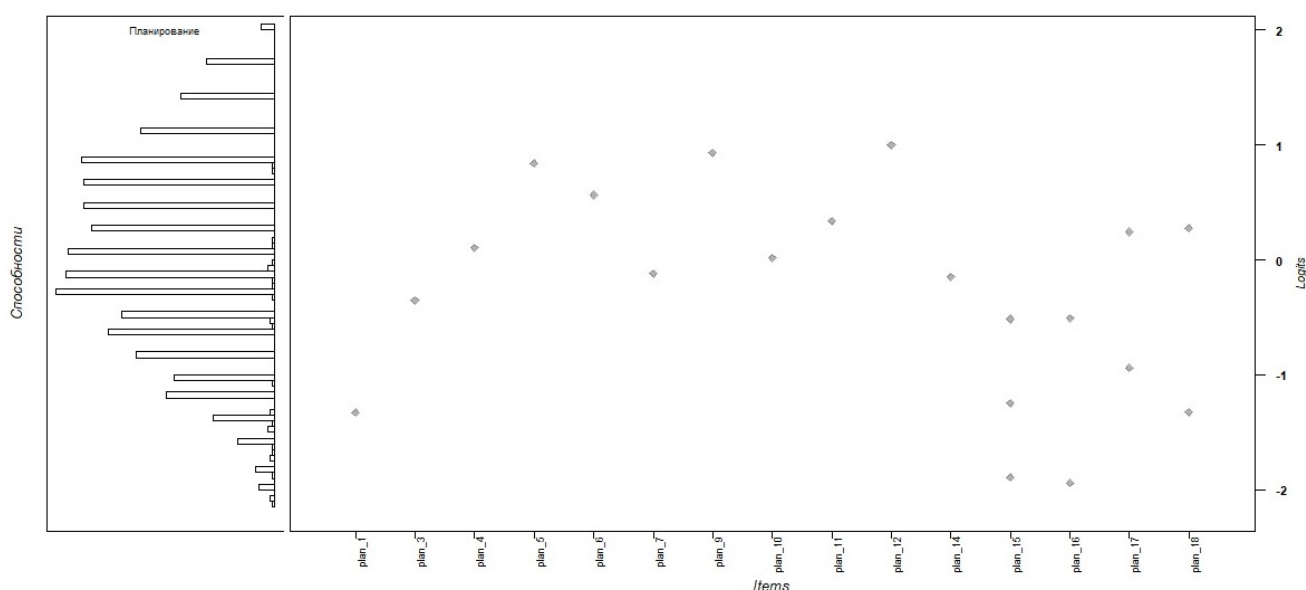


Рисунок 1. Карта переменных

Дополнительный анализ в рамках классической теории тестирования

Для лучшего представления о работе теста был дополнительно проведён анализ итоговой версии в рамках КТТ. В таблице 5 представлены основные параметры теста. В Приложении 3 представлены трудность и дискриминативность отдельных заданий (рассчитанные в рамках КТТ).

Надёжность теста в рамках классической теории тестирования также можно назвать достаточно высокой: ответы участников на разные задания позволяют сделать вывод, что они достаточно хорошо согласованы между собой. Трудность заданий распределена в широком диапазоне. По средней трудности теста, распределению суммарного балла и среднему баллу можно заключить, что тест примерно соответствует уровню подготовленности респондентов. Средняя дискриминативность является высокой, минимальная не опускается ниже рекомендованного порога: на все задания чаще отвечают верно участники с более высоким суммарным баллом по тесту. Распределение суммарного балла близко к нормальному; есть некоторая смещённость вправо, свидетельствующая о небольшом преобладании более высоких баллов. В таблице 6 представлены общие характеристики параметров ответных опций. В Приложении 4 приведены параметры ответных опций (посчитанные в рамках КТТ).

Таблица 5. Общие параметры инструмента (классическая теория тестирования)

Параметр	Среднее	Минимум	Максимум	Рекоменд. значение
Балл по тесту	11,3	0	20 (из 20)	–
Трудность	.53	.31	.81	[.1; 0,9] ¹³
Дискриминативность	.48	.21	.66	≥ .29 ^{14 15 16}

Параметр	Значение	Рекоменд. значение
Стандартное отклонение	4.36	–
Ошибка измерения	3.89	–
Надёжность: альфа Кронбаха	.78	≥ .612 ¹⁷
Надёжность: омега Макдональда	.78	≥ .613 ¹⁸

На рисунке 2 представлено распределение суммарного балла по тесту.

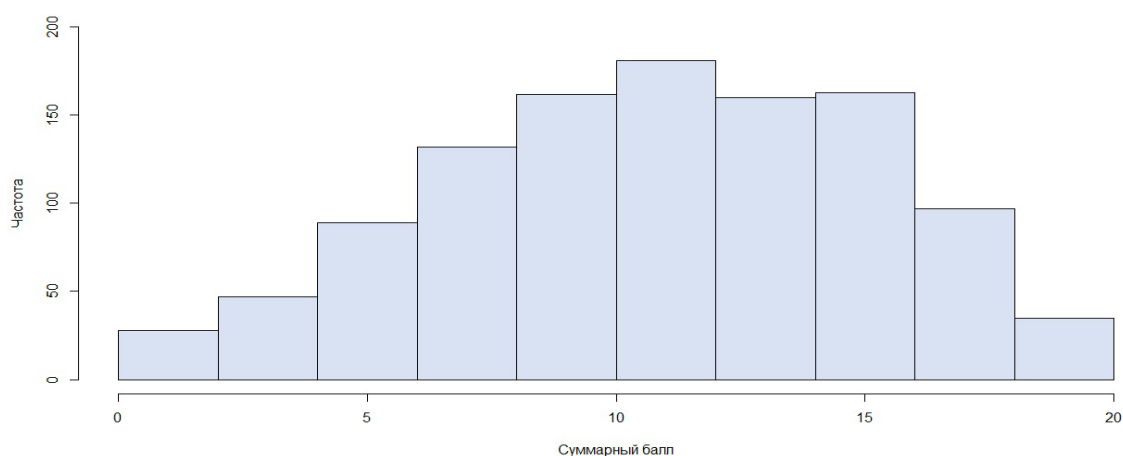


Рисунок 2. Распределение суммарного балла

Анализ ответных опций свидетельствует о том, что неправильные варианты ответа в заданиях соответствуют критериям: а) участники их выбирают, б) их выбирают участники с более низким уровнем подготовленности. Некоторые дистракторы нескольких заданий (plan_1, plan_7, plan_10) выбирают не очень большая доля участников (1%–2%). Принято решение не удалять их, поскольку они выполняют свою функцию, привлекая внимание менее подготовленных респондентов, и их удаление может сделать задания легче. Верные варианты ответа всех заданий обладают высокой дискриминативностью и достаточным уровнем трудности, что подтверждает правильное функционирование заданий.

Анализ в рамках как классической, так и современной теории тестирования говорит о хороших психометрических параметрах теста. Анализ в КТТ дополнительно продемонстрировал хорошее функционирование ответных опций итогового банка.

13 Аббакумов, Д. Ф. (2023). Трудность тестового задания. Большая российская энциклопедия. URL: <https://bigenc.ru/c/trudnost-testovogo-zadaniia-da0711> (дата обращения: 27.03.2025)

14 Thorndike, R. M., Thorndike-Christ, T. M. (2013). Measurement and evaluation in psychology and education. Pearson Education Limited.

15 Matlock-Hetzel, S. (1997). Basic Concepts in Item and Test Analysis.

16 Crocker, L., Algina, J. (1986). Introduction to classical and modern test theory. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlan-do, FL 32887.

17 Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J., Sjöberg, A. (2013). EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. Version 4.2, 6.

18 Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J., Sjöberg, A. (2013). EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. Version 4.2, 6.

Таблица 6. Статистика по параметрам ответных опций

Типы ответов	Дискриминативность ответных опций				Трудность ответных опций			
	Сред.	Мин.	Макс.	Рекомендуемые значения	Сред.	Мин.	Макс.	Рекомендуемые значения
Дистракторы	-.21	-.37	-.05	< 014 15	.11	.01	.45	> .05 ^{21 22}
Верные ответы	.42	.21	.52	> .2 ^{23 24 25} (для единственного верного ответа)	.40	.21	.76	[.1; .9] ²⁶ (для единственного верного ответа)

Анализ дополнительных свидетельств валидности

Конвергентная валидность

Свидетельства конвергентной валидности получены при помощи анализа взаимосвязей с другими конструктами. Взаимосвязь проверена при помощи коэффициента корреляции Пирсона. Результаты корреляционного анализа представлены в таблице 7.

Таблица 7. Корреляции с другими конструктами

Конструкт	Определение	r Пирсона	p
Коммуникативная грамотность	Владение культурными нормами общения, чёткое и структурированное формулирование мыслей, учёт особенностей собеседников	.53	.000
Анализ информации	Анализ и корректная работа с различного рода информацией, установление взаимосвязей между разрозненными данными	.43	.000
Абстрактное и системное мышление	Логическое мышление, выявление взаимосвязи в неструктурированных данных, формирование целостной картины происходящего	.42	.000
Управление	Склонность влиять и делать так, чтобы другие следовали определённой точке зрения	.15	.000
Вызов	Склонность заниматься чем-то достаточно сложным, требующим усилий, но интересным	.15	.000
Самоорганизация	Склонность погружаться в детали работы и следовать заданным срокам	.11	.001
Тревожность	Склонность к тревожным переживаниям и способность отключаться от имеющихся трудностей в работе	-.14	.000

Предполагалось, что способность планирования будет связан с другими когнитивными способностями человека. Предположение подтвердилось. Баллы по тесту организационного планирования, действительно, в первую очередь связаны с результатами по другим когнитивным тестам. Это связано как с некоторым содержательным пересечением конструктов, так и с тем, что может задействоваться некий общий мыслительный процесс, связанный именно с решением задач.

Предполагалось, что планирование продемонстрирует небольшую позитивную взаимосвязь с близкими по смыслу личностными чертами: самоорганизацией, ценностями управления и вызова, и небольшую отрицательную взаимосвязь — с тревожностью. Предположение также подтвердилось. Взаимосвязь с личностными чертами — менее выраженная, чем с когнитивными. Отчасти это можно объяснить эффектом включения разных когнитивных процессов, эффектом формата, ограничениями формата самоотчётных методик. При этом взаимосвязи, хотя и небольшие, также представляют интерес, поскольку ложатся в концепцию номологической сети конструкта (Kock, Verbekova, Assaf, Josiassen, 2024). Проинтерпретировать взаимосвязи можно следующим образом.

- Коммуникативная грамотность ($r = 0,53$). Планирование требует эффективного обмена информацией, согласования целей и ресурсов с другими.
- Анализ информации ($r = 0,43$). Умение систематизировать и интерпретировать данные помогает выстраивать реалистичные планы.
- Абстрактное и системное мышление ($r = 0,42$). Оценка сложных систем и взаимосвязей критична для определения наилучшего пути достижения целей.

- Управление ($r = 0,15$). Ориентация на лидирующую позицию способствует координации и реализации планов.
- Вызов ($r = 0,15$). Ориентация на преодоление трудностей мотивирует продуманное планирование.
- Самоорганизация ($r = 0,11$). Умение управлять собой способствует порядку и последовательности в планах. Это содержательно близкий планированию конструкт, измеряющий похожий феномен, концентрирующийся на личностных особенностях.
- Тревожность ($r = -0,14$). Высокий уровень тревожности снижает способность сосредоточиться и принимать обоснованные решения, что мешает планированию.

Критериальная валидность

Кроме того, в рамках проведённых работ нами были получены свидетельства критериальной валидности. Данные о критериях собирались при помощи анкеты, предъявляемой до начала тестирования. Анализ проведён при помощи ANOVA. Включены следующие переменные.

- Пол: мужской и женский;
- Возраст: младшая возрастная группа (до 44 лет), средняя возрастная группа (от 45 до 54 лет), старшая возрастная группа (55 и более лет). Интервалы приняты, отталкиваясь от интервалов возраста, выделенных Институтом статистики ЮНЕСКО (UNESCO Institute of Statistics, 2015), и при учёте того, что людей младше 35–40 лет в выборке практически нет;
- Номинация проекта: руководители и бухгалтеры.

Результаты анализа представлены в таблице 8.

Таблица 8. Результаты модели ANOVA

Предиктор	<i>df</i>	<i>F</i>	<i>p</i>	Размер эффекта (partial η^2)
Константа	1	658	.000	
Пол	1	3.59	.003	.01
Возраст	2	6.17	.000	.03
Номинация: руководитель	1	5.86	.023	.01

Значимы все включённые в анализ предикторы. С точки зрения размера эффекта все предикторы значимо вкладываются в зависимую переменную ($\eta^2 \geq 0,1$, что соответствует 1% объяснённой дисперсии зависимой переменной).

В таблице 9 приведены результаты попарных сравнений по группам, представленным в независимых переменных. Сравнения проведены при помощи теста Тьюки. Приведены только статистически значимые результаты. Результаты сравнения представлены в *T*-баллах (шкала со средним значением, равным 500 баллам, стандартным отклонением, равным 100 баллам, и минимумом и максимумом, равным 200 и 800 баллам соответственно).

Таблица 9. Попарные сравнения по полу, возрасту и должности

Переменная	Пара сравнения	Разница (<i>T</i> -баллы)	<i>p</i>
Пол	Мужской — женский	33	.004
Возраст	Младшая возрастная группа — старшая возрастная группа	45	.001
	Средняя возрастная группа — старшая возрастная группа	40	.001
Должность	Руководители — бухгалтеры	20	.035

На основе анализа литературы предполагалось, что планирование у работающих мужчин не будет значимо отличаться от результатов работающих женщин (Eagly, Woo, Diekman, 2012; Powell, Pharris, Hardy, 2020). Однако это предположение не подтвердилось. Мужчины могут

демонстрировать более высокие результаты в планировании из-за культурных и социальных факторов, где им чаще приписываются роли, связанные со стратегическим мышлением и управлением (Hsu, Badura, Newman, Speach, 2021).

Ожидалось, что более молодые специалисты в организациях покажут более высокие результаты по тесту. Современная культура подчёркивает важность эффективного использования времени, что мотивирует молодых сотрудников организаций развивать способность планирования. Молодые сотрудники чаще используют приложения и программы для планирования, что помогает им эффективно организовывать свои задачи и ценить способность планирования (Gouda, 2020). Наконец, у более молодых групп выше скорость обработки информации и когнитивная гибкость, что способствует успешному планированию; с возрастом же некоторые когнитивные способности снижаются, что возможно объясняет наблюдаемые различия (Harada, Love, Triebel, 2013). Предположение подтвердилось.

Ожидалось, что руководители покажут более высокий результат по сравнению с исполнительскими позициями. Предположение подтвердилось. Руководители, как правило, сосредоточены на постановке целей и распределении ресурсов, что требует хорошо развитого планирования.

Содержательная валидность

Свидетельства содержательной валидности были получены в ходе когнитивных интервью, а также в ходе дополнительной экспертной оценки теста.

Когнитивные интервью

В качестве одного из этапов разработки и для получения первых свидетельств содержательной валидности была проведено семь когнитивных интервью, участники которых самостоятельно оценивали свои навыки планирования от уровня «Ниже среднего» до уровня «Очень высокие». Респонденты подбирались таким образом, чтобы в когнитивных интервью приняли участие как те, кто оценивает свои навыки высоко, так и те, кто оценивает их достаточно низко. В ходе сессии было необходимо:

- решить задание;
- кратко написать решение;
- оценить свою уверенность в правильности ответа от 1 (совсем не уверен) до 5 (полностью уверен);
- оценить воспринимаемую трудность решения задания от 1 (очень простое) до 5 (очень сложное);
- соотнести задание с измеряемым им конструктом (планированием или другим, если респондент сочтёт нужным);
- соотнести задание с измеряемым индикатором (индикаторами планирования — планированием времени, оценкой ресурсов для выполнения плана, ранжированием задач, корректировкой выполнения плана, — либо другими, если респондент сочтёт нужным).

На основе обратной связи респондентов в задания были внесены изменения, а часть заданий, вызвавших затруднения или неоднозначные интерпретации, были исключены. В целом участники когнитивных интервью корректно соотносили задания с заявленными конструктами и индикаторами, что также может выступать свидетельствами содержательной валидности.

Экспертная оценка заданий

Для дальнейшего сбора свидетельств содержательной валидности задания теста были предложены на рецензирование трём внешними экспертами в области управления

проектами. Эксперты обладают значительным опытом работы в данной сфере, включая преподавание курсов по управлению проектами в ведущих вузах, а также практическую деятельность в крупных компаниях. Им были предоставлены все материалы теста, включая теоретическую рамку, инструкции и задания с указанием индикаторов.

Эксперты независимо друг от друга оценивали задания по следующим критериям:

- корректность ключей к заданиям;
- соответствие заданий измеряемому конструкту
- соотнесение заданий с конкретными индикаторами планирования (планирование времени, оценка ресурсов, приоритизация задач, корректировка выполнения плана);
- соответствие заданий реальным ситуациям, требующим применения навыков планирования.

Эксперты оставили комментарии и предложения по улучшению инструмента. Они подтвердили, что содержание заданий и теоретическая рамка теста носят универсальный характер и могут быть применены для оценки навыков планирования у широкого круга специалистов. Все вопросы теста, которые вошли в финальную версию, по мнению экспертов, понятны, соответствуют индикаторам, а также направлены на измерение навыков планирования.

Заключение

Результаты проведённого исследования подтверждают, что ТОП является надёжным инструментом для оценки навыков планирования с подтверждёнными свидетельствами валидности. Психометрический анализ продемонстрировал корректное функционирование заданий, их хорошую дифференцирующую способность, а также подтвердил одномерную структуру инструмента.

Анализ валидности подтвердил ожидаемые связи ТОП с другими конструктами — когнитивными и личностными. Исследование критериальной валидности почти во всех случаях выявило ожидаемые различия в результатах по возрасту и профессиональному статусу, кроме результатов по полу. Когнитивные интервью с респондентами и экспертная оценка содержания инструмента показали, что задания корректно отражают ключевые аспекты планирования.

Важным преимуществом ТОП над самоотчётными методиками, при помощи которых зачастую измеряется планирование, является то, что он способен более объективно оценить уровень планирования, исключая влияние субъективных факторов, таких, например, как социальная желательность и так называемый «эффект мужского высокомерия и женской скромности» (*male hubris, women humility effect*) (Reilly et al., 2022; Syzmanowicz, Furnham, 2011). Кроме того, использование формата теста способностей позволяет избежать статистических артефактов, таких как «эффект потолка», встречающийся в диагностических опросниках.

Научно-практическая значимость ТОП заключается в сочетании объективности теста способностей с экологической валидностью «горячего» контекста. Методика позволяет измерять когнитивный компонент планирования в условиях, приближённых к организационной реальности, преодолевая ограничения как субъективных опросников, так и лабораторных нейропсихологических проб. Кроме того, ТОП создан с соблюдением строгих требований к разработке и валидации тестов, и его психометрические характеристики позволяют говорить о нём как о надёжном и достаточно объективном инструменте оценки, практически пригодном в организационном контексте.

Таким образом, ТОП является научно обоснованным инструментом для оценки навыков планирования. Инструмент применим для использования как в исследовательских, так и в практических целях.

Ограничения исследования

Ограничениями исследования является смещение географического состава выборки в сторону Центрального и Сибирского федеральных округов, а также преобладание среди респондентов профильных специалистов. Выводы исследования, основанные на данных конкретной выборки, требуют дополнительного изучения для оценки их применимости к другим группам населения. Перенос этих результатов на другие группы требует учёта их специфики и возможных различий. Необходимы дальнейшие исследования функционирования теста на более разнообразных выборках.

Будущие исследования могут быть направлены на адаптацию теста к различным профессиональным группам (особенно к таким, деятельность которых непосредственно не связана с планированием), уточнение нормативных данных и изучение влияния культурных особенностей. Дальнейшее изучение взаимосвязей способности планирования с личностными и когнитивными характеристиками также представляет перспективный вектор исследования данного диагностического инструмента. Возможно проведение анализа различий между участниками с разным уровнем управленческой ответственности, чтобы подтвердить, что тест действительно измеряет способность, необходимую на разных уровнях управления, например, в стратегическом управлении.

Литература

- Аббакумов, Д. Ф. (2023). Трудность тестового задания. В сб.: *Большая российская энциклопедия*. URL: <https://bigenc.ru/c/trudnost-testovogo-zadaniia-da0711> (дата обращения: 27.03.2025)
- Адаменко, А. А., Хорольская, Т. Е., Петров, Д. В. (2020). Роль стратегического планирования в развитии предпринимательской деятельности. *Естественно-гуманитарные исследования*, 3(29), 24–28.
- Батурин, Н. А., Вучетич, Е. В., Костромина, С. Н., Кукаркин, Б. А., Куприянов, Е. А., Лурье, Е. В., Митина, О. В., Науменко, А. С., Орел, Е. А., Полетаева, Ю. С., Попов, А. Ю., Потапкин, А. А., Симоненко, С. И., Сеницына, Ю. Д., Шмелев, А. Г. (2015). Российский стандарт тестирования персонала (временная версия, созданная для широкого обсуждения в 2015 году). *Организационная психология*, 5(2), 67–138.
- Вилкова, К. А. (2020). Измерение саморегулируемого обучения: обзор инструментов. *Современная зарубежная психология*, 9(2), 123–133.
- Моросанова, В. И., Кондратюк, Н. Г. (2020). К вопросу о факторной структуре краткой версии опросника В. И. Моросановой «Стиль саморегуляции поведения — ССПМ». В сб.: Т. Н. Банщикова, Е. А. Фомина, В. И. Моросанова (ред. колл.), *Личностные и регуляторные ресурсы достижения образовательных и профессиональных целей в эпоху цифровизации (20-23 октября 2020 г.)* (178–188). М.: Знание-М.
- Смарт Бизнес Солюшнс. Организационный тест. Полная и экспресс версия. *HR-маркет*. URL: <https://sbsc.ru/hrm/orgtest.html>
- Россия — страна возможностей. Департамент оценки и методологии. (2026). *Каталог инструментов оценки и диагностики*. URL: <https://seafire.rsv.ru:9443/d/3e36d466a8534ec8ba2d/> (дата обращения: 10.02.2026)
- Руководство к своду знаний по управлению проектами (Руководство РМВОК-6)*. (2020). М. : Олимп-бизнес.

- Штроо, В. А. (2025). *Методы активного социально-психологического обучения : учебник и практикум для академического бакалавриата*. М.: Юрайт.
- Ясин, М. И., Бажданова, Ю. В. (2024). Методика измерения долгосрочности и краткосрочности мотивации. *Организационная психология*, 14(1), 96–108.
- A Guide to the Project Management Body of Knowledge — Seventh Edition*. (2025). Moscow : ABC.
- Aarts, H., Dijksterhuis, A. P., Midden, C. (1999). To plan or not to plan? Goal achievement or interrupting the performance of mundane behaviors. *European Journal of Social Psychology*, 29(8), 971–979.
- AERA, APA, NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Washington, DC: AERA
- Allan, J. L., Sniehotta, F. F., Johnston, M. (2013). The best laid plans: Planning skill determines the effectiveness of action plans and implementation intentions. *Annals of Behavioral Medicine*, 46(1), 114–120.
- Andrich, D., Marais, I. (2019). A course in Rasch measurement theory. Measuring in the educational, social and health sciences, 41(8).
- Antony, J. (2000). Ten key ingredients for making SPC successful in organisations. *Measuring Business Excellence*, 4(4), 7–10.
- Barati, A. A., Asadi, A., Kalantari, K., Azadi, H., Mamoorian, M. (2014). Analyzing the impacts of agricultural land use change according to the experts opinion of agricultural land organization in Iran. *Iranian Journal of Agricultural Economics and Development Research*, 45(4), 639–650.
- Bertua, C., Anderson, N., Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational psychology*, 78(3), 387–409.
- Bock, R. D., Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443–459.
- Bock, R. D., Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied psychological measurement*, 6(4), 431–444.
- Cao, D. S., Schüz, N., Xie, G. R., Lippke, S. (2013). Planning skills moderate the intention–planning cognitions–behaviour relation: a longitudinal study on physical activity in Chinese adolescents. *Research in Sports Medicine*, 21(1), 12–23.
- Collins, A., Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology*, 10(3), e1001293.
- Costa, P. T., McCrae, R. R. (2014). The NEO Inventories. In *Personality assessment* (229–260). Routledge.
- Costa, P. T. (1996). Work and personality: Use of the NEO-PI-R in industrial / organisational psychology. *Applied Psychology*, 45(3), 225–241.
- Crocker, L., Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.
- Eagly, A. H., Woo, W., Diekmann, A. B. (2012). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes, H. M. Trautner (Eds.), *The developmental social psychology of gender* (123–174). New York: Psychology Press.
- Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J., Sjöberg, A. (2013). *EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. Version 4.2*, 6.
- Evers, A., Muñiz, J., Hagemester, C., Høstmælingen, A., Lindley, P., Sjöberg, A., Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema*, 25(3), 283–291.
- Frederiksen, N., Saunders, D. R., Wand, B. (1957). *The in-basket test*. *Psychological Monographs*, 71(438). <https://doi.org/10.1037/h0093706>

- Gouda, H. (2020). Exploring the relevant skills needed for the digital age employees. *European Journal of Business and Management*, 12(2), 58–67.
- Gross, R. G., Grossman, M. (2010). Executive resources. *Continuum: Lifelong Learning in Neurology*, 16(4), 140–152.
- Haladyna, T. M., Downing, S. M. (1988). *Functional Distractors: Implications for Test-Item Writing and Test Design*.
- Harada, C. N., Love, M. C. N., Triebel, K. (2013). Normally cognitive aging. *Clinics in geriatric medicine*, 29(4), 737.
- Hayes-Roth, B., Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science*, 3, 245—262.
- Hsu, N., Badura, K. L., Newman, D. A., Speech, M. E. P. (2021). Gender, “masculinity,” and “femininity”: A meta-analytic review of gender differences in agency and communion. *Psychological Bulletin*, 147(10), 987.
- Joshi, A., Kale, S., Chandel, S., Pal, D. K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4), 396–403
- Karat, J. (1982). A model of problem solving with incomplete constraint knowledge. *Cognitive Psychology*, 14, 538–559.
- Kerzner, H. (2019). *Using the Project Management Maturity Model: Strategic Planning for Project Management*. John Wiley & Sons.
- Kock, F., Berbekova, A., Assaf, A. G., Josiassen, A. (2024). Developing a scale is not enough: on the importance of nomological validity. *International Journal of Contemporary Hospitality Management*, 36(10), 3233–3244.
- Köhler, C., Robitzsch, A., Hartig, J. (2020). A bias-corrected RMSD item fit statistic: An evaluation and comparison to alternatives. *Journal of Educational and Behavioral Statistics*, 45(3), 251–273.
- Linacre, J. M. (2025). *A User's Guide to WINSTEPS® MINISTEP Rasch-Model Computer Programs. Program Manual 5.9.1*. URL: <https://www.winsteps.com/a/Winsteps-Manual.pdf> (дата обращения: 27.03.2025)
- Masters, G. N. (2016). Partial credit model. In W. J. Linden, R. K. Hambleton (Eds.), *Handbook of item response theory* (109–126). Chapman and Hall/CRC.
- Matlock-Hetzel, S. (1997). *Basic Concepts in Item and Test Analysis*.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel psychology*, 60(3), 683–729.
- Morris, R., Ward, G. (Eds.). (2004). *The cognitive psychology of planning*. New York : Psychology Press.
- Nassif, V. M. J., Hashimoto, M., Amaral, D. J. (2014). Entrepreneurs self-perception of planning skills: evidences from Brazilian entrepreneurs. *Revista Ibero Americana de Estratégia*, 13(4), 107–121.
- The OPC (2025). Graduate Situational Judgement Test (GSJT). *opc assessment*. URL: <https://www.theopc.co.uk/assessment/test/gsjt/> (дата обращения: 28.10.2025)
- Owen, A. (1997). Cognitive planning in humans: Neuropsychological, neuro anatomical and neuropharmacological perspectives. *Prog Neurobiol*, 53, 431–450
- Papa-Gusho, L., Biçaku-Çekrezi, R. (2015). Factors that affect effective planning skills of the teacher in the classrooms. *Academic Journal of Interdisciplinary Studies*, 4(3), 560–564.
- Pellerin, R., Perrier, N. (2019). A review of methods, techniques and tools for project planning and control. *International Journal of Production Research*, 57(7), 2160–2178.
- Phillips, L. H., MacLeod, M. S., Kliegel, M. (2005). Adult aging and cognitive planning. *The cognitive psychology of planning*, 111–134.

- Powell, J. W., Pharris, L. J., Hardy, M. M. (2020). A Comparison of time management skills among accounting, business, and information systems students by age and gender. *Issues in Information Systems*, 21(3), 1–10.
- Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of experimental child psychology*, 108(3), 621–637.
- Reilly, D., Neumann, D. L., Andrews, G. (2022). Gender differences in self-estimated intelligence: Exploring the male hubris, female humility problem. *Frontiers in Psychology*, 13.
- Roshanai, A., Janeslätt, G., Marie White, S., Lidstroem Holmqvist, K., Holmfur, M. (2021). Psychometric properties of the original version of the assessment of time management skills. *The American Journal of Occupational Therapy*, 75(Supplement_2), 7512500038p1-7512500038p1.
- Shallice, T. (1982). specific impairments of planning. *philosophical transactions of the Royal Society of London. B, Biological Sciences*, 298(1089), 199–209.
- Sorel, O., Pennequin, V. (2008). Aging of the planning process: The role of executive functioning. *Brain and cognition*, 66(2), 196–201.
- Syzmanowicz, A., Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta-analyses. *Learning and Individual Differences*, 21(5), 493–504.
- Testa, S., Toscano, A., Rosato, R. (2018). Distractor efficiency in an item pool for a statistics classroom exam: Assessing its relation with item cognitive level classified according to Bloom's taxonomy. *Frontiers in psychology*, 9, 1585.
- Thorndike, R. M., Thorndike-Christ, T. M. (2013). *Measurement and evaluation in psychology and education*. Pearson Education Limited.
- UNESCO Institute of Statistics (2015). Researchers by age. *UNESCO UIS*. URL: <https://uis.unesco.org/node/3080143> (дата обращения: 31.03.2025)
- Verulava, T. (2023). Strategic management and planning for health care organizations in Georgia. *Asia Pacific Journal of Health Management*, 18(3), 41–48.
- Viswesvaran, C., Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and psychological measurement*, 59(2), 197–210.
- Vita, S., Coplin, H., Feiereisel, K. B., Garten, S., Mechaber, A. J., Estrada, C. (2013). Decreasing the ceiling effect in assessing meeting quality at an academic professional meeting. *Teaching and learning in medicine*, 25(1), 47–54.
- Wakita, T., Ueshima, N., Noguchi, H. (2012). Psychological distance between categories in the Likert scale: Comparing different numbers of options. *Educational and psychological measurement*, 72(4), 533–546.
- Ward, L. B. (1959). The business in-basket test: A method of assessing certain administrative skills. *ETS Research Bulletin Series*, 1959(1), i-49.
- Zelazo, P. D., Cunningham, W. (2007). Executive function: Mechanisms underlying emotion regulation. In J. J. Gross, R. A. Thompson (Eds.), *Handbook of emotion regulation* (135–158). Guilford Press.

Поступила 31.03.2025

Приложения

Приложение 1. Пример заданий — демонстрационный вариант (в тесте не используются)

Перед сотрудниками частного вуза стоит задача организовать конференцию на 150 человек. Вуз обратился в компанию, специализирующуюся на организации мероприятий. Бюджет вуза, заложенный на проведение конференции: 500 000 рублей. Вуз хочет, чтобы за эти деньги получилось организовать конференцию настолько по высшему разряду, насколько возможно: задачи сэкономить заложенный бюджет не стоит.

В таблице ниже представлены характеристики предложенных пакетов.

Пакет	Зал (чел.)	Кофе-брейк	Обед	Съемка	Стоимость пакета (руб.)
Пакет «А»	120	Премиум	Стандарт	Фото, видео	500 000
Пакет «Б»	160	Стандарт	Эконом	Фото	450 000
Пакет «В»	150	Эконом	Эконом	Не предоставляется	380 000
Пакет «Г»	160	Стандарт	Стандарт	Фото	500 000
Пакет «Д»	155	Премиум	Премиум	Фото, видео	600 000

Какой пакет в большей степени подходит под приведенные условия? Выберите один вариант ответа.

- Пакет «А»
- Пакет «Б»
- Пакет «В»
- Пакет «Г»
- Пакет «Д»

Менеджер съемочной группы рекламного клипа составил календарь со сроками выполнения разных этапов одного из предстоящих проектов, учитывающий очередность задач: начать следующую возможно по завершении предыдущей. Однако финальный срок не устроил заказчика.

	Сентябрь													
	5 чт	6 пт	7 сб	8 вс	9 пн	10 вт	11 ср	12 чт	13 пт	14 сб	15 вс	16 пн	17 вт	18 ср
Сценарий клипа														
Выбор актеров и площадки														
Съемки клипа														
Монтаж кадров														
Запись звуковой дорожки														
Подготовка документации														

К какой дате станет возможным завершение проекта, если сотрудники будут работать по субботам, а запись звуковой дорожки начнут одновременно с монтажом кадров? Выберите один вариант ответа.

- 12 сентября
- 14 сентября
- 15 сентября
- 16 сентября

Первого июня менеджер проекта согласовал с руководством ключевые цели на ближайшие 2 месяца:

- Своевременная подготовка квартального отчета.
- Разработка и запуск собственного сайта проекта.

Какие ТРИ задачи следует выполнить команде в первую очередь, исходя из их важности и срочности? Выберите подходящие варианты ответа.

- Задача 1 — Подготовка первой версии квартального отчета. Срок: до 10 июня
- Задача 2 — Согласование бюджета на разработку сайта проекта. Срок: до 15 июня
- Задача 3 — Проведение опроса сотрудников для сбора данных по использованию текущих информационно-коммуникационных ресурсов компании. Срок: до 30 июля
- Задача 4 — Тестирование прототипа сайта с ключевыми пользователями. Срок: до 15 июля
- Задача 5 — Сбор отчетных данных по отделам для подготовки квартального отчета. Срок: до 5 июня
- Задача 6 — Обучение сотрудников новым инструментам работы с сайтом проекта. Срок: до 25 июля
- Задача 7 — Оформление окончательной версии годового отчета. Срок: до 30 июля

Приложение 2. Параметры заданий теста (IRT)

Задание	Трудность	M	SE	Infit	Связь со способностью
plan_1_O	-1.33	.76	.08	.99	.41
plan_3_O	-.35	.58	.07	.98	.45
plan_4_K	.11	.48	.07	1.05	.39
plan_5_O	.84	.33	.07	1.03	.40
plan_6_PL	.56	.39	.07	.96	.49
plan_7_O	-.12	.53	.07	.98	.46
plan_9_PL	.93	.31	.07	1.17	.23
plan_10_PL	.02	.51	.07	1.07	.38
plan_11_PR	.34	.43	.07	1.01	.44
plan_12_K	1.00	.31	.07	.92	.52
plan_14_PR	-.15	.54	.07	.94	.51
plan_15_PR	-1.21	.81	.10	.99	.61
plan_16_PR	-1.23	.77	.09	.96	.58
plan_17_K	-.35	.60	.07	.96	.61
plan_18_K	-.52	.63	.08	.94	.65

Приложение 3. Параметры заданий теста (КТТ)

Задание	M	Дискриминативность
plan_1_O	.76	.43
plan_3_O	.58	.47
plan_4_K	.48	.39
plan_5_O	.33	.39
plan_6_PL	.39	.49
plan_7_O	.53	.47
plan_9_PL	.31	.21
plan_10_PL	.51	.37
plan_11_PR	.43	.44
plan_12_K	.31	.5
plan_14_PR	.54	.52
plan_15_PR	.81	.63
plan_16_PR	.77	.59
plan_17_K	.60	.62
plan_18_K	.63	.66

Приложение 4. Параметры ответных опций (КТТ)

Задание	Параметры задания				
	Ключ	Варианты ответа	n	Труд.	Дискр.
plan_1_O	*	1	814	.76	.43
		2	14	.01	-.15
		3	163	.15	-.37
		4	40	.04	-.17
		5	38	.04	-.16
plan_3_O	*	1	64	.06	-.24
		2	205	.19	-.27
		3	619	.58	.47
		4	90	.08	-.20
		5	62	.06	-.18
		6	28	.03	-.17

plan_4_K	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	104	.10	-.19
		2	516	.48	.39
		3	225	.21	-.17
		4	221	.21	-.34
plan_5_O	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	37	.03	-.14
		2	359	.34	.39
		3	345	.32	-.17
		4	127	.12	-.30
plan_6_PL	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	414	.39	.49
		2	418	.39	-.18
		3	80	.08	-.25
		4	85	.08	-.26
plan_7_O	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	45	.04	-.17
		2	565	.53	.47
		3	237	.22	-.29
		4	196	.18	-.29
plan_9_PL	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	54	.05	-.10
		2	340	.32	.21
		3	483	.45	-.13
		4	195	.18	-.23
plan_10_PL	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	247	.23	-.21
		2	159	.15	-.16
		3	93	.09	-.24
		4	536	.50	.37
		5	15	.01	-.17
plan_11_PR	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	131	.12	-.26
		2	155	.14	-.16
		3	465	.43	.44
		4	74	.07	-.26
		5	186	.17	-.12
plan_12_K	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	52	.05	-.22
		2	210	.20	-.22
		3	76	.07	-.22
		4	180	.17	-.24
		5	327	.31	.50
plan_14_PR	Ключ *	Варианты ответа	n	Труд.	Дискр.
		1	39	.04	-.25
		2	58	.05	-.30
		3	65	.06	-.27
		4	203	.19	-.27
		5	573	.54	.52
		Варианты ответа	n	Труд.	Дискр.
		6	129	.12	-.13

plan_15_PR	Ключ	Варианты ответа	n	Труд.	Дискр.	
		1	126	.03	-.25	
		2	291	.07	-.05	
		3	208	.05	-.18	
		*	4	827	.21	.36
		*	5	597	.15	-.11
		*	6	956	.24	.45
		*	7	103	.03	-.15
plan_16_PR	Ключ	Варианты ответа	n	Труд.	Дискр.	
		1	208	.09	-.23	
		2	163	.07	-.33	
		*	3	810	.36	.31
			4	56	.02	-.23
			5	125	.06	-.28
		*	6	835	.37	.44
			7	55	.02	-.17
plan_17_K	Ключ	Варианты ответа	n	Труд.	Дискр.	
		1	137	.07	-.20	
		*	2	704	.37	.40
		*	3	312	.17	-.34
			4	568	.30	.43
plan_18_K	Ключ	Варианты ответа	n	Труд.	Дискр.	
		*	1	593	.30	.41
			2	113	.06	-.32
			3	326	.16	-.22
			4	220	.11	-.15
		*	5	745	.37	.44



Development and validation of the “Organizational Planning Test”

Anastasia SADOVA

ORCID: <https://orcid.org/0000-0001-8472-5305>

Autonomous non-commercial organization «Russia — the country of opportunities», Moscow, Russia

Olga SOBOLEVA

ORCID: <https://orcid.org/0000-0002-2348-2925>

Autonomous non-commercial organization «Russia — the country of opportunities», Moscow, Russia

Alina IVANOVA

ORCID: <https://orcid.org/0000-0003-3340-7651>

HSE University, Moscow, Russia

Ekaterina STEPASHKINA

ORCID: <https://orcid.org/0000-0002-4340-3753>

Autonomous non-commercial organization «Russia — the country of opportunities», Moscow, Russia

Abstract. *Purpose.* This study presents the development and validation of the “Organizational Planning Test” (OPT), designed to assess planning competencies in working professionals (specialists and managers). *Methodology.* Employing contemporary psychometric approaches (classical test theory and item response theory), the test development followed a rigorous multi-stage procedure: (1) theoretical framework construction, (2) initial item pool generation, (3) think-aloud protocols, (4) expert review, (5) pilot testing ($N = 1,094$; 74% female, 26% male; $M = 49$, $SD = 9.2$), and (6) psychometric evaluation with validity evidence collection. *Findings.* Psychometric analyses using the partial credit model demonstrated: (1) Essential unidimensionality (first contrast eigenvalue = 1.33), (2) Excellent item fit (Infit MNSQ = .92–1.17), (3) Adequate discrimination parameters (.23–.65), and (4) Strong reliability ($\alpha = .78$, $\omega = .78$). Convergent validity evidence emerged through significant correlations with: information analysis ($r = .43$, $p < .001$), systems thinking ($r = .42$, $p < .001$), self-organization ($r = .15$, $p < .001$), and anxiety ($r = -.14$, $p < .001$). ANOVA confirmed criterion validity through expected group differences: gender ($\eta^2 = .01$, $p = .003$), age ($\eta^2 = .03$, $p = .000$), and position level ($\eta^2 = .01$, $p = .023$). *Implications for practice.* The OPT constitutes a psychometrically robust measure of planning ability, with demonstrated applications for: (1) personnel selection, (2) competency assessments, (3) professional development programs, and (4) cognitive research. Future studies should explore cross-occupational adaptability and relationships with broader cognitive-personality constructs. *Value of the results.* The instrument provides HR practitioners and organizational psychologists with a validated diagnostic tool for workplace planning competencies, while contributing to research on cognitive predictors of job performance.

Keywords: planning ability, cognitive abilities, psychometrics, item response theory, organizational psychology, psychological assessment.

References

- Aarts, H., Dijksterhuis, A. P., Midden, C. (1999). To plan or not to plan? Goal achievement or interrupting the performance of mundane behaviors. *European Journal of Social Psychology*, 29(8), 971–979.
- Abbakumov, D. F. (2023). Trudnost' testovogo zadaniya [Test item difficulty]. *Bol'shaya russiyskaya entsiklopediya*. Retrieved from <https://bigenc.ru/c/trudnost-testovogo-zadaniia-da0711> (in Russian)
- Adamenko, A. A., Khorolskaya, T. E., Petrov, D. V. (2020). Rol' strategicheskogo planirovaniya v razvitiu predprinimatelskoy deyatel'nosti [The role of strategic planning in the development of entrepreneurial activity]. *Estestvenno-gumanitarnye issledovaniya*, (3(29)), 24–28. (in Russian)
- AERA, APA, NCME. (2014). Standards for educational and psychological testing. *American Educational Research Association, American Psychological Association, National Council on Measurement in Education*. Washington, DC: AERA
- Allan, J. L., Sniehotta, F. F., Johnston, M. (2013). The best laid plans: Planning skill determines the effectiveness of action plans and implementation intentions. *Annals of Behavioral Medicine*, 46(1), 114–120.
- Andrich, D., Marais, I. (2019). A course in Rasch measurement theory. *Measuring in the educational, social and health sciences*, 41(8).
- Antony, J. (2000). Ten key ingredients for making SPC successful in organisations. *Measuring Business Excellence*, 4(4), 7–10.
- Barati, A. A., Asadi, A., Kalantari, K., Azadi, H., Mamoorian, M. (2014). Analyzing the impacts of agricultural land use change according to the expert's opinion of agricultural land organization in Iran. *Iranian Journal of Agricultural Economics and Development Research*, 45(4), 639–650.
- Baturin, N. A., Vuchetich, E. V., Kostromina, S. N., Kukarkin, B. A., Kupriyanov, E. A., Lurye, E. V., Mitina, O. V., Naumenko, A. S., Orel, E. A., Poletaeva, Yu. S., Popov, A. Yu., Potapkin, A. A., Simonenko, S. I., Sinitsyna, Yu. D., Shmelev, A. G. (2015). Rossiyskiy standart testirovaniya personala (vremennaya versiya, sozdannaya dlya shirokogo obsuzhdeniya v 2015 godu) [Russian standard for personnel testing (a temporary version created for broad discussion in 2015)]. *Organizational psychology*, 5(2), 67–138. (in Russian)
- Bertua, C., Anderson, N., Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational psychology*, 78(3), 387–409.
- Bock, R. D., Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443–459.
- Bock, R. D., Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied psychological measurement*, 6(4), 431–444.
- Cao, D. S., Schüz, N., Xie, G. R., Lippke, S. (2013). Planning skills moderate the intention–planning cognitions–behaviour relation: a longitudinal study on physical activity in Chinese adolescents. *Research in Sports Medicine*, 21(1), 12–23.
- Collins, A, Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology*, 10(3), e1001293.
- Costa, P. T. (1996). Work and personality: Use of the NEO-PI-R in industrial/organisational psychology. *Applied Psychology*, 45(3), 225–241.
- Costa, P. T., McCrae, R. R. (2014). The NEO Inventories. In *Personality assessment* (229–260). Routledge.
- Crocker, L., Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.

- Eagly, A. H., Woo, W., Diekmann, A. B. (2012). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes, H. M. Trautner (Eds.), *The developmental social psychology of gender* (123–174). NY: Psychology Press.
- Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J., Sjöberg, A. (2013). *EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. Version 4.2*, 6.
- Evers, A., Muñiz, J., Hagemester, C., Høstmælingen, A., Lindley, P., Sjöberg, A., Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema*, 25(3), 283–291.
- Frederiksen, N., Saunders, D. R., Wand, B. (1957). The in-basket test. *Psychological Monographs*, 71(438).
- Gouda, H. (2020). Exploring the relevant skills needed for the digital age employees. *European Journal of Business and Management*, 12(2), 58–67.
- Gross, R. G., Grossman, M. (2010). Executive resources. *Continuum: Lifelong Learning in Neurology*, 16(4), 140–152.
- Haladyna, T. M., Downing, S. M. (1988). *Functional Distractors: Implications for Test-Item Writing and Test Design*.
- Harada, C. N., Love, M. C. N., Triebel, K. (2013). Normal cognitive aging. *Clinics in geriatric medicine*, 29(4), 737.
- Hayes-Roth, B., Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science*, 3, 245–262.
- Hsu, N., Badura, K. L., Newman, D. A., Speech, M. E. P. (2021). Gender, “masculinity”, and “femininity”: A meta-analytic review of gender differences in agency and communion. *Psychological Bulletin*, 147(10), 987.
- Joshi, A., Kale, S., Chandel, S., Pal, D. K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4), 396–403.
- Karat, J. (1982). A model of problem solving with incomplete constraint knowledge. *Cognitive Psychology*, 14, 538–559.
- Kerzner, H. (2019). *Using the Project Management Maturity Model: Strategic Planning for Project Management*. John Wiley & Sons.
- Kock, F., Berbekova, A., Assaf, A. G., Josiassen, A. (2024). Developing a scale is not enough: on the importance of nomological validity. *International Journal of Contemporary Hospitality Management*, 36(10), 3233–3244.
- Köhler, C., Robitzsch, A., Hartig, J. (2020). A bias-corrected RMSD item fit statistic: An evaluation and comparison to alternatives. *Journal of Educational and Behavioral Statistics*, 45(3), 251–273.
- Linacre, J. M. (2025). *A User's Guide to WINSTEPS® MINISTEP Rasch-Model Computer Programs. Program Manual 5.9.1*. URL: <https://www.winsteps.com/a/Winsteps-Manual.pdf> (дата обращения: 27.03.2025)
- Masters, G. N. (2016). Partial credit model. In *Handbook of item response theory* (109–126). Chapman and Hall/CRC.
- Matlock-Hetzel, S. (1997). *Basic Concepts in Item and Test Analysis*.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel psychology*, 60(3), 683–729.
- Morosanova, V. I., Kondratyuk, N. G. (2020). K voprosu o faktornoy strukture kratkoy versii oprosnika V. I. Morosanovoy «Stil' samoregulyatsii povedeniya — SSPM» [On the factor structure of the short version of V. I. Morosanova's questionnaire “Behavioral Self-Regulation Style — SSPM”]. In T. N. Banshchikova, E. A. Fomina, V. I. Morosanova (Eds.), *Personal and regulatory resources for*

- achieving educational and professional goals in the digital age (October 20-23, 2020)* (178–188). Znanie-M. (in Russian)
- Morris, R., Ward, G. (Eds.). (2004). *The cognitive psychology of planning*. NY: Psychology Press.
- Nassif, V. M. J., Hashimoto, M., Amaral, D. J. (2014). Entrepreneurs self-perception of planning skills: evidence from Brazilian entrepreneurs. *Revista Ibero Americana de Estratégia*, 13(4), 107–121.
- Owen, A. (1997). Cognitive planning in humans: Neuropsychological, neuro anatomical and neuropharmacological perspectives. *Prog Neurobiol.*, 53, 431–450.
- Papa-Gusho, L., Biçaku-Çekrezi, R. (2015). Factors that affect effective planning skills of the teacher in the classrooms. *Academic Journal of Interdisciplinary Studies*, 4(3), 560–564.
- Pellerin, R., Perrier, N. (2019). A review of methods, techniques and tools for project planning and control. *International Journal of Production Research*, 57(7), 2160–2178.
- Phillips, L. H., MacLeod, M. S., Kliegel, M. (2005). Adult aging and cognitive planning. *The cognitive psychology of planning*, 111–134.
- Powell, J. W., Pharris, L. J., Hardy, M. M. (2020). A Comparison of time management skills among accounting, business, and information systems students by age and gender. *Issues in Information Systems*, 21(3), 1–10.
- Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of experimental child psychology*, 108(3), 621–637.
- Project Management Institute (PMI). (n.d.). *Rukovodstvo k svodu znaniy po upravleniyu proektami (PMBOK® Guide)* [A guide to the project management body of knowledge (PMBOK® Guide)]. (in Russian)
- Project Management Institute. (2021, July). *A Guide to the Project Management Body of Knowledge (руководство PMBOK®) — Seventh Edition and The Standard for Project Management*. Project Management Institute.
- Reilly, D., Neumann, D. L., Andrews, G. (2022). Gender Differences in Self-Estimated Intelligence: Exploring the Male Hubris, Female Humility Problem. *Frontiers in Psychology*, 13.
- Roshanai, A., Janeslätt, G., Marie White, S., Lidstroem Holmqvist, K., Holmefur, M. (2021). Psychometric properties of the original version of the assessment of time management skills. *The American Journal of Occupational Therapy*, 75(Supplement_2), 7512500038p1-7512500038p1.
- Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Biological Sciences*, 298(1089), 199–209.
- Smart Business Solutions. (n. d.). Organizatsionnyi test. Polnaya i ekspress versiya [Organizational Test. Full and Express Version]. *HR-market*. URL: <https://sbsc.ru/hrm/orgtest.html> (in Russian)
- Sorel, O., Pennequin, V. (2008). Aging of the planning process: The role of executive functioning. *Brain and cognition*, 66(2), 196–201.
- Stroh, W. A. (2025). *Metody aktivnogo sotsial'no-psikhologicheskogo obucheniya : uchebnik i praktikum dlya akademicheskogo bakalavriata* [Methods of active social and psychological learning: A textbook and workshop for the academic bachelor's degree]. M.: Yurayt. (in Russian)
- Syzmanowicz, A., Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta-analyses. *Learning and Individual Differences*, 21(5), 493–504.
- Testa, S., Toscano, A., Rosato, R. (2018). Distractor efficiency in an item pool for a statistics classroom exam: Assessing its relation with item cognitive level classified according to Bloom's taxonomy. *Frontiers in psychology*, 9, 1585.
- The OPC (2025). Graduate Situational Judgement Test (GSJT). *opc assessment*. URL: <https://www.theopc.co.uk/assessment/test/gsjt/> (дата обращения: 28.10.2025)

- Thorndike, R. M., Thorndike-Christ, T. M. (2013). *Measurement and evaluation in psychology and education*. Pearson Education Limited
- UNESCO Institute of Statistics (2015). Researchers by age. *UNESCO UIS*. Retrieved from <https://uis.unesco.org/node/3080143>
- Verulava, T. (2023). Strategic management and planning for health care organizations in Georgia. *Asia Pacific Journal of Health Management, 18*(3), 41–48.
- Vilkova, K. A. (2020). Izmerenie samoreguliruemogo obucheniya: obzor instrumentov [Measuring self-regulated learning: A review of instruments]. *Sovremennaya zarubezhnaya psikhologiya, 9*(2), 123–133. (in Russian)
- Viswesvaran, C., Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and psychological measurement, 59*(2), 197–210.
- Vita, S., Coplin, H., Feiereisel, K. B., Garten, S., Mechaber, A. J., Estrada, C. (2013). Decreasing the ceiling effect in assessing meeting quality at an academic professional meeting. *Teaching and learning in medicine, 25*(1), 47–54.
- Wakita, T., Ueshima, N., Noguchi, H. (2012). Psychological distance between categories in the Likert scale: Comparing different numbers of options. *Educational and psychological measurement, 72*(4), 533–546.
- Ward, L. B. (1959). The business in-basket test: A method of assessing certain administrative skills. *ETS Research Bulletin Series, 1959*(1), i–49.
- Yasin, M. I., Bazhdanova, Yu. V. (2024). Metodika izmereniya dolgosrochnosti i kratkosrochnosti motivatsii [Methodology for measuring long-term and short-term motivation]. *Organizational psychology, 14*(1), 96–108. (in Russian)
- Zelazo, P. D., Cunningham, W. (2007). Executive function: Mechanisms underlying emotion regulation. In J. J. Gross and B. Q. Ford (Eds.), *Handbook of emotion regulation* (135–158). Guilford Press.

Received 31.03.2025