

Федеральное государственное бюджетное научное учреждение  
«Институт стратегии развития образования»

*На правах рукописи*

МАЛЫГИН Алексей Александрович

**МЕТОДОЛОГИЯ И ТЕХНОЛОГИЯ  
МНОГОСТАДИЙНОГО ОЦЕНИВАНИЯ ВЫПУСКНИКОВ  
ПРИ ИТОГОВОЙ АТТЕСТАЦИИ В ВЫСШЕМ ОБРАЗОВАНИИ**

5.8.7. Методология и технология профессионального образования

Диссертация на соискание ученой степени  
доктора педагогических наук

Научные консультанты:  
академик РАО, доктор философских наук,  
кандидат педагогических наук, профессор  
Иванова Светлана Вениаминовна  
доктор педагогических наук, профессор  
Чельшкова Марина Борисовна

Москва – 2024

## Оглавление

<b>Введение .....</b>	<b>4</b>
<b>Глава 1. Обоснование многостадийного оценивания при итоговой аттестации в высшем образовании .....</b>	<b>27</b>
1.1. Становление и развитие подходов к оцениванию при аттестации в отечественном образовании.....	27
1.2. Особенности и целевые ориентиры современного оценивания выпускников при итоговой аттестации в высшем образовании.....	47
1.3. Понятийно-терминологический аппарат многостадийного оценивания при итоговой аттестации в высшем образовании .....	70
Выводы по первой главе.....	91
<b>Глава 2. Методология многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании .....</b>	<b>94</b>
2.1. Цели и функции многостадийного оценивания при итоговой аттестации...	94
2.2. Принципы многостадийного оценивания.....	101
2.3. Бипарадигмальный подход и теория тестов как методологические основания многостадийного оценивания уровня компетентности выпускников при итоговой аттестации .....	107
2.4. Методы и инструментарий многостадийного оценивания уровня компетентности выпускников при итоговой аттестации.....	137
2.5. Структурная модель реализации многостадийного оценивания уровня компетентности выпускников при итоговой аттестации.....	147
Выводы по второй главе.....	154
<b>Глава 3. Технология многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании .....</b>	<b>157</b>
3.1. Разработка инструментария для многостадийного оценивания уровня компетентности выпускников.....	157
3.2. Стратегии и алгоритмы реализации многостадийного оценивания.....	188
3.3. Методика определения надежности результатов многостадийного оценивания .....	205

3.4. Обеспечение содержательной валидности результатов многостадийного оценивания.....	211
Выводы по третьей главе.....	216
<b>Глава 4. Опытнo-экспериментальная работа по реализации технологии многостадийного оценивания выпускников.....</b>	<b>219</b>
4.1. Организационно-педагогические условия реализации многостадийного оценивания выпускников.....	219
4.2. Шкалирование результатов и установление критериальных баллов для многостадийного оценивания.....	233
4.3. Классификационные решения об аттестации выпускников на основе многостадийного оценивания.....	248
Выводы по четвертой главе.....	259
<b>Заключение.....</b>	<b>261</b>
<b>Список литературы.....</b>	<b>270</b>
<b>Приложение 1. Виды моделей современной теории тестов (IRT).....</b>	<b>308</b>
<b>Приложение 2. Макет паспорта компетенции.....</b>	<b>310</b>
<b>Приложение 3. Различные таксономии для операционализации конструкторов.....</b>	<b>316</b>
<b>Приложение 4. Пример структуры спецификации измерителя для аттестационного испытания.....</b>	<b>320</b>
<b>Приложение 5. Требования к заданиям разного типа.....</b>	<b>322</b>
<b>Приложение 6. Конструктор заданий.....</b>	<b>327</b>
<b>Приложение 7. Примеры оценочных заданий для различных стадий оценивания (направление подготовки 44.03.02 Психолого-педагогическое образование, образовательная программа «Психология образования»).....</b>	<b>328</b>
<b>Приложение 8. Расчеты показателей надежности по различным методикам.....</b>	<b>338</b>
<b>Приложение 9. Практические рекомендации по математико-статистической обработке и интерпретации результатов тестирования.....</b>	<b>339</b>
<b>Приложение 10. Анкета для выпускника.....</b>	<b>351</b>
<b>Приложение 11. Анкета для работодателя.....</b>	<b>355</b>
<b>Приложение 12. Основные характеристики компетенции и индикаторы достижения компетенции.....</b>	<b>358</b>
<b>Приложение 13. Отчет по результатам социологического опроса (анкетирования).....</b>	<b>362</b>
<b>Приложение 14. Расчеты вероятностей правильных ответов для 17 заданий по однопараметрической модели IRT при различных значениях уровня подготовленности.....</b>	<b>369</b>
<b>Приложение 15. Протоколы государственной экзаменационной комиссии.....</b>	<b>371</b>

## ВВЕДЕНИЕ

**Актуальность темы исследования.** Со второго десятилетия XXI века в нашей стране на законодательном уровне реализуется компетентностный подход к подготовке специалистов, который лежит в основе федеральных государственных образовательных стандартов, имеющих уже не одну версию обновлений, по направлениям подготовки и специальностям всех уровней высшего образования. Эти инновации привели к ряду изменений в структуре и содержании образования, однако практически не сказались на контрольно-оценочной деятельности в организациях высшего образования России. Фрагментарные попытки исследовательских коллективов, учебно-методических объединений, университетов и отдельных ученых и практиков актуализировать и обновлять фонды оценочных средств и технологии их применения, к сожалению, не привели к существенным изменениям, столь необходимым в аттестационных испытаниях высокой значимости. Одним из таких аттестационных испытаний является государственный экзамен в процедуре итоговой аттестации выпускников, подводящей черту их многолетнего обучения и выявляющей, по своему назначению, уровень сформированности установленной стандартом и образовательной программой компетентности.

В соответствии с положениями Федерального закона № 273-ФЗ от 29.12.2012 г. «Об образовании в Российской Федерации» итоговая аттестация проводится в целях определения соответствия результатов освоения обучающимися основных образовательных программ требованиям федерального государственного образовательного стандарта. Таким образом, согласно Закону при итоговой аттестации необходимо оценить компетентность выпускников, которая складывается из совокупности универсальных, общепрофессиональных и профессиональных компетенций.

Традиционные формы и методы проведения итоговой аттестации приходят в явное противоречие не только с требованиями Закона «Об образовании в Российской Федерации», но и с многочисленными попытками вузов расширить

перечень профессиональных компетенций, овладение которыми декларируется в образовательных программах, но уровень сформированности которых никак не подтверждается при итоговой аттестации выпускников. Число профессиональных компетенций во многих образовательных программах намного больше двадцати, и ряд университетов продолжает увеличивать их количество, вводя дополнительные компетенции.

Стагнация контрольно-оценочной деятельности в высшем образовании в какой-то мере имеет исторические предпосылки, поскольку профессиональное педагогическое сообщество не относилось должным образом к необходимости формирования определенной культуры оценивания и обращения к образовательным измерениям. Однако не только требования компетентностного подхода к трактовке качества результатов обучения, но и постоянное увеличение объемов информации в образовании, бурное развитие технологий на основе машинного обучения приводят к необходимости изменений в итоговой аттестации выпускников, проведения аттестации на основе новой методологии, новых форм организации контрольно-оценочной деятельности в образовательных организациях высшего образования и нового инструментария для ее осуществления. В основе таких изменений должно лежать многостадийное оценивание, позволяющее в несколько этапов образовательных измерений с применением различного инструментария говорить о надежных и валидных результатах аттестационных процедур. Только такие существенные изменения позволят получить достоверные и обоснованные выводы об уровне сформированности компетентности выпускников, дополненные оценками вторичного анализа результатов итоговой аттестации для улучшения содержания образовательных программ, уточнения формулировок компетенций с целью их корреляции с трудовыми функциями, изложенными в профессиональных стандартах, что весьма актуально для современной образовательной практики и будет способствовать в конечном счете повышению качества подготовки выпускников вузов и их востребованности на рынках труда.

В теоретическом плане актуальность рассматриваемых в диссертации вопросов определяется недостаточной научной разработанностью методологических и технологических основ многостадийного оценивания на основе образовательных измерений.

**Степень разработанности темы исследования.** К настоящему времени в педагогической науке получены определенные как фундаментальные, так и прикладные результаты в области образовательного оценивания для дальнейшего развития отечественной системы образования. Так, научные аспекты проблемы контрольно-оценочной деятельности рассматриваются в контексте управления качеством образования на разных его уровнях (А. Е. Бахмутский, А. Г. Бермус, В. А. Болотов, Р. Е. Булат, О. А. Граничина, В. И. Звонников, Л. В. Ишкова, Е. Ю. Левина, М. Л. Левицкий, В. П. Панасюк, С. А. Сафонцев, В. В. Сериков, С. Ю. Трапицын и др.). К этому же направлению исследований следует отнести концептуально-методологические работы, в которых изучаются и обосновываются различные аспекты мониторинга и оценки качества образования (В. И. Андреев, Е. В. Бебенина, М. В. Гуськова, М. Ю. Демидова, Н. Ф. Ефремова, В. И. Звонников, С. В. Иванова, Г. С. Ковалева, М. И. Кузнецова, А. Н. Майоров, О. А. Решетникова, Т. В. Третьякова, Ю. А. Шихов и др.).

Построение содержания профессиональной подготовки специалистов на компетентностной основе и разработка подходов и моделей доказательной аргументации оценки заданных стандартами результатов представлены в исследованиях С. Н. Беловой, В. А. Болотова, О. В. Галустян, В. В. Гриншкун, Н. Ф. Ефремовой, С. А. Писаревой, Н. В. Чекалевой, М. Б. Чельшковой и др.

Формирование и развитие российской национальной системы квалификаций и закрепление на законодательном уровне системы независимой оценки квалификаций в соответствии с требованиями профессиональных стандартов и рынков труда потребовало своего теоретико-методологического обоснования (В. И. Блинов, Е. Ю. Есенина, С. А. Ефимова, Т. В. Семенова, Ж. М. Сизова, А. А. Факторович и др.).

Образовательные измерения, получившие интенсивное развитие с начала XXI века, дополняют и обогащают проводимые исследования в области психологической и педагогической диагностики и психометрики, обосновывают необходимость применения обширного математического аппарата в аттестационных процедурах различного уровня значимости, что актуализировано и представлено в трудах Н. А. Батурина, Л. Ф. Бурлачука, М. Ю. Демидовой, К. Д. Дятловой, И. Н. Елисеева, Е. Ю. Кардановой, Г. С. Ковалевой, Т. Н. Малаховой, Е. А. Михайлычева, Н. Н. Найденовой, О. А. Решетниковой, А. Г. Шмелева, М. Б. Чельшковой и др.

Адаптивный подход к освоению содержания образования и оцениванию результатов обучения, учитывающий специфику предметной области, личностные особенности обучающихся и условия цифровой трансформации образовательного процесса, рассматривается в работах Д. Ф. Аббакумова, Ю. В. Вайнштейн, В. В. Гриншкуна, В. Т. Горбачева, Л. С. Куравского, Н. Т. Минко, Ю. Б. Поповой, Р. В. Терюхи, М. В. Хлопотова и др.

Малоизученное перспективное направление многостадийного оценивания в части сочетания различного оценочного инструментария для повышения надежности и валидности результатов затрагивается в ряде исследований по аттестации и аккредитации В. И. Звонникова, Т. Н. Малаховой, Т. В. Семеновой, Ж. М. Сизовой, М. Б. Чельшковой и др.

Зарубежные исследования, посвященные вопросам оценивания качества результатов обучения студентов и сформированности компетенций выпускников, представлены в работах, которые опубликованы уже в XXI веке (S. Blömeke, J. Garden, P. Griffin, B. Erdwien, B. Haasler, L. Heinemann, J. J. Fremer, C. Kuhn, T. Martens, A. Maurer, C. N. Mills, R. J. Mislevy, M. T. Potenza, F. Rauner, W. C. Ward, O. Zlatkin-Troitschankaia и др.).

Отдельно следует выделить работы, сфокусированные на развитии и применении классической теории тестов (J. Algina, L. Crocker, J. Keeves, R. Linn, F. Lord, M. Novick, M. H. Stone, B. D. Wright и др.), на характеристике математических моделей современной теории тестов (D. Andrich, D. L. Bandalos,

К. А. Becker, В. А. Bergstrom, А. Birnbaum, R. D. Bock, R. D. Gibbons, F. Lord, Т. М. Haladyna, R. K. Hambleton, M. Novick, H. J. Rogers, H. Swaminathan и др.), адаптивного тестирования (С. А. Glas, А. А. Davier, Т. J. Eggen, G. Kingsbury, W. J. Linden, R. J. Owen, N. A. Thompson, H. Wainer, D. Weiss, A. Zara, D. Yan и др.), шкалирования результатов оценочных процедур и установления стандартов выполнения тестов (W. A. Angoff, R. L. Brennan, G. J. Cizek, M. J. Kolen и др.).

Особое развитие в первые десятилетия XXI века получил доказательный подход в сфере социальных и гуманитарных наук, в том числе – в современных обучающих и оценочных технологиях, проводимый на основе аппарата образовательных измерений (С. М. Авдеева, И. В. Брун, Н. П. Бусыгина, Н. Ф. Ефремова, Т. Н. Малахова, Е. А. Орел, И. Л. Углова, М. Б. Чельшкова, S. Messick, R. Mislevy и др.).

Вместе с тем, несмотря на значительное число теоретических и практических работ, посвященных совершенствованию контрольно-оценочной деятельности в контексте обновления содержания и форм профессионального образования, сочетанию дистанционных образовательных технологий с очной формой обучения, стандартизации образования на различных уровнях, существует определенная проблемная ситуация, обусловленная сложившимися противоречиями между практическими потребностями оценивания результатов освоения студентами образовательных программ высшего образования и уровня сформированности компетентности выпускников на итоговой аттестации и реальными возможностями используемых технологий и инструментария. На обобщенном уровне это можно обозначить как **противоречия** между:

– ролью оценки и значением оценивания результатов освоения обучающимися основных профессиональных образовательных программ и неразработанностью теоретико-методологических основ итоговой аттестации выпускников в высшем образовании, нацеленной на соблюдение принципов объективности и независимости при ее реализации;



– наличием исследований в области итоговой аттестации в прикладных целях и отсутствием общего методологического подхода к организации и проведению итоговой аттестации в высшем образовании;

– теоретической разработанностью бипарадигмального подхода к образовательным измерениям при аттестационных процедурах и отсутствием технологии и доказательных практик его результативности при итоговой аттестации выпускников в высшем образовании;

– наличием множества различных фондов оценочных средств для проведения итоговой аттестации в высшем образовании и их несоответствия задачам оценивания уровня сформированности компетентности выпускников, обеспечивающим получение надежных, валидных и аутентичных результатов;

– широким внедрением дистанционных образовательных технологий и происходящей цифровой трансформацией образовательного процесса и слабым применением современных технологий образовательных измерений, повышающих эффективность контрольно-оценочных процедур и сопоставимость результатов в условиях итоговой аттестации выпускников.

**Проблема исследования** заключается в неразработанности методологии многостадийного оценивания и технологии его применения при итоговой аттестации выпускников в высшем образовании, что обеспечило бы получение надежных, валидных и аутентичных результатов (т. е. оценок) для обоснованных классификационных решений об уровне сформированности их компетентности.

Актуальность, противоречия и проблема послужили основанием для определения **темы диссертационного исследования «Методология и технология многостадийного оценивания выпускников при итоговой аттестации в высшем образовании».**

**Объект исследования** – итоговая аттестация в высшем образовании.

**Предмет исследования** – многостадийное оценивание уровня компетентности выпускников при итоговой аттестации в высшем образовании.

**Цель исследования** – разработать и представить методологию многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании и технологию его реализации.

В соответствии с целью исследования ставятся следующие **задачи**:

1. Обосновать многостадийное оценивание для установления уровня компетентности выпускников при итоговой аттестации в высшем образовании.

2. Определить состав понятийно-терминологического аппарата многостадийного оценивания, дополнить и уточнить ряд терминов и категорий.

3. Представить авторскую методологию многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании.

4. Разработать и описать структурную модель многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании.

5. Осуществить разработку технологии многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании.

6. Выявить организационно-педагогические условия реализации методологии многостадийного оценивания уровня сформированности компетентности выпускников при итоговой аттестации.

7. Определить процедуру установления критериальных (пороговых) баллов, с помощью которых проводится классификация выпускников на аттестованных и не аттестованных для каждой стадии оценивания уровня их компетентности при итоговой аттестации.

8. Выполнить опытно-экспериментальную работу по апробации созданной методологии, применению модели по разработанной технологии многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании.

**Гипотеза исследования.** Для достижения поставленной цели в качестве гипотезы исследования выдвинуто предположение о том, что процедура оценивания при итоговой аттестации в высшем образовании приведет к

обоснованным классификационным решениям об уровне сформированности компетентности выпускников, если будет разработана и обоснована методология многостадийного оценивания выпускников при итоговой аттестации в высшем образовании:

которая будет включать цели, функции, принципы, основания, методы и средства;

основанием для методологии многостадийного оценивания будут выступать бипарадигмальный подход к образовательным измерениям и теория тестов, которые соответствуют современным требованиям к содержанию и организации итоговой аттестации в высшем образовании в условиях реализации компетентностного подхода;

для практической реализации будет создана и описана структурная модель многостадийного оценивания уровня компетентности выпускников;

технология многостадийного оценивания будет представлять собой взаимосвязанную и взаимообусловленную последовательность этапов образовательных измерений, инструментарий для которых на каждом этапе релевантен оцениваемому уровню сформированности компетентности аттестуемого;

методики, стратегии и алгоритмы многостадийного оценивания будут учитывать трудность и дифференцирующую способность каждого задания, входящего в оценочный инструментарий, обеспечивая тем самым получение надежных и валидных оценок (результатов);

будут выполняться организационно-педагогические условия реализации многостадийного оценивания уровня компетентности выпускников при итоговой аттестации;

принятие классификационных решений на аттестованных и не аттестованных по результатам многостадийного оценивания уровня компетентности выпускников будет основываться на критериально-ориентированном подходе в образовательных измерениях и с учетом требований образовательных и профессиональных стандартов одновременно.

Гипотеза может быть подтверждена апробацией разработанной методологии многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании на основе реализации предлагаемой технологии, совокупности методик и соблюдения организационно-педагогических условий.

**Научная новизна исследования:**

1. Обосновано многостадийное оценивание уровня компетентности выпускников при итоговой аттестации в высшем образовании, которая относится к процедурам высокой значимости и по результатам которой принимаются ответственные управленческие решения о присвоении квалификации выпускникам, делаются выводы о качестве подготовки специалистов по основным профессиональным образовательным программам.

2. Определен состав понятийно-терминологического аппарата многостадийного оценивания, дополнены и уточнены термины (оценочный инструментарий, включающий само оценочное средство и шкалу, на которую переносится результат измерения с использованием этого средства; тест как измеритель для аттестационных процедур высокой значимости, разрабатываемый в соответствии с требованиями теории; шкала, применяемая для многостадийного оценивания и отражающая уровни количественных и качественных измерений; критериальный (пороговый) балл, с помощью которого проводится классификация аттестуемых выпускников) и категории (диагностика, образовательные измерения, уровни сформированности компетентности, квалиметрия).

3. Представлена авторская методология многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании, включающая цели, функции, общие и специальные принципы, основания, методы и инструментарий.

4. Для реализации авторской методологии разработана структурная модель многостадийного оценивания уровня компетентности выпускников при итоговой аттестации, включающая методологический, содержательный, технологический компоненты и компонент принятия управленческих решений, обоснование применения и состав каждого из компонентов.

5. Осуществлена разработка технологии многостадийного оценивания уровня компетентности выпускников с позиций ее адекватности современным задачам итоговой аттестации при компетентностной трактовке результатов освоения образовательных программ. Показано, что данная технология в совокупности применяемых методик позволяет повысить эффективность оценочного процесса, оптимизировать процедуры генерации оценочных заданий по количеству и трудности, темпу их выполнения аттестуемыми выпускниками на разных этапах оценивания уровня сформированности компетентности без потери точности самих результатов.

6. Выявлены и представлены организационно-педагогические условия реализации методологии многостадийного оценивания уровня сформированности компетентности выпускников при итоговой аттестации.

7. Определена процедура установления критериальных (пороговых) баллов и классификации на ее основе выпускников на аттестованных и не аттестованных для каждой стадии оценивания.

8. Выполнена опытно-экспериментальная работа по апробации созданной методологии, применению модели по специально разработанной технологии многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании.

#### **Теоретическая значимость результатов исследования:**

Внесен научный вклад в теоретическую педагогику в части методологических основ контрольно-оценочной деятельности в высшем образовании, уточнен понятийно-терминологический аппарат в сфере образовательных измерений. Введен в научный оборот термин «многостадийное оценивание» и уточнены производные от него понятия, связанные с уровнем сформированности компетентности и ошибкой классификации аттестуемого.

Расширены научные представления об организации и проведении итоговой аттестации выпускников в условиях реализации компетентностного подхода в высшем образовании и о необходимости определения уровня сформированности компетентности на основе авторской методологии многостадийного оценивания.

Обосновано использование бипарадигмального подхода, предусматривающего сочетание количественных и качественных уровней образовательных измерений при оценивании компетентности, и теории тестов (классической и современной) для обеспечения надежных, валидных и аутентичных результатов итоговой аттестации в высшем образовании.

Предложены научные требования к проектированию и разработке оценочного инструментария, базирующиеся на образовательных измерениях и способствующие повышению качества результатов итоговой аттестации выпускников в высшем образовании.

### **Практическая значимость исследования:**

Материалы исследования используются в образовательном процессе. Диссертантом разработаны и ведутся учебные занятия по дисциплинам и модулям для университетских образовательных программ высшего образования *уровня бакалавриата*: «Оценивание в образовании», «Образовательные стандарты и программы», «Современные средства оценки учебных достижений»; *уровня магистратуры*: «Измерения в образовании», «Нормативно-правовое обеспечение психолого-педагогической деятельности»; разработаны программы *дополнительного профессионального образования* профессиональной переподготовки («Управление образованием») и повышения квалификации специалистов, обеспечивающих реализацию мероприятий по развитию национально-региональной системы независимой оценки качества образования и создание национальных механизмов оценки качества («Независимая оценка качества образования: порядок проведения, анализ результатов и принятие управленческих решений», «Моделирование системы оценки качества образования с учетом результатов независимых мониторинговых исследований»).

Разработан и апробирован аттестационный инструментарий для многостадийного оценивания уровня сформированности компетентности выпускников по нескольким направлениям подготовки высшего образования уровня бакалавриат и специалитет (44.03.02 Психолого-педагогическое образование, 39.03.01 Социология, 31.05.01 Лечебное дело).

Предложены и реализованы стратегии и алгоритмы процедур многостадийного оценивания при итоговой аттестации выпускников, обеспечивающие дифференциацию и индивидуализацию аттестуемых за счет предъявления им заданий, которые соответствуют уровню подготовленности каждого из них.

Определены организационно-педагогические условия, соблюдение которых обеспечивает принятие обоснованных классификационных решений об уровне компетентности выпускников на каждом этапе многостадийного оценивания при итоговой аттестации в высшем образовании.

Результаты работы получены также в рамках НИР «Разработка модели и методического обеспечения оценки результатов обучения студентов на основе адаптивных технологий педагогических измерений», выполнявшейся под руководством автора в 2013–2014 гг. при поддержке гранта Президента Российской Федерации; внедрены в практику выполнения проекта «Разработка и апробация модели и технологии создания и применения измерительных средств в соответствии с требованиями профессиональных стандартов для процедур оценки и сертификации квалификаций» ФЦПРО на 2011–2015 годы; в деятельность учебно-методических объединений по образованию в области менеджмента и по профессиональному образованию; в деятельность Ивановского государственного университета и Методического центра аккредитации специалистов здравоохранения. В целом материалы диссертации имеют практическую значимость для совершенствования содержания и условий реализации основных профессиональных образовательных программ, повышения качества высшего образования, проведения сравнительных исследований в области качества подготовки специалистов по одинаковым или смежным направлениям подготовки, функционирования внешней и внутренней систем оценки качества образования, развития аккредитационного мониторинга.

**Основные этапы исследования.** Работа выполнялась на протяжении 2012–2024 гг.

1 этап (2012–2019 гг.) – подготовительный и аналитико-поисковый. Данный этап включал в себя анализ практических проблем, выявленных при реализации нового поколения компетентностно-ориентированных образовательных стандартов и при оценивании результатов обучения в этих условиях (обучение профессорско-преподавательского состава; формирование фондов оценочных средств; анализ литературных источников; постановка и обоснование исследовательской проблемы; определение перечня задач исследовательской работы; моделирование объектов исследования). Предпосылкой аналитической работы послужил опыт профессиональной деятельности по развитию внутривузовской системы сертификации педагогических измерительных материалов, формированию банка калиброванных заданий и фондов оценочных средств по дисциплинам математического цикла, проектированию оценочных процедур и технологий в различных условиях. Этот опыт был получен автором диссертации в Ивановском государственном химико-технологическом университете, в Ивановском государственном университете, в Институте качества высшего образования НИТУ МИСиС на кафедре педагогических измерений; в учебно-методическом объединении по образованию в области менеджмента на базе Государственного университета управления; в учебно-методическом объединении по профессиональному образованию при Российском государственном профессионально-педагогическом университете; Ивановском государственном медицинском университете; Методическом центре аккредитации специалистов здравоохранения, а также за время деятельности автора в качестве руководителя и исполнителя научно-исследовательских работ.

2 этап (2020–2022 гг.) – аналитико-синтетический, нацеленный на обобщение практического опыта, анализ литературы по теме исследования, разработку концептуальной и методологической части результатов исследования, проведение анализа полученных результатов, подготовку статей, учебных пособий, монографии и глав в монографиях, апробацию результатов исследования на конференциях и совещаниях.



3 этап (2023–2024 гг.) – заключительный, состоящий в обобщении данных анализа, формулировании выводов, подведении итогов исследования, оформлении диссертационного исследования, разработке методик и рекомендаций по использованию результатов многостадийного оценивания при итоговой аттестации выпускников в высшем образовании.

**Опытно-экспериментальной базой исследования** выступали Ивановский государственный университет, Методический центр аккредитации специалистов здравоохранения, Ивановская государственная медицинская академия (с 2024 года – Ивановский государственный медицинский университет). Всего в исследовании приняло участие 1724 обучающихся, в том числе выпускников образовательных программ высшего образования, 250 человек из числа педагогических работников и 485 человек из числа профессорско-преподавательского состава, 75 представителей работодателей.

**Методологическую и теоретическую основу исследования** составили:

– методология педагогического исследования (В. И. Загвязинский, С. В. Иванова, В. В. Краевский, А. М. Новиков, В. М. Полонский и др.);

– методология компетентностного подхода (В. И. Байденко, А. Г. Бермус, В. А. Болотов, В. И. Звонников, И. А. Зимняя, В. В. Сериков, Ю. Г. Татур, А. В. Хуторской, P. Griffin, L. Spenser и др.);

– теоретические основы оценки качества образования и его результатов (А. Е. Бахмутский, В. А. Болотов, Н. Ф. Ефремова, В. И. Звонников, Г. С. Ковалева, А. Н. Майоров, В. П. Панасюк, И. Д. Рудинский, Н. А. Селезнева, А. И. Субетто, Т. В. Третьякова и др.);

– психолого-педагогические теории, раскрывающие принципы и методы диагностики индивидуальных склонностей, способностей и возможностей обучающихся (Б. Г. Ананьев, Н. А. Батулин, Л. Ф. Бурлачук, Л. С. Выготский, В. Н. Дружинин, К. Ингенкамп, А. Г. Шмелев и др.);

– теория образовательных измерений (М. В. Гуськова, Н. Ф. Ефремова, В. И. Звонников, В. М. Кадневский, А. А. Маслак, В. И. Михеев, Ю. М. Нейман,

М. Б. Челышкова, J. Algina, D. Bandalos, R. Berk, A. Binet, R. Brennan, L. Cook, L. Crocker, F. Galton, J. Keeves, R. Linn и др.);

– теория тестов, в которую входят *классическая теория тестов*, обеспечивающая концептуальные основы надежности и валидности результатов образовательных измерений, и *современная теория тестов*, устанавливающая функциональные взаимосвязи между наблюдаемыми результатами оценивания и латентными характеристиками обучающихся (А. Анастаси, Н. Ф. Ефремова, В. И. Звонников, Е. Ю. Карданова, А. Н. Майоров, М. Б. Челышкова, J. Algina, D. Andrich, A. Birnbaum, R. Bock, L. Crocker, R. Fischer, R. Hambleton, F. Lord, G. Rasch, S. Reise, H. Swaminathan и др.);

– теория адаптивного тестирования и многостадийности оценивания (Л. С. Куравский, Р. В. Терюха, М. Б. Челышкова, С. А. Glas, А. А. Davier, T. J. Eggen, G. Kingsbury, W. J. Linden, D. Magis, R. J. Owen, N. A. Thompson, H. Wainer, D. Weiss, D. Yan, A. Zara и др.).

При выполнении работы использовались следующие **методы исследования**:

– теоретические: анализ, классификация, сравнение, индукция и дедукция, абстрагирование, доказательство, систематизация и моделирование;

– эмпирические: констатирующий эксперимент, включенное наблюдение, беседы, анкетирование, тестирование, математическое моделирование на основе вероятностных моделей IRT, математико-статистические методы обработки результатов данных.

На защиту выносятся следующие **основные положения**:

1. Применение многостадийного оценивания для определения уровня компетентности выпускников при итоговой аттестации в высшем образовании в форме государственного экзамена как одного из двух аттестационных испытаний обосновывается тем, что: многостадийное оценивание представляет собой определенную последовательность взаимосвязанных и взаимообусловленных, научно обоснованных этапов и процедур создания оценочного инструментария, его применения для оценки результатов выполнения заданий; такое оценивание, проводимое в несколько этапов (стадий), отвечает требованиям компетентностного

подхода, на котором основываются ФГОС высшего образования, а также профессиональных стандартов, в которых трудовые функции специалистов описаны через необходимые знания, умения и трудовые действия; многостадийность при аттестационной процедуре высокой значимости способствует повышению надежности, валидности и аутентичности результатов аттестуемых выпускников, обеспечивая тем самым обоснованные классификационные решения об уровне сформированности их компетентности.

2. Определение состава понятийно-терминологического аппарата многостадийного оценивания потребовалось в силу новизны самого понятия, а также довольно размытого понятийно-терминологического аппарата контрольно-оценочной деятельности, аттестации и измерительной функции в образовании. Такая работа была проведена для данного исследования и может быть полезной для иных исследований в области образовательного оценивания и измерений. Было сформулировано понятие «многостадийное оценивание», а также уточнены и дополнены в исследовательском контексте производные термины (оценочный инструментарий, тест, шкала, критериальный (пороговый) балл) и категории (диагностика, образовательные измерения, уровни сформированности компетентности, квалиметрия).

3. Представлена авторская методология многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании, включающая *цели*, к числу которых относятся – повышение качества высшего образования и подготовки специалистов, сопоставимость оценок и уровня квалификации выпускников, обеспечение преемственности требований к качеству высшего образования на всех уровнях, снижение доли субъективности и несправедливости аттестационных процедур, повышения их эффективности, обоснованности аттестационных решений, обеспечение благоприятной атмосферы при прохождении аттестационного испытания, снижающей невербальное психологическое давление; *функции* – оценочная, информационная, сравнительная, прогностическая; *принципы* – нормативные (объективность, независимость), общедидактические (научность, релевантность, достоверность, справедливость,

эффективность) и специальные (дифференциация, индивидуализация, последовательность, критериальность, открытость, сочетание количественного и качественного уровней измерений, кумулятивность); *основания* – компетентностный подход, бипарадигмальный подход и теория тестов; *методы и инструментарий*, предусматривающие их применение сообразно этапу оценивания и уровню образовательных измерений.

4. Для реализации авторской методологии разработана структурная модель многостадийного оценивания уровня компетентности выпускников при итоговой аттестации, состоящая из методологического, содержательного, технологического и управленческого компонентов. Методологический компонент содержит цели, функции и принципы научной организации многостадийного оценивания с опорой на бипарадигмальный подход и теорию тестов (классическую и современную). Содержательный компонент включает в себя компетентностную трактовку к результатам освоения обучающимися образовательных программ и требования профессиональных стандартов, которые должны быть положены в основу выделения профессионально значимой совокупности компетенций. Технологический компонент содержит совокупность методик и алгоритмов реализации многостадийного оценивания, в числе которых проектирование и разработка аттестационного инструментария, обеспечение надежности и валидности результатов и их шкалирование на каждом этапе оценивания. Компонент принятия управленческих решений при итоговой аттестации предусматривает обращение к критериально-ориентированному подходу в образовательных измерениях, установление пороговых баллов (стандартов выполнения) для каждой стадии оценивания соответствующего уровня сформированности компетентности выпускника и обоснование аттестационного вывода о каждом аттестуемом (выставление оценки по пятибалльной шкале согласно нормативным требованиям).

5. Созданная технология на основе методологии многостадийного оценивания уровня компетентности выпускников, которая представляет собой совокупность средств, методик, алгоритмов и классификационных решений и

являет собой, по сути, технологический компонент модели. Основным средством выступает оценочный инструментарий для аттестационных процедур с высоким уровнем значимости, разработанный в соответствии с требованиями теории тестов, обладающий системообразующими статистическими характеристиками и обеспечивающий получение надежных и валидных результатов. В числе методик – совокупность правил и последовательность действий реализации многостадийного оценивания, создания оценочного инструментария, релевантного уровню компетентности, определения надежности и обеспечения содержательной валидности результатов.

6. Для реализации методологии многостадийного оценивания уровня сформированности компетентности выпускников при итоговой аттестации необходимо выполнение следующих организационно-педагогических условий:

- повышение уровня квалификации педагогических работников в области объективного оценивания и образовательных измерений;

- наличие банка калиброванных оценочных заданий, разработанных с учетом требований профессиональных стандартов;

- наличие компьютерных программ либо программно-инструментальной среды (сервиса, платформы), обеспечивающих реализацию алгоритмов многостадийного оценивания, а также получение обратной связи от выпускников и работодателей для развития и совершенствования качества подготовки специалистов и аттестационных процедур;

- наличие спецификации к оценочному инструментарию для аттестационных процедур, которая обеспечивает содержательную валидность результатов образовательных измерений;

- наличие согласованного регламента для проведения многостадийного оценивания при итоговой аттестации.

7. Определена процедура установления критериальных (пороговых) баллов, заранее доводимая до аттестуемых и обеспечивающая принятие обоснованного классификационного решения об уровне компетентности выпускников на каждой стадии оценивания. Установление критериальных (пороговых) баллов проводится

с помощью итеративного структурированного метода оценки заданий и с учетом минимизации ошибки классификации каждого выпускника при отнесении его к категории «аттестованный» или «не аттестованный».

8. Проведенная опытно-экспериментальная работа с участием 1724 обучающихся, в том числе выпускников по образовательным программам высшего образования, 250 педагогических работников, 485 человек из числа профессорско-преподавательского состава, 75 представителей работодателей показала успешные результаты реализации методологии на основе применения модели во всех ее компонентах, при особом внимании к технологическому компоненту, потому что основными целями опытно-экспериментальной работы были (1) апробация технологии, совокупности методик, инструментария для проведения всех стадий многостадийного оценивания и (2) обоснованность принятия решений по результатам аттестационных испытаний в рамках итоговой аттестации выпускников университета.

**Степень достоверности результатов исследования** обеспечена комплексным использованием современных теоретических и эмпирических методов, адекватных целям и задачам исследования; всесторонним анализом проблемы при определении исходных теоретико-методологических позиций; непротиворечивостью полученных выводов относительно основных педагогических идей и результатов проверки состоятельности предложенных положений и технологий организации и проведения многостадийного оценивания при итоговой аттестации выпускников в высшем образовании.

Достижению должного научного уровня диссертационного исследования и обоснованности его результатов способствовало использование системного подхода, соблюдение логики общенаучного подхода с опорой на методологию педагогической науки. Обоснованность технологических подходов, предложенных в диссертации к решению ряда поставленных в ней задач, обеспечивалась при помощи эмпирических проверок и внедрения в образовательную практику разработок, выполненных на основании результатов исследования. Достоверность выводов по результатам внедренческой работы подтверждалась проверкой

соответствия результатов многостадийного оценивания критериям надежности, валидности и аутентичности.

**Апробация результатов исследования.** Основные положения диссертации, теоретические и практические результаты представлялись и обсуждались на:

заседаниях ученого совета (2022, 2023 гг.), лаборатории дидактики общего и профессионального образования (2023, 2024 гг.) ФГБНУ «Институт стратегии развития образования»; заседаниях Отделения философии образования и теоретической педагогики Российской академии образования (2023, 2024 гг.);

заседаниях методического совета Ивановского государственного университета (2014–2024 гг.);

*международных мероприятиях:* III Конгрессе РОПРЯЛ «Динамика языковых и культурных процессов в современной России» (г. Санкт-Петербург, 2012 г.); школе-семинаре «Физика в системе высшего и среднего образования» (г. Москва, 2012 г.); XXXII Международном научном семинаре преподавателей математики университетов и педагогических вузов «Современные подходы к оценке и качеству математического образования в школе и вузе» (г. Екатеринбург, 2013 г.); Международных научно-методических конференциях «Современное образование: содержание, технологии, качество» (г. Санкт-Петербург, 2013, 2014 гг.); Международных научно-методических конференциях «Проблемы высшего образования» (г. Хабаровск, 2014, 2015, 2017, 2018 гг.); V Международной конференции ЕАОКО «Управление качеством образования на основе оценки» (Киргизия, п. Кара-Ой, 2016 г.); XIV Международной научно-практической конференции «Тенденции развития образования: кто и как использует и оценивает образовательные стандарты» (г. Москва, 2017 г.); 22-й Международной научно-практической конференции «Инновации в профессиональном и профессионально-педагогическом образовании» (г. Екатеринбург, 2017 г.); IV Международной научно-практической конференции «Электронное обучение в непрерывном образовании» (г. Ульяновск, 2017 г.); Международной научно-практической конференции «Глобальное европейское пространство: проблемы интеграции сквозь призму межкультурной

коммуникации» (г. Ростов-на-Дону, 2017 г.); V Международной научно-практической конференции «Научно-методические подходы к формированию образовательных программ подготовки кадров в современных условиях» (г. Москва, 2017 г.); Международной конференции Евразийской ассоциации оценки качества образования «Мониторинги, рейтинги, рэнкинги как инструменты управления качеством образования» (г. Санкт-Петербург, 2017 г.); XVI Международной научно-практической конференции «Тенденции развития образования. Эффективность образовательных институтов» (г. Москва, 2019 г.); Международных научно-практических конференциях «Образовательное пространство в информационную эпоху» (ЕЕИА-2019, ЕЕИА-2023) (г. Москва); II Международной научно-практической конференции «Непрерывное образование: теория и практика реализации» (г. Екатеринбург, 2019 г.); Международной научно-практической конференции «Университет, открытый регионам: интеграционные механизмы взаимодействия педагогического вуза и регионов» (г. Санкт-Петербург, 2022 г.); Международной научной конференции «Научное наследие великого русского педагога К. Д. Ушинского» (г. Москва 2023 г.); III Международной научно-практической конференции «Инновационное развитие образования: методы и регуляторы» (г. Москва, 2023 г.); IX Международном форуме по педагогическому образованию «Качество педагогического образования в условиях современных вызовов» (г. Казань, 2023 г.); II и III Международной Ассамблеи Российской академии образования (2023, 2024 гг.);

*конференциях всероссийского уровня:* VI Всероссийской научно-практической конференции с международным участием «Тестирование в сфере образования: проблемы и перспективы развития» (г. Красноярск, 2013 г.); Всероссийской научно-методической конференции «Университетский комплекс как региональный центр образования, науки и культуры» (г. Оренбург, 2019 г.); XVII Пастуховских чтениях «ДПО как ключевой элемент кадрового обеспечения развития национальной экономики» (г. Ярославль, 2019 г.); XVI Форуме Гильдии экспертов в сфере профессионального образования «Новые технологии оценки качества образования» (г. Москва, 2020 г.); научно-методической конференции



«Современное университетское образование: вызовы и проблемы, ценности инновации, технологии и качество» (г. Иваново, 2021 г.); IV и V Всероссийском научно-образовательном Форуме с международным участием «Миссия университетского педагогического образования в XXI веке» (г. Ростов-на-Дону, 2022, 2023 гг.); Data-Driven Life Conference (г. Санкт-Петербург, г. Иваново, 2022 г.); II Всероссийской научно-практической конференции с международным участием «Современные тенденции развития общего и вузовского образования» (г. Ярославль, 2022 г.); научно-практической конференции Розовские чтения «Качество педагогического образования: состояние и проблемы» (г. Москва, 2023 г.); научной конференции «Традиции и инновации в классическом университете» (г. Иваново, 2023 г.); Втором профессорском педагогическом форуме «Университет в эпоху педагогических новаций» (г. Санкт-Петербург, 2023 г.); научно-практической конференции «Нормативные правовые механизмы регулирования электронного обучения и применения дистанционных технологий в высшем образовании» (г. Москва, 2023 г.); XVI Всероссийской научно-методической конференции с международным участием «Актуальные проблемы совершенствования высшего образования» (г. Ярославль, 2024 г.);

межвузовском исследовательском проекте «Новая дидактика» (г. Ярославль, 2023 г.) и просветительском проекте «ЛекторРАО» (г. Москва, 2023 г.).

Наиболее значимые научные результаты диссертационного исследования отражены в монографиях и главах монографий, научных статьях, учебных и учебно-методических пособиях. По теме исследования опубликована 61 работа, в том числе: 2 авторские монографии и 2 главы в коллективных монографиях, 20 статей в ведущих реферируемых журналах, включенных в список ВАК, 4 статьи в изданиях, индексируемых в международных базах цитирования.

### **Соответствие диссертации паспорту научной специальности.**

Проведенное исследование соответствует следующим пунктам паспорта научной специальности 5.8.7. Методология и технология профессионального образования, направления исследования:

п. 6. Педагогические измерения в профессиональном образовании;

п. 13. Проектирование содержания, методов, дидактических систем и технологий профессионального образования. Системы проектирования и оценивания результатов профессионального образования.

**Структура диссертации.** Диссертация состоит из введения, четырех глав, заключения, списка литературы и приложений.

# ГЛАВА 1. ОБОСНОВАНИЕ МНОГОСТАДИЙНОГО ОЦЕНИВАНИЯ ПРИ ИТОГОВОЙ АТТЕСТАЦИИ В ВЫСШЕМ ОБРАЗОВАНИИ

## 1.1. Становление и развитие подходов к оцениванию при аттестации в отечественном образовании

Одним из центральных вопросов, который касается любого человека и который никого не оставляет равнодушным в современном мире, был и остается вопрос, связанный с оценкой в образовании. Непосредственно для образовательного процесса роль оценки трудно переоценить, хотя, как будет показано ниже, в истории отечественной педагогики были этапы безотметочного обучения. Как несколько веков назад, так и сегодня, зачастую оценку считают основным «регулятором» образовательных отношений [24, 28, 84, 98 и др.]. В управлении образованием это направление обозначается как контрольно-надзорная и оценочная деятельность. На теоретико-методологическом уровне – это образовательное оценивание как процесс и оценка как результат в самых разнообразных контекстах – от выставления отметок до принятия стратегических решений, оказывающих влияние как на личность обучающегося, так и на повышение качества образования профессиональных кадров, и, следовательно, на экономику страны и ее дальнейшее развитие [3, 4, 23, 132 и др.].

Рассматривая процесс становления и развития подходов к организации и содержанию контрольно-оценочной деятельности в нашей стране, следует отметить, что она менялась и преобразовывалась по мере развития системы образования, общества, технологий, под влиянием различных политических, социальных и экономических факторов. Обращаясь к историческому контексту, к XVIII столетию, в частности, прежде всего следует выделить документ 1721 года, известный как Духовный регламент. В нем впервые была введена обязанность учителя – оценивать, «как кто учится», согласно Регламенту надо было указывать две характеристики: «лениво» и «прилежно». Позднее, в 1758 году,

М. В. Ломоносов в своем «Регламенте Академической гимназии» предлагал выставлять отметки буквенными аббревиатурами. Эти исторические факты говорят о зарождении педагогической терминологии и подходов к введению оценочной деятельности в образовании [28, 207, 220].

Разумеется, такое педагогическое явление, как оценка (оценивание), имеет свою предысторию и стало известно задолго до появления государственных документов в XVIII веке. Это явление рассматривается и историками педагогики, которые изучают закономерности развития образования и влияния на него оценочной деятельности, и исследователями в области оценочной деятельности, педагогических измерений, которые прокладывают путь инновационным оценочным методикам и доказывают эффективность тех или иных подходов, стремясь к получению объективной информации и принятия решений на ее основе [3, 71, 111, 221, 236, 245 и др.].

Обращаясь к XIX веку как веку становления научной педагогики, связанному прежде всего с именем К. Д. Ушинского, мы находим документальные подтверждения формирования единых взглядов и общих подходов к оценке. Так, в 1802 году по указу Императора Александра I было создано Министерство народного просвещения, которое в числе других задач вырабатывало единые подходы к оценке (и оцениванию) и фиксировало их в различных документах. В 1818 году Министерство выпустило распоряжение «Начертание подробнейших правил, касательно испытаний в таких учебных заведениях, коих воспитанники при выпуске имеют право на получение классного чина», в котором вводилась четырехбалльная система оценивания успехов учащихся, позднее замененная на 12-балльную.

В 1837 году Министром народного просвещения С. С. Уваровым было предложено шестибалльное оценивание. В обосновывающем этот шаг документе «Правила для желающих поступать в университеты» было записано: «для определения познаний употребляются цифры: 0, 1, 2, 3, 4 и 5, которые в том же порядке означают: совершенное незнание, слабые, посредственные, достаточные, хорошие, отличные сведения» [207, 213]. Исследователи полагают, что

«ожидаемого единообразия вследствие очередного нововведения Министерства не последовало, и учебные заведения продолжали самостоятельный поиск новых подходов к процедуре оценивания» [220, с. 196].

Далее в 1845 году в России применялась 115-балльная система оценивания, которая описана в «Положении об испытаниях в учебных заведениях Министерства народного просвещения». Она не получила широкого распространения в силу своей громоздкости и трудности перевода в 5-балльную шкалу.

Обращаясь к богатому наследию К. Д. Ушинского как следующей точке для рассмотрения изменений подходов и процедур в контрольно-оценочной деятельности в образовании, стоит выделить его статью «Несколько слов об училищных отметках (баллах), употребляемых в наше время во всей России», написанную в 1861 году, то есть более 160 лет назад [249]. В частности, одно из глубоких замечаний, высказанное от лица преподавателя гимназии Густава Александровича, не теряет своей актуальности и сегодня: «Отвлеченное значение, которое имеют цифры 1, 2, 3, 4, 5, хотя не показывает прямого влияния на учение, но составляет также важные неудобства для преподавания. Действительно, эти цифры показывают только, как тот или другой ученик ответил урок, но нисколько не знакомят нас с их способностями, с причиною их медленных успехов, с их хорошими или дурными качествами... Мы из этих отметок не можем видеть, в чем именно состоит недостаток ученика: в слабой ли памяти, в незрелости ли размышления, в недостатке ли красноречия... А не зная причин, которые действуют, мы не можем уничтожить само действие. Простые же отвлеченные цифры, конечно, не могут нам ответить на все эти вопросы, от которых, однако, зависят не только успехи, но, может быть, и вся будущность молодого человека, воспитание которого нам поручено» [249, с. 248].

Критический взгляд К. Д. Ушинского на существующую систему оценивания состоял в том, что нельзя «уничтожить все баллы, и притом без предварительного обдуманного постепенного перехода... Прежде всего нужно парализовать вредное влияние ежедневных отметок в форме цифр...

Первоначально следует приступать к обучению без баллов в высших и элементарных классах, а затем в средних... Баллы за хорошее поведение и прилежание должны быть совсем выведены из употребления. Баллы в виде цифр должны быть навсегда заменены подробными письменными замечаниями, касающимися поведения и успехов ученика» [249, с. 266].

Несмотря на развернувшуюся в 80-е гг. XIX века дискуссию педагогов за сохранение или отмену балльной системы оценивания в школе, в учебных заведениях началось применение 4-балльной шкалы: 2, 3, 4, 5, означавшие соответственно слабые, достаточные, хорошие и отличные знания [221].

Таким образом, в XIX в. после введения различных вариантов оценивания сложилась практика выставления отметок в форме цифр, у которой были как свои сторонники, так и противники. В защиту тех, кто критиковал «баллопроизводство», методика которого описана в цитируемой работе К. Д. Ушинского [249, с. 231], следует признать правомерность высказываемых замечаний еще и в силу недопустимости проведения ряда арифметических действий с качественными данными на порядковой шкале измерений, профессиональное и общественное понимание которого придет намного позднее.

Университеты на определенном историческом этапе обладали большей автономией, что отразилось в первом университетском уставе 1804 года, и не имели общих требований оценивания знаний студентов, устанавливая собственные правила и критерии оценки. Как указывается в работе Д. В. Пещерова, архивные документы императорских университетов – ведомости об успеваемости и поведении студентов, экзаменационные ведомости и листы, учебные таблицы, аттестаты и дипломы – свидетельствуют об использовании 3-балльной словесной шкалы оценивания: неудовлетворительно, удовлетворительно и весьма удовлетворительно – как высшая оценка [207]. Некоторые дисциплины оценивались как зачет или незачет. Однако четвертый устав императорских университетов 1884 года – документ, отражающий политические и социальные условия в государстве и взаимодействие правительства и академического сообщества, – вызвал тогда у современников диаметрально противоположные

реакции, создал существенно обновленное правовое поле отечественной высшей школы. Одно из нововведений касалось отделения функции контроля за качеством обучения от собственно учебы. Уже тогда была предпринята попытка на уровне государственного управления высшим образованием ввести единые экзаменационные требования, по которым должны были проверяться студенты всех университетов [195]. Сами экзамены проводились назначаемой персонально Министром народного просвещения независимой государственной комиссией. В указанном документе укреплялась идея университета как стандартизированного национального учреждения, находящегося под контролем государства и на его финансовом обеспечении.

Наступивший XX век ознаменовался отменой оценок, которая произошла уже в молодой советской России. Постановлением Наркомпроса РСФСР в мае 1918 г. «Об отмене отметок» была упразднена балльная система оценки «познаний и поведения учащихся» во всех школах страны. В это время информационной основой об уровне достижений в учебе являлись отзывы партийных комитетов и педагогических советов школ об исполнении учебной работы. В стране продолжался поиск новых подходов к оценке знаний, поскольку существовавшая до революции система оценивания была признана «уродливой». Инструкция Наркомпроса 1922 года «Об учете учебной работы и проверке знаний учащихся» также содержала отказ от выставления оценок. Такой опыт безотметочного обучения не привел к позитивным изменениям, поскольку личная ответственность как учеников за выполнение учебной программы, так и педагогов за результаты обучения резко снизилась.

Особо следует отметить, что в условиях становления новой советской школы одной из важных задач стала разработка стандартизированных способов сравнения успеваемости, что нашло свое отражение в развитии диагностической практики 20–30-х гг. XX века – оценке личности учащегося и оценке педагогического метода. В этот период происходит бурное и бесконтрольное использование тестов, которые частично заимствовались из зарубежных образовательных систем и разрабатывались отечественными психологами и педологами (П. П. Блонский,

А. П. Болтунов, Л. С. Выготский, А. Ф. Лазурский, А. Р. Лурия, А. П. Нечаев, Ф. Е. Рыбаков и др.) [22, 33, 83]. Направление диагностики обученности и собственно педологии сначала получило официальную поддержку в 1931 году постановлением Наркомпроса РСФСР «Об организации педологической работы в Республике», но через 5 лет, в 1936 году, другим постановлением ЦК ВКП(б) «О педологических извращениях в системе Наркомпроса» метод тестов был признан «буржуазным оружием дискриминации учащихся». В результате многие годы, вплоть до конца XX века, отрицалась сама возможность и образовательных измерений, и педагогической диагностики, и в целом тестирования как объективного метода оценивания в образовании. Вместе с тем необходимо заметить, что за рубежом исследования и поиск оптимальных способов повышения объективности и обоснованности оценочных суждений как в психологии, так и в образовании не останавливались и привели к появлению сначала классической теории тестов (Classical Test Theory), а позднее – современной теории тестов (Item Response Theory), к теоретическим и практическим положениям которых отечественные исследователи обратятся намного позднее [5, 20, 77, 108, 114, 118, 125, 135, 165, 177, 180, 203, 253, 263, 266 и др.].

Вместе с осмыслением ошибочного выбора в пользу безотметочной системы спустя почти двадцать лет постановлением Совнаркома СССР и ЦК ВКП(б) в 1935 г. «Об организации учебной работы и внутреннем распорядке в начальной, неполной средней и средней школе» возвращается пятибалльная система оценивания. Для всех школ Советского Союза разрабатывается единый порядок оценивания результатов обучения и вводится пятиступенчатая оценка успеваемости учащихся – словесные отметки-аналоги вместо цифр: «очень плохо», «плохо», «посредственно», «хорошо», «отлично».

Однако в январе 1944 года постановлением Совнаркома РСФСР «О введении цифровой пятибалльной системы оценки успеваемости и поведения учащихся начальной, семилетней и средней школы» цифры от 1 до 5 снова появляются в оценивании, заменяя словесные отметки. Наряду с вышеназванным постановлением наркомом просвещения РСФСР В.П. Потемкиным в феврале



1944 года была утверждена «Инструкция о применении цифровой пятибалльной системы оценки успеваемости и поведения учащихся начальной, семилетней и средней школы», в которой были даны рекомендации, за что выставлять те или иные баллы. Одновременно оценка приобретает особый статус, поскольку в том же 1944 году утверждено «Положение об экзаменах на аттестат зрелости». В Положении излагался порядок награждения медалями: если выпускник имеет «5» по всем предметам и «отличное поведение», то он награждается золотой медалью. Серебряную медаль получали выпускники за «отличное поведение», оценку «5» по всем предметам, отнесенным к экзаменам на аттестат зрелости, и оценку «4» не более чем по трем из остальных предметов. Что касается вузов, то там с 1978 года оценки также стали влиять на получение дипломов с отличием при соблюдении следующих условий: учащийся получил оценку «5» (пять) не менее, чем по 75 % предметов учебного плана и по поведению, оценку «4» (четыре) по остальным предметам и сдал все государственные экзамены с оценкой «5» (пять), а также проявил себя в общественной работе.

Наряду с отметками в школах и вузах вводилось и безотметочное оценивание. Инструкции министерства, уставы школ предусматривали по отдельным дисциплинам, например по физической культуре, выставление оценки «зачтено» или «не зачтено».

Советский период также характеризуется жестким контролем со стороны органов управления образованием, что объясняется стремлением усилить надзор за выполнением декретов и распоряжений партии и молодого правительства в области народного просвещения и образования. Более того, в годы Великой Отечественной войны был усилен контроль за выполнением требований всеобуча, для руководителей школ и инспекторов вводилась обязанность следить за посещением учащихся, бороться с отчислением учащихся из школ, что прямо сказывалось на оценочной деятельности и выставлении отметок. Также в 1944 году вышел приказ Наркомпроса РСФСР «О социалистическом соревновании в школе», который отменял соревнования среди учащихся и учителей по учебе. Согласно этому приказу запрещалась «неправильная и вредная практика оценки работы

школы и учителя на основе средних процентов успеваемости учащихся», на учителей возлагалась персональная ответственность за правильность оценки успеваемости учеников.

Е. Л. Болотова справедливо отмечает, что нормативные документы только констатировали применение словесной, словесно-цифровой или цифровой системы оценивания, но не раскрывали сущность каждого балла, правил определения итоговой оценки [28]. Такое положение дел приводило к манипулированию отметками ради положительных отчетов об успеваемости и призывам к ликвидации неуспеваемости, как и не изжитые до сих пор издержки «процентомании», и отнюдь не способствовало развитию оценочной деятельности на основе принципов объективности и сопоставимости.

Без преувеличения можно утверждать, что начало 90-х гг. XX века явилось периодом развития научно обоснованных подходов в управлении учебным процессом. Этому способствовали многие обстоятельства, в частности: принятие новых законов Российской Федерации «Об образовании» (1992) и «О высшем и послевузовском профессиональном образовании» (1996), введение понятий «государственная (итоговая) аттестация», «государственная аккредитация», «принцип объективности», «качество подготовки выпускников», «образовательные стандарты» и др., а также получение образовательными учреждениями права самостоятельно выбирать систему оценок, формы, порядок и периодичность промежуточной аттестации обучающихся.

С 1995 года Россия становится участником международных сравнительных исследований – TIMSS (Trends in Mathematics and Science Study – Международное исследование качества математического и естественно-научного образования), с 2000 года – PISA (Programme for International Student Assessment – Международная программа по оценке образовательных достижений учащихся), затем с 2001 года – PIRLS (Progress in International Reading Literacy Study – Международное исследование качества чтения и понимания текста). По результатам этих исследований начался пересмотр содержания образования, методов и средств педагогического контроля, позволяющих оценивать различные виды деятельности

и компетенции обучающихся не только на репродуктивном уровне, но и продуктивном (творческом) уровне, а также начали вырабатываться рекомендации по улучшению обучения [58, 117, 208]. В 1996 году вводится централизованное тестирование выпускников школ в добровольном порядке по единым измерительным материалам и стандартизированной процедуре, предусматривающее дальнейшее поступление с его результатами в высшее учебное заведение.

В системе высшего образования также начинают назревать коренные изменения, которые позднее будут охарактеризованы как смена образовательных парадигм вслед за нормативными изменениями, произошедшими в связи с принятием новой законодательной базы. Одну из ведущих позиций в этих вопросах занял Исследовательский центр проблем качества подготовки специалистов, созданный на базе НИИ при Московском государственном институте стали и сплавов в 1994 году [128, 184]. Именно его коллектив среди основных результатов научно-исследовательских работ предложил концептуальные основы управления качеством высшего образования, разработал научно-методические основы проектирования образовательных стандартов высшей школы, технологии самообследования, аттестации и аккредитации высших учебных заведений [11, 91, 238]. Здесь же вырабатывались методологические подходы к оценке качества образования и результатов обучения с опорой на педагогические измерения [73, 200, 235, 262].

В 1994 году на смену типовым учебным планам пришли первые государственные образовательные стандарты высшего профессионального образования (ГОС ВПО). Следует отметить, что обязательность наличия федеральных государственных образовательных стандартов зафиксирована частью 5 статьи 43 Конституции Российской Федерации и призвана обеспечить единство образовательного пространства по всей стране, а также преемственность между уровнями образования. Позднее они будут иметь общепринятое название «стандарты первого поколения», фиксирующие единые требования к обязательному минимуму содержания и уровню подготовки выпускников. Также

ГОС ВПО содержали требования к структуре образовательных программ, условиям их реализации, нормативы учебной нагрузки на студента, перечень дидактических единиц, знаний, умений и владений по дисциплинам. Такая «нормированная» операционализация минимальных требований к результатам обучения в ГОС ВПО способствовала введению жесткого контроля как за содержательными элементами получаемого высшего профессионального образования, так и за процессами подготовки специалистов.

Как следует из нормативных документов, на протяжении всего двадцатого столетия государство контролирует вопросы оценивания успеваемости обучающихся в школе и в высших учебных заведениях, определяя регламенты, условия, порядки, периодичность выставления оценок. Отметки, рассматриваемые как условно-формальное отражение оценочного процесса, кроме рычага психологического и педагогического воздействия на обучающихся, обладают еще и юридической силой. Это положение вещей достаточно ярко охарактеризовал Ш. А. Амонашвили: «Отметка, которой приписывается лишь невинная роль простого отражателя и фиксатора результата оценки, на практике становится для ребенка источником радости или горя. Не замечать эту действительность – значит допустить серьезный психологический просчет в анализе оценочной стороны обучения» [3, с. 12].

Новый политический уклад в стране в конце XX века вместе с научно-техническим развитием, сопровождающимся информатизацией на всех уровнях, привел к изменениям и в системе образования. Расширение объема содержания образования, повышение доли междисциплинарных и творческих аспектов подготовки обучающихся, разработка и реализация государственных образовательных стандартов, повышение качества образования позволили и теоретикам, и практикам принципиально с других позиций прийти к общему мнению, что для принятия ответственных и обоснованных управленческих решений необходима достоверная информация о качестве образования, которую можно получить благодаря использованию объективных методов оценивания [88, 98, 174]. И действительно, конец 90-х гг. XX века можно охарактеризовать как

время новых научных дискуссий о качестве образования, в которых обсуждается не только управление качеством, но и его оценка. Сама контрольно-оценочная деятельность, как отмечалось выше, привлекала к себе гораздо больший общественный интерес, чем методы и содержание обучения, поскольку напрямую касалась каждого обучающегося, его семьи и общества в целом, характеризовала состояние самой системы образования.

В контексте дальнейшего развития оценочной деятельности наступивший XXI век можно обозначить как период, когда общество всерьез поставило вопрос о качестве образования и сформулировало проблему объективности оценок. Это было обусловлено многими причинами и в число ключевых В. А. Болотов включает следующие:

- «– нарастающий кризис в обществе в целом;
- несоответствие и несопоставимость оценок учащихся в аттестатах об общем образовании даже внутри одной школы;
- резкий рост числа медалистов среди выпускников;
- субъективный отбор абитуриентов на программы высшего образования в институты, академии, университеты;
- разрыв между полученными квалификациями выпускников высших учебных заведений и требованиями рынка труда к молодым специалистам» [23, с. 289].

В феврале 2001 года было принято постановление Правительства Российской Федерации «Об организации эксперимента по введению единого государственного экзамена» и начал разворачиваться эксперимент по введению единого государственного экзамена (ЕГЭ) как прозрачной и справедливой общенациональной оценочной процедуры, выполняющей сразу двойную функцию – функцию аттестации выпускников средней школы и функцию зачисления по его результатам в вуз [64, 76, 88]. Этот эксперимент, являющийся не только строго психолого-педагогическим, но и социальным по своему предназначению, был призван обеспечить справедливость и доступность получения высшего образования. Пожалуй, ни одно нововведение в отечественном образовании не

вызывало таких противоречивых суждений, разногласий и споров, как это. Прошло уже более 20 лет, но дискуссии о нем не утихают до сих пор. Это вполне естественно, так как единый экзамен касается миллионов людей, вовлечённых в орбиту системы образования. Острота критики не уменьшилась, а даже возросла с 2010 года в связи с переходом ЕГЭ в штатный режим функционирования. Частично критика была спровоцирована неоправданным использованием результатов ЕГЭ, поскольку на их основе нередко делали выводы о работе учителей или эффективности школ и даже о деятельности глав субъектов Федерации. Однако была и необоснованная критика тех, кто не испытывал доверия к результатам экзамена и даже не пытался разобраться в их явном преимуществе в плане высокой объективности, обоснованности и прогностичности, удобства для поступления в вузы. Следует признать, что даже определенная часть научно-педагогических работников воспринимает любые тесты как задания с выбором ответов, требующие от экзаменуемых исключительно знания заученной информации, при этом не всегда осознавая, что в существовании «плохих» заданий с ограниченными возможностями виноваты не сами задания, а авторы, не умеющие их разрабатывать [146]. Другая часть привержена консервативным традициям отечественной школы и понимает любой экзамен как привычное и крайне субъективное собеседование с экзаменуемыми по вопросам билета. По сути, сторонники этих крайних позиций не отдают себе отчета в том, что традиционный экзамен не дает практически никаких возможностей для объективного оценивания результатов обучения.

Несмотря на трудности и критику, можно констатировать, что в России удалось создать форму прохождения конкурсных испытаний в вузы, учитывающую как лучший зарубежный опыт, так и своеобразие российских условий, а сам эксперимент по применению точных методов педагогического контроля и технологий образовательного оценивания, основанных на последних достижениях науки, признать состоявшимся [64]. За прошедшие годы с момента появления ЕГЭ развивался и его инструментарий. Содержание заданий приобрело бóльшую проблемность, жесткая стандартизация форм заданий с выбором сменилась приоритетом заданий с развернутыми ответами, появились задания,

проверяющие освоение учащимися метапредметных умений [56]. Совершенствовался банк оценочных средств, в котором основную долю заняли калиброванные задания, обеспечивающие получение валидных и надежных результатов измерений.

Наряду с ЕГЭ в системе профессионального образования стартовал проект «Федеральный интернет-экзамен в сфере профессионального образования» (ФЭПО), который сначала был ориентирован на помощь вузам при подготовке к аттестационному тестированию в рамках государственной аккредитации, а сегодня позволяет проводить внешнюю независимую оценку результатов обучения студентов в рамках требований федеральных государственных образовательных стандартов [25].

В это время в России начинает формироваться сообщество профессионалов, ученых, специалистов в области образовательных измерений, разработчиков заданий, специалистов по компьютерным технологиям, успешно реализующих задачи центров оценки качества образования и развивающих культуру оценивания в общем и высшем образовании [8, 12, 24, 40, 72, 87, 102, 104 и др.]. К числу заслуг таких сформировавшихся коллективов можно отнести не только создание и функционирование Федерального института педагогических измерений (ФИПИ), а также кафедр, центров и лабораторий по проблемам объективного оценивания и измерений в образовании при университетах и научных институтах, но и воспроизводство научно-педагогических кадров соответствующей направленности [217]. Речь идет не только о подготовке кандидатских и докторских диссертаций [14, 15, 19, 49, 55, 57, 62, 67, 88, 93, 106, 108, 126, 225, 243, 265 и др.], посвященных актуальным вопросам обучения и оценивания, но и об открытии новых образовательных программ на уровне магистратуры. Например, назовем магистерские программы Института образования НИУ ВШЭ «Обучение и оценивание как наука» (ранее «Измерения в психологии и образовании»), «Доказательное развитие образования»; магистерскую программу Донского государственного технического университета «Измерения и оценка качества образования»; магистерскую программу Московского городского педагогического

университета «Оценка и мониторинг образовательных результатов в системе общего образования»; магистерскую программу МГУ имени М. В. Ломоносова «Проектирование системы оценки качества образования» и другие.

Кроме того, исследователи и практики образования на регулярной основе в открытом формате проводят международный научный семинар «Измерения и анализ данных в психологии и образовании» на базе МГППУ. Содержание семинара связано с проблематикой измерений в психологии и образовании, с исследованиями в русле доказательного подхода; с вопросами разработки и валидации современного психодиагностического инструментария в соответствии с международными стандартами, анализом больших данных и управлением образованием в контексте процессов его цифровой трансформации, оценкой эффективности применения цифровых технологий в образовании и др. Эти направления деятельности (как и многочисленные конференции по проблемам оценивания в образовании) по праву можно отнести к достижениям по формированию и развитию единого образовательного пространства России [96, 100]. Также к числу успешных практик развития системы независимой оценки профессиональных педагогических компетенций студентов и иных лиц, готовящихся осуществлять педагогическую деятельность, следует отнести опыт Московского городского университета (МГПУ) и функционирующего на его базе соответствующего Центра независимой оценки компетенций студентов.

Хотя пока и не массово, но именно во многом благодаря развитию объективных методов оценивания, пришло понимание субъективности и несопоставимости сложившейся в стране системы выставления оценок на основе экспертных шкал. Глубоко пустившая свои корни и привычная еще с XIX века в системе образования пятибалльная шкала, являющаяся с позиций метрологии и квалиметрии порядковой шкалой, не допускает проведение математических действий сложения, деления и нахождения среднего значения, что, конечно, на практике повсеместно нарушается [132]. Ровно об этом предупреждал в своих работах и К. Д. Ушинский [249]. Фактически пятибалльная шкала сводится к трехуровневой: «отлично» – выше среднего уровня, «хорошо» – средний уровень и



«удовлетворительно» – ниже среднего. Оценка «неудовлетворительно», которая крайне нежелательна всем участникам образовательного процесса, особенно в условиях итоговой аттестации, интерпретируется как «недопустимо ниже среднего» [64]. При этом у каждого экзаменатора свое собственное понимание среднего уровня подготовки, поэтому в действительности получаются субъективные, достаточно грубые и несопоставимые данные об уровне сформированности заданных образовательными программами и стандартами результатов обучения при аттестации как в общем образовании, так и в профессиональном.

Во втором десятилетии XXI века вводились следующие версии федеральных государственных образовательных стандартов (ФГОС), получивших в профессиональном сообществе название «стандарты трех групп требований» – требования к структуре образовательной программы, условиям ее реализации и результатам её освоения. Это второе поколение стандартов для общего образования и третье поколение стандартов для высшего образования, которые под влиянием общественно-политических, социальных изменений и технологического развития страны прошли уже не одну итерацию обновлений.

В основу современных образовательных стандартов общего образования положен системно-деятельностный подход, а образовательные стандарты для профессионального образования, то есть для уровней среднего профессионального и высшего образования, разработаны в парадигме компетентностного подхода. Как в случае общего образования, так и профессионального, образовательные результаты сформулированы в новых трактовках, отличных от тех, которые фиксировались в предыдущих поколениях стандартов в знаниевой парадигме.

За время начала массовой реализации ФГОС ВО на основе компетентностного подхода, т. е. с 2011 года, что совпало по времени с интенсивным развитием компьютерных и информационных технологий, контрольно-оценочная деятельность в высшем образовании получила новый импульс. Это связано не только с формальной стороной вопроса – несколькими версиями ФГОС, уточнением перечня компетенций, относящихся к результатам

освоения студентами образовательных программ, обязательным наличием фондов оценочных средств для осуществления текущего педагогического контроля, промежуточной и итоговой аттестации, функционированием внутренней системы оценки качества образования – но и с изменением понимания сущности компетентностного подхода и его связи с профессиональными стандартами. В частности, компетентностный подход находит свое отражение в формулировках результатов освоения образовательных программ, представленных в федеральных государственных образовательных стандартах по различным направлениям подготовки и специальностям. Это, в свою очередь, влияет на определение содержания обучения, на выбор методов обучения, средств формирования и оценивания образовательных результатов в компетентностной парадигме [65, 112, 148, 173, 185, 218 и др.]. Исследованию компетентностного подхода и характеризующих его дефиниций, анализу условий внедрения и доказательствам его состоятельности для управления качеством подготовки специалистов только за последние десять лет было посвящено, по нашим подсчетам, свыше 1500 докторских и кандидатских работ [16, 46, 50, 70, 130, 179, 233 и др.].

Внимание государства к вопросам качества образования и готовности выпускников системы высшего образования к профессиональной деятельности в очередной раз было подтверждено новыми нормативными документами, повлекшими за собой целый ряд концептуальных и содержательных изменений в контрольно-оценочной деятельности. Так, Указом Президента Российской Федерации от 7 мая 2012 года № 597 «О мероприятиях по реализации государственной социальной политики» была инициирована разработка профессиональных стандартов – документов, определяющих характеристику квалификации (т. е. уровня знаний, умений, профессиональных навыков и опыта работы), необходимой работнику для осуществления определенного вида профессиональной деятельности, в том числе – для выполнения определенной трудовой функции. Эти документы заменили сегодня по многим отраслям и видам экономической деятельности привычные квалификационные справочники [113]. По замыслу разработчиков профессиональных стандартов, соответствие

требованиям рынка труда, выраженное в этих документах в формулировках обобщенных трудовых функций и трудовых функций (необходимых знаний, необходимых умений и трудовых действий) работника, должно оцениваться через независимую процедуру сертификации квалификаций [21, 66, 69, 119, 186, 199].

В основу оценки квалификаций были положены принципы:

- добровольности оценки квалификаций;
- защиты прав лиц, желающих подтвердить свою квалификацию;
- доступности процедур оценки квалификаций;
- обеспечения конфиденциальности в отношении персональных данных, полученных при проведении оценки квалификаций.

В целях возможной дальнейшей полномасштабной реализации такой процедуры в период с 2012 г. по 2013 г. выполнялся проект «Развитие системы оценки качества профессионального образования на основе создания и внедрения механизмов сертификации квалификаций специалистов и выпускников образовательных учреждений с учетом интеграции требований федерального государственного образовательного стандарта и профессиональных стандартов» в рамках федеральной целевой программы развития образования на период 2011–2015 гг. В разработке и реализации данного проекта также принимал участие и автор диссертации. Среди полученных результатов была разработка методического и технологического обеспечения для системы сертификации квалификаций, включающей в себя экспертно-методические центры и центры оценки квалификации. По замыслу участников работ, первая группа центров была призвана осуществлять научно-методическую поддержку системы сертификации квалификаций, а вторая – непосредственно организовывать ее функционирование. Внедрению результатов выполнения проекта и практическому развитию системы сертификации квалификаций в России на тот момент препятствовало отсутствие достаточного числа профессиональных стандартов, проблемы правового и организационного характера и дефицит специалистов-профессионалов в сфере оценочных технологий, реализуемых на основе образовательных измерений. Необходимость обращения к последним была обусловлена высоким уровнем

ответственности при выдаче сертификатов о соответствии определенному квалификационному уровню конкретного профессионального стандарта и вытекающими отсюда требованиями к качеству оценочных процессов, их высокой надежностью и валидностью для обеспечения обоснованности управленческих решений. С точки зрения подходов к проектированию образовательных программ и оцениванию в системе высшего образования как подсистеме общества, в которой происходит подготовка человека к эффективному осуществлению своих трудовых функций, сказанное выше означает, что в данной сфере указанные процессы должны реализовываться гораздо быстрее и полномасштабнее, чем это происходит в иной социальной сфере. По тем видам экономической деятельности, где отсутствуют профессиональные стандарты или требуется их актуализация, вузам необходимо согласовывать формулировки профессиональных компетенций с представителями профессиональных сообществ и работодателями.

Представленные в этом параграфе изменения, касающиеся, в первую очередь, подходов к организации и проведению оценочных процессов в образовании, а также изменения в сфере труда, с которыми неразделимо связана сфера профессионального образования, свидетельствуют об эволюционном развитии в области педагогического контроля и оценки. К началу XXI века среди специалистов в области образовательных измерений утвердилось понимание того, что традиционных оценочных средств и привычных процедур аттестации явно недостаточно для оценивания уровня компетентности студентов, выпускников или работающих специалистов даже в тех случаях, когда задания предусматривают творческий уровень их выполнения и свободно конструируемый ответ [6, 35, 40, 56, 59, 61, 111, 246 и др.]. Достижения в области образовательных измерений позволяют сегодня не только ставить вопросы о содержательной и прогностической валидности, о надежности результатов оценивания, справедливости получаемых оценок – но и отвечать на них [75, 85, 137, 141, 206, 231 и др.].

Стремительное развитие компьютерных мощностей, интернета и методов анализа данных позволяют на иных принципах организовывать аттестационные

процедуры и оценивание. Стало возможно, хотя и не так распространено и широко доступно, проведение многомерных измерений сложных конструктов – таких, как критическое мышление, коммуникация, решение проблем, – не только с помощью заданий с выбором ответов, но и с использованием новых форматов заданий [34, 57, 63, 68, 105, 116, 135, 144, 197, 318 и др.]. Получило распространение компьютерное тестирование, основанное на современной теории тестов (Item Response Theory, IRT) [114, 158, 180, 262, 286, 294, 301, 329, 331 и др.].

Ряд исследователей и практиков стали применять математические методы и алгоритмы, используя различные модели IRT, в реализации компьютерного адаптивного тестирования [81, 129, 142, 167, 239, 252, 332, 339, 344 и др.]. Центральное место среди отечественных работ занимает фундаментальное исследование М. Б. Чельшковой, посвященное теоретико-методологическим и технологическим основам адаптивного тестирования в образовании, которое во многом опередило время [262]. Вслед за основополагающими положениями адаптивного тестирования автору настоящего диссертационного исследования удалось еще до массового перехода к электронному обучению и применению дистанционных образовательных технологий в высшем образовании обосновать эффективность использования адаптивного тестирования учебных достижений студентов в текущем педагогическом контроле и при промежуточной аттестации [81, 139]. И это послужило основой следующего этапа развития и изучения методов компьютерного тестирования и способов оценивания результатов обучения в уже изменившихся социально-экономических условиях и образовательных парадигмах [80, 82, 136, 155, 159, 205, 228 и др.].

Первый массовый выпуск по компетентностно-ориентированным образовательным программам бакалавриата пришелся на 2015 год, а по программам магистратуры на 2013 год. Вне зависимости от уровня высшего образования при итоговой аттестации предусматривается оценивание уровня освоения выпускниками образовательных программ, т. е. оценивание уровня их компетентности как совокупности сформированных универсальных, общепрофессиональных и профессиональных компетенций. При этом в

нормативных документах, регламентирующих порядок проведения государственной итоговой аттестации по образовательным программам высшего образования, зафиксировано, что результаты каждого государственного аттестационного испытания (государственного экзамена, защиты выпускной квалификационной работы) определяются оценками «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно» [189]. Эти оценки выставляет создаваемая для проведения итоговой аттестации экзаменационная комиссия, состоящая, как правило, из числа ведущих преподавателей, представителей работодателей по соответствующему виду экономической деятельности и возглавляемая независимым от образовательной организации председателем. Сама процедура проведения итоговой аттестации практически не изменилась, несмотря на произошедший переход на принципиально другие по содержанию и результатам образовательные стандарты и необходимость учета требований профессиональных стандартов. И если защита выпускной квалификационной работы позволяет экспертным способом оценивать уровень сформированности тех или иных компетенций у выпускника, то подходы к организации и проведению государственного экзамена не обеспечивают получение сопоставимых оценок. Подтверждением этому являются программы ГИА, доступные для анализа и размещаемые всеми образовательными организациями высшего образования в соответствующих разделах официальных сайтов. Несопоставимость оценок усиливается еще одним обстоятельством, заключающимся в том, что перечень компетенций, которые должны быть сформированы у каждого выпускника вуза, является единым в рамках каждого направления подготовки, и отличается только набором профессиональных компетенций. Формулировка последних должна быть увязана с направленностью (профилем) образовательной программы и соответствующим профессиональным стандартом (при наличии). Следовательно, оценочный инструментарий должен быть приведен к единому и однозначному соответствию, чтобы процедура, в том числе, отвечала законодательным, дидактическим и специальным принципам оценивания при итоговой аттестации, которые будут описаны в следующей главе.

Таким образом, в условиях сближения сферы образования и сферы труда через компетентностный подход и профессиональные стандарты по-прежнему актуальной остается проблема получения надежных, валидных и сопоставимых результатов для принятия ответственных и обоснованных управленческих решений в части выставления выпускникам университетов аттестационных оценок и присвоения им квалификации. И если на уровне общего образования эти вопросы во многом сняты благодаря ЕГЭ, то в системе профессионального образования они остаются не до конца решенными и изученными как с теоретико-методологических позиций, так и в практической плоскости [133, 134]. Не упрощает ситуацию и введение демонстрационного экзамена в рамках итоговой аттестации, который на уровне среднего профессионального образования закреплён нормативно, а на уровне высшего образования пока вводится в инициативном порядке рядом университетов [56, 232]. Эти и другие обстоятельства заставляют сегодня пересматривать сложившиеся подходы к оцениванию в высшем образовании и определять целевые ориентиры уже в современных условиях и на ближайшую перспективу.

## **1.2. Особенности и целевые ориентиры современного оценивания выпускников при итоговой аттестации в высшем образовании**

В широком смысле главной целью любого оценивания в образовании является получение всеми заинтересованными лицами и участниками образовательных отношений информации о качестве образования. При этом само понятие «качество образования» постоянно проходит этапы изменений, трактуется и интерпретируется обучающимися, родителями, педагогами, работодателями, общественными объединениями, чаще всего исходя из собственных целевых установок и представлений. Вместе с тем оно нормативно определено в Федеральном законе № 273-ФЗ «Об образовании в Российской Федерации» (далее – Закон об образовании). В статье 2 Закона «качество образования» определяется

как «комплексная характеристика образовательной деятельности и подготовки обучающегося, выражающая степень их соответствия федеральным государственным образовательным стандартам и (или) потребностям физического или юридического лица, в интересах которого осуществляется образовательная деятельность, в том числе степень достижения планируемых результатов образовательной программы» [188].

Общее определение качества, данное в Законе об образовании, должно быть дифференцировано в соответствии с видом и уровнем образования, отличающимся по своим требованиям. Эти требования зафиксированы в образовательных стандартах (федеральных государственных образовательных стандартах – ФГОС либо в самостоятельно устанавливаемых образовательных стандартах – СУОС), а также содержатся в субъективном восприятии, вносящем неопределенность в трактовку качества. Строго говоря, субъективность всегда присутствует в любых оценочных процедурах. Для получения объективной (надежной) и сопоставимой информации об уровне качества образования одна из важных задач – минимизировать эту субъективность, определив соответствующие подходы и технологии, которые и представлены в данной диссертации. Другими словами, качество связано с процессом и результатом образования, при этом каждый из компонентов рассматривается по-разному, в зависимости от того, кто проводит и кто участвует в оценочной процедуре [88]. Если в высшем образовании при оценивании уровня подготовленности студента на первый план выходит соответствие требованиям рынка труда и профессиональным стандартам, то совокупность показателей качества образования обретает конкретный перечень профессиональных компетенций [101, 199].

В данном исследовании оценивание рассматривается в контексте итоговой аттестации завершивших освоение основных профессиональных образовательных программ студентов, по результатам которого, как правило, и делаются выводы о качестве высшего образования и степени подготовленности выпускников к профессиональной деятельности. Однако, как отмечалось во множестве исследований по качеству образования, его результаты носят отсроченный



характер, а выводы должны основываться на лонгитюдных исследованиях [24, 43, 54, 127 и др.]. В частности, такие исследования должны предусматривать получение обратной связи (минимум на протяжении 3–5 лет) от выпускников университетов, завершивших обучение по образовательным программам и получивших соответствующую квалификацию и документ об образовании. Это позволит отследить становление выпускника в профессии, его карьерный путь, успехи и неудачи. Сами результаты исследований должны быть положены в основу мероприятий по улучшению и обновлению содержания образовательных программ высшего образования, коррекции процесса обучения и релевантности формируемых образовательных результатов, прогностичности компетенций, аттестационных оценочных средств и процедур в каждом университете. В свою очередь, все это будет способствовать снижению уровня взаимных претензий как со стороны работодателей, которые чаще всего не учитывают содержание образовательных и профессиональных стандартов, выдвигают собственные требования к выпускникам и формально включаются в образовательный процесс, так и со стороны университетов, которые несут ответственность за подготовку квалифицированных кадров для сферы труда и выполняют государственное задание.

Применительно к высшему образованию качество его результатов, вслед за В. И. Звонниковым и М. Б. Челышковой, можно трактовать как совокупность характеристик, определяющих результаты последовательного и эффективного формирования компетенций студента и его компетентности как выпускника [87]. Делать выводы о качестве высшего образования можно по тому, насколько успешно и эффективно выпускник после освоения образовательной программы осуществляет профессиональную деятельность уже в статусе молодого специалиста, неся всю ответственность за результаты своего труда, с одной стороны, и соответствуя требованиям работодателей в условиях высокой неопределенности, с другой стороны [131]. В целом можно соотнести цели оценивания на итоговой аттестации с каждым субъектом, который прямо или косвенно включен в этот процесс (табл. 1).

**Соотношение субъектов и их целей при оценивании в итоговой аттестации**

<b>Субъект</b>	<b>Цели оценивания на итоговой аттестации</b>
Студент / выпускник	Получение желаемой оценки и диплома о высшем образовании для дальнейшего трудоустройства и получение престижной работы
Преподаватель (при условии его участия в процедуре оценивания на итоговой аттестации)	Получение информации о результатах аттестации каждого студента, его способности применять знания, умения и полученный практический опыт деятельности в решении квазипрофессиональных задач, нестандартно мыслить и ответственно относиться к учебному процессу
Заведующий кафедрой	Получение количественной и качественной информации об уровне подготовленности выпускников образовательной программы, содержательных и организационных выводов об образовательном процессе на реализуемых кафедрой программах
Руководитель образовательной программы	Получение обратной связи от работодателей, принимающих участие в итоговой аттестации, для улучшения содержания образовательной программы
Администратор (ректор, проректор)	Получение информации о качестве образовательной деятельности и качестве подготовки выпускников, их востребованность на рынке труда и дальнейшее трудоустройство
Государственный регулятор, заказчик, учредитель	Соблюдение установленных нормативными документами порядков, трудоустройство выпускников вуза, выполнение государственного задания
Государственный контролер (Рособрнадзор)	Формирование условий для развития культуры соблюдения обязательных требований в сфере образования для гарантии высокого качества

Субъект	Цели оценивания на итоговой аттестации
	образования и предотвращении вреда охраняемым законом ценностям в результате осуществления образовательной деятельности
Работодатель	Отбор выпускников, которые показали высокие результаты на итоговой аттестации, для трудоустройства

На данный момент ситуация в высшем образовании допускает организацию аттестационных процедур в различных режимах – очном, дистанционном, смешанном (т. е. сочетающим очное присутствие с применением дистанционных образовательных технологий, электронного обучения). При этом в любом из режимов проведения аттестации встает вопрос о доверии и справедливости получаемых оценок [231]. Роль принимаемых управленческих решений на основе полученных оценок возрастает по мере продвижения от текущего педагогического контроля к промежуточной и итоговой аттестации. Вполне обоснованной представляется позиция создания в этом психологически напряженном процессе ситуации успеха для каждого испытуемого согласно его способностям [3, 4]. Это означает, что каждому студенту должны предъявляться только посильные для выполнения оценочные задания, которые бы соответствовали его реальному уровню подготовленности. Таким образом, задача повышения эффективности оценочных процедур и качества результатов аттестации (объективность – надежность, обоснованность – валидность и аутентичность) студентов и выпускников представляет не только теоретический, но также и практический интерес.

Компетенции как латентные характеристики проявляются только в деятельности. Эта латентная характеристика, являющаяся причиной того, насколько способны студенты и выпускники выполнять профессиональные задачи, выступает целью образовательного измерения [88]. Но в результате прохождения

студентами аттестационной процедуры получают наблюдаемые оценки их способностей или навыков выполнения квазипрофессиональных задач, по которым и делаются выводы об уровне сформированности латентных компетенций.

Все аттестационные процедуры можно классифицировать по нескольким основаниям. В первую очередь, в основе классификации лежат *дидактические задачи* образовательного процесса [193, 237]. Формирование образовательных результатов, управление этим процессом, определение достоинств и недостатков тех или иных методов обучения, установление взаимосвязи между целями и результатами обучения, оценка достижений и способностей обучающихся, а также эффективность деятельности образовательных организаций и отдельных преподавателей связывают с педагогическим контролем и образовательным оцениванием. В педагогической науке и образовательной практике сложилась классификация по видам педагогического контроля, в рамках которого и проходит аттестация обучающихся, – текущий, тематический, рубежный и итоговый контроль [259]. При реализации оценочных процедур на любом этапе обучения гуманистически ориентированные организаторы и педагоги стремятся, как правило, к соблюдению принципов индивидуализации и дифференциации. Сегодня это становится возможным за счет применения современных информационно-коммуникационных технологий, алгоритмов машинного обучения, при этом соблюдается экономическая целесообразность, минимизируются человеческие, временные и финансовые ресурсы. Но эти же достижения ставят и другую психолого-дидактическую проблему – доверия к самостоятельности полученных студентами результатов в ходе выполнения оценочных заданий, курсовых и квалификационных работ [115, 215].

*Нормативные основания* классификации аттестационных процедур связаны с закрепленной за образовательными организациями функцией осуществления текущего контроля успеваемости и промежуточной аттестации обучающихся, установлением их форм, периодичности и порядка проведения. Кроме того, статьей 59 «Закона об образовании в Российской Федерации» устанавливаются принципы проведения итоговой аттестации – объективности и независимости оценки

качества подготовки обучающихся [188]. Действующие образовательные стандарты содержат требования к механизмам оценки качества образовательной деятельности и подготовки обучающихся (системы внутренней и внешней оценки). Развивающаяся в нашей стране система независимой оценки квалификаций, сопряженная с профессиональными стандартами и непосредственно потребностями рынка труда, также оказывает влияние на содержание оценочного инструментария при аттестации [120, 131, 216].

Наконец, еще одно основание, согласно которому можно разделить все аттестационные процедуры в отношении обучающихся и выпускников, связано с интерпретацией результатов оценивания и уровнем принятия управленческих решений (перевод на следующий уровень образования, выдача документов об образовании, допуск к профессиональной деятельности и др.).

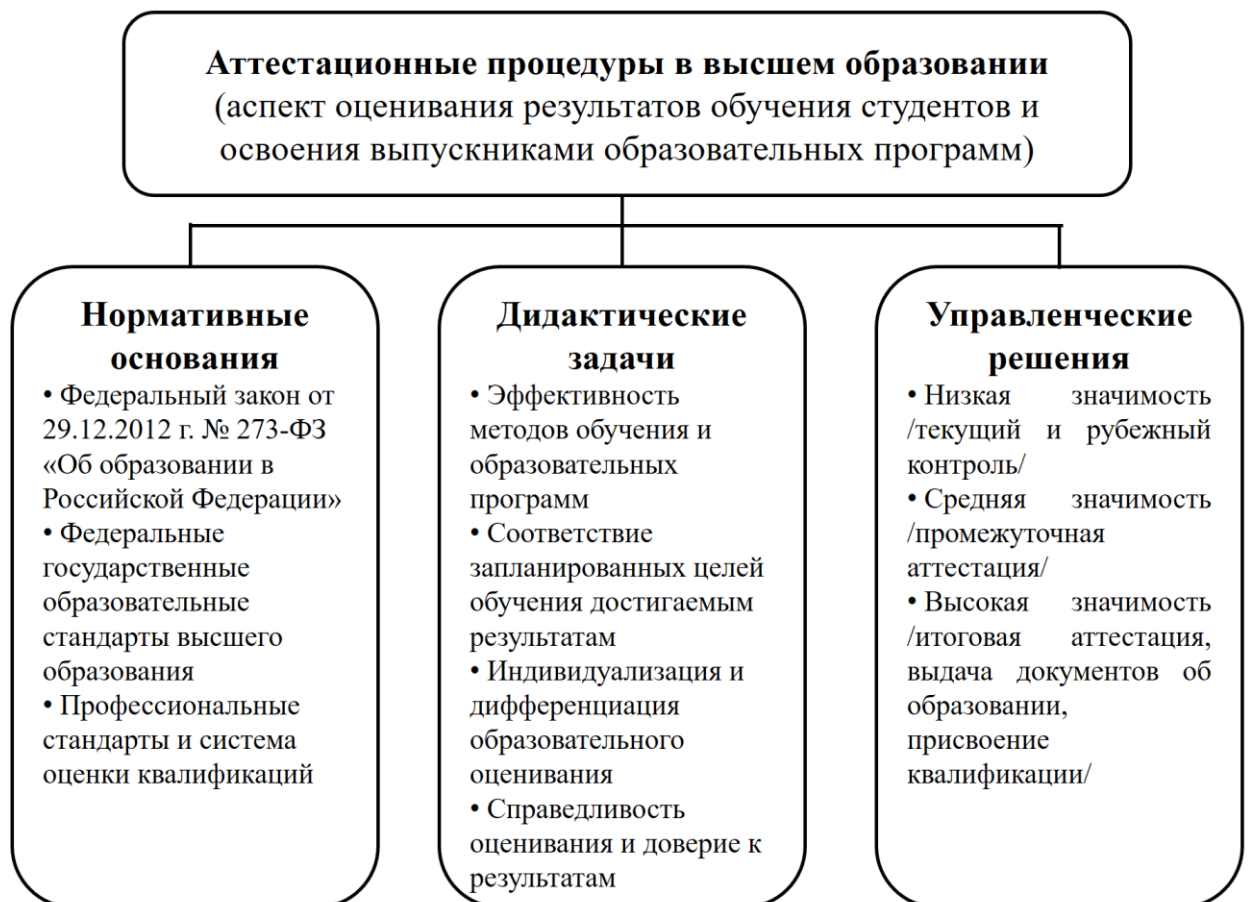


Рис. 1. Классификация аттестационных процедур по различным основаниям

В соответствии с описанной классификацией аттестационных процедур, представленной на рис. 1, изменяются требования к качеству оценочного инструментария. Эти требования возрастают по мере изменения роли принимаемых решений. Такая зависимость проиллюстрирована на рис. 2.

Безусловно, что дидактический интерес представляет вопрос организации и проведения промежуточной и итоговой аттестации. Он обусловлен не только новыми компьютерными и информационными возможностями, но и стремлением создать для каждого аттестуемого студента ситуацию успеха при выполнении оценочных заданий. Учитывая психологические факторы, неизбежные в любой контрольно-оценочной процедуре, социальную роль и воспитывающую функцию оценки, необходимо стремиться к созданию максимально комфортных условий для прохождения аттестации [152]. Вместе с тем следует помнить и о необходимости получения достоверных данных, достижения объективности, сопоставимости и обоснованности выставляемых студентам и выпускникам оценок.



Рис. 2. Зависимость требований к качеству оценочного инструментария от значимости аттестационных процедур

В случае итоговой аттестации на первый план выходит соблюдение принципов объективности и независимости, продиктованных на законодательном уровне в статье 59 «Закона об образовании в Российской Федерации». Абсолютная объективность, равно как и независимость, – это труднодостижимые показатели в любых оценочных процедурах, в том числе – и в образовательных измерениях, к которым принято обращаться, если речь идет о стандартизации контрольно-оценочных процедур. Однако можно и нужно стремиться к минимизации случайных и систематических ошибок, которые присутствуют в любых измерениях, для повышения эффективности самого оценивания, обеспечивая сопоставимость и доверие к получаемым оценкам. Принцип независимости должен соблюдаться через участие представителей работодателя в составе государственной экзаменационной комиссии, которую обязательно возглавляет внешнее по отношению к организации лицо, имеющее ученую степень доктора наук либо большой практический опыт в соответствующей сфере труда. Однако, как свидетельствует практический опыт, в любых оценочных процедурах, если в них принимает участие человек, всегда привносится субъективность в результаты. И если на уровне общего образования эти принципы итоговой аттестации соблюдаются за счет стандартизированных и относительно независимых процедур ОГЭ и ЕГЭ, то на уровне высшего образования этот вопрос до сих пор остается открытым. И тому есть несколько причин.

Во-первых, ставшее возможным с внедрением ФГОС большое многообразие образовательных программ и профилей подготовки, высокая степень свободы у вузов при выборе форм, методов и средств формирования и оценивания заданных образовательными стандартами результатов обучения в компетентностной трактовке внесли значительную неопределенность в проведение итоговой аттестации выпускников. Формы и методы проведения государственного экзамена, если он включен образовательной организацией в состав итоговой аттестации, варьируются в широких пределах: от традиционного прохождения экзамена по двум-трем вопросам всей образовательной программы до компьютерного тестирования с использованием разнообразного инструментария, далеко не всегда

позволяющего оценить уровень сформированности профессиональных компетенций и не отвечающего по своим характеристикам требованиям объективности и обоснованности результатов экзаменов «с высокими ставками».

Во-вторых, неопределенность в итоговой аттестации усиливает разброс требований к качеству подготовки выпускников, обусловленный металатентной природой компетенций и отсутствием грамотно сформулированных индикаторов по многим направлениям подготовки, которые должны в операциональной, т. е. имеющей однозначную трактовку при измерениях, форме отображать наблюдаемые признаки проявления компетенций. По замыслу разработчиков ФГОС, эти индикаторы входят в состав примерных основных образовательных программ, разработка которых в настоящее время не ведется.

В-третьих, ситуация неопределенности в итоговой аттестации усугубляется отсутствием должного внимания со стороны научно-педагогического сообщества к созданию фондов оценочных средств и их применению в стандартизированных условиях при осуществлении оценочных процедур. Следует признать, что пока достаточно сильны консервативные традиции высшей школы, и государственный экзамен воспринимается как привычное и крайне субъективное собеседование по вопросам билета. По сути, классический экзамен не предоставляет никаких возможностей для оценивания уровня сформированности компетенций студентов. Даже если он заменяется выполнением определенного набора тестовых заданий с выбором одного или нескольких ответов, то говорить о надежности и валидности получаемых результатов тоже не приходится. Это связано с большой разобщенностью профессионального сообщества, закрытостью друг от друга, тогда как правильнее было бы двигаться в сторону формирования единых банков оценочных средств и их совместного использования, чтобы пройти широкую апробацию на репрезентативных выборках обучающихся, что также предлагалось и исследователями, и практиками. На уровне высшего образования такая работа по составлению банка оценочных заданий, называемых также калиброванными заданиями, начиналась в 2012 году в учебно-методическом объединении по образованию в области менеджмента, которое тогда возглавлял доктор



педагогических наук, профессор В. И. Звонников. По замыслу авторов этой целесообразной идеи (проф. В. И. Звонников, проф. М. Б. Челышкова), предполагалось открывать доступ к использованию заданий с последующей обработкой результатов их выполнения всем участникам, предоставившим свои фонды оценочных средств в общий банк [87, 261]. Ее реализация позволяла бы тогда говорить и о согласованности требований разных вузов к результатам освоения студентами образовательных программ по одним и тем же направлениям подготовки, и о формировании единого образовательного пространства, а вместе с ним – об открытости и прозрачности аттестационных процедур в высшем образовании [99].

Сложившаяся ситуация в оценивании при проведении итоговой аттестации приводит к тому, что ее результаты отличаются высоким субъективизмом (то есть низкой надежностью), низкой пригодностью для выводов об уровне сформированности компетенций (низкой валидностью) и полной несопоставимостью. На пути к преодолению описанной проблемы предлагается рассмотреть оценивание в четырех областях (морально-нравственной, нормативной, содержательной и инструментально-технологической), определяя в каждой из них целевые ориентиры (рис. 3).

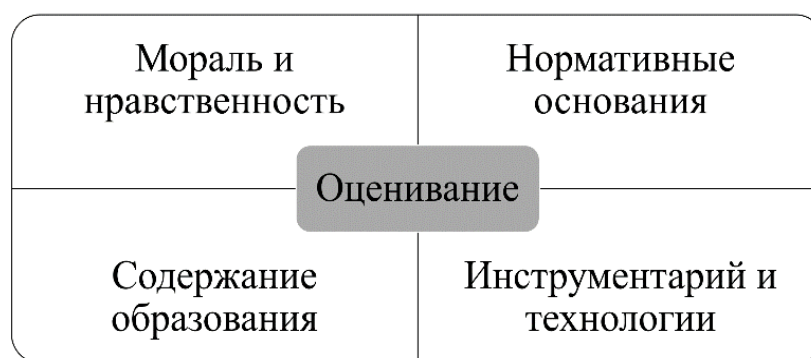


Рис. 3. Области целевых ориентиров оценивания при итоговой аттестации

*Морально-нравственные целевые ориентиры* рассматриваются нами в контексте открывшихся сегодня возможностей персональных когнитивных

помощников/ассистентов, нейросети Chat GPT и др., позволяющих по запросу на естественном языке получать ответы на поставленные вопросы [52, 131 и др.]. Для процедур итоговой аттестации, включающих в себя государственный экзамен и защиту выпускной квалификационной работы, это означает получение ответов на вопросы экзаменационного билета или несамостоятельное написание дипломной работы. Нравственность как внутреннее качество человека, интериоризация морали, конечно, выходит на первый план. Если мы говорим о таких нравственных качествах студента, как честность, порядочность, добросовестность, которые не позволят выдать за свой интеллектуальный труд выполненную роботом-помощником курсовую или выпускную квалификационную работу. Одновременно моральные установки как набор определенных этических норм и правил, стремление получения объективной и справедливой оценки о результатах обучения, учебных достижениях должны формироваться как в университетской среде, так и в обществе в целом. Именно поэтому так важно, стремясь к достижению объективности и сопоставимости, не подавлять студента, а создавать ему ситуацию успеха при оценивании, формировать и соблюдать морально-нравственные установки у всех участников образовательных отношений. Открытость и прозрачность самих аттестационных процедур будут способствовать формированию доверия к результатам, повышению мотивации у студентов, рассчитывающих на получение справедливой и объективной оценки.

С другой стороны, нельзя совсем игнорировать появляющиеся инструменты генеративного искусственного интеллекта, которые можно использовать во благо [51, 133 и др.]. Так, например, А. В. Резаев и Н. Д. Трегубова отмечают, что в условиях трудоемкости экзаменов и зачетов можно с помощью технологий искусственного интеллекта (ИИ) «составлять предварительные вопросы, помогать проводить собеседование и проверку письменных работ, но готовить данные для предварительных вопросов, а также администрировать и оценивать ответы всё-таки должны преподаватели» [215, с. 29].

Целевые ориентиры в области *нормативных оснований* отражены выше и здесь необходимо лишь зафиксировать ряд ключевых документов и позиций в них,

которые следует учитывать при организации оценивания в итоговой аттестации [188, 189]:

– статья 59 «Итоговая аттестация» Федерального закона № 273-ФЗ от 29.12.2012 г. «Об образовании в Российской Федерации» в части соблюдения принципов объективности и независимости оценки качества подготовки обучающихся;

– федеральные государственные образовательные стандарты высшего образования в части требований к результатам освоения студентами образовательных программ и сформированности у них совокупности компетенций, а также применяемых механизмов оценки качества образовательной деятельности и подготовки обучающихся по соответствующим программам – функционирование внутренней и внешней систем оценки качества образования, аккредитационный мониторинг и др.;

– профессиональные стандарты, учет которых необходим при проектировании образовательных программ и определении совокупности профессиональных компетенций;

– положения Порядка проведения государственной итоговой аттестации по образовательным программам высшего образования, предусматривающие проведение аттестационного испытания – государственного экзамена по одной или нескольким дисциплинам (модулям) образовательной программы, результаты освоения которых имеют определяющее значение для профессиональной деятельности выпускников, и выявление уровня подготовленности студентов к решению профессиональных задач.

В области *содержания образования* речь, прежде всего, идет о целях и соотносимых с ними результатах обучения в компетентностной трактовке, сопряжении того, чему мы учим и что формируем, с тем, что и как оценивается в итоге. Компетентностный подход, положенный в основу реализуемых уже более десяти лет образовательных стандартов, как было отмечено выше, вызвал немало содержательных и организационных вопросов. Так, в силу металатентной природы компетенций как образовательных результатов, а также отсроченного характера их

проявления традиционные средства и формы оценивания не позволяют дать ответ на два ключевых вопроса. Тот ли результат обучения оценивается? Можно ли по полученным оценкам предсказать дальнейшие успехи студента в учебной или профессиональной деятельности? Кроме того, при проектировании образовательных программ, определении результатов их освоения студентами, разработчикам необходимо учитывать требования соответствующих профессиональных стандартов [101]. Тем самым и логика компетентностного подхода, нацеленного на сближение с требованиями рынка труда и конкретных работодателей, и нормы положения об итоговой аттестации, предусматривающие выявление уровня подготовленности выпускника к решению профессиональных задач, заставляют обращаться к специальным методам образовательных измерений, обеспечивающих содержательную, конструктивную и прогностическую валидность результатов оценивания, а также их надежность. Такой положительный опыт был накоплен в рамках развития системы аккредитации специалистов здравоохранения [199].

В таблице 2 представлено сравнительное описание уровней значимости решений, характеристик оценочных средств и особенностей их применения, позволяющее выделить инструментарий для итоговой аттестации в отдельную группу. Разработка такого инструментария и условия его применения требуют значительной работы по обеспечению требуемого качества.

**Сравнительное описание уровней значимости решений,  
характеристик оценочных средств и особенностей их применения**

<b>Особенности аттестацион- ной процедуры</b>	<b>Уровень значимости решений</b>		
	<b>низкий (инструментарий для текущего контроля)</b>	<b>средний (инструментарий для промежуточной аттестации)</b>	<b>высокий (инструментарий для итоговой аттестации)</b>
Последствия применения	Незначительные	Значительные частично	Значительные
Решения, принимаемые по результатам применения инструментария	Неважные, поддаются изменению	Могут быть частично изменены	Высокой значимости, исключающие изменение без должного обоснования
Выводы для аттестуемых	Необходима корректировка программы обучения	Устранение выявленных пробелов в знаниях, умениях и дефицитов компетенций	Невыполнение оценочных заданий инструментария приводит к отрицательным (неудовлетворительн ым) решения об аттестации
Причины для обмана со стороны аттестуемых	Незначительные	Средние	Высокие

<b>Особенности аттестацион- ной процедуры</b>	<b>Уровень значимости решений</b>		
	<b>низкий (инструментарий для текущего контроля)</b>	<b>средний (инструментарий для промежуточной аттестации)</b>	<b>высокий (инструментарий для итоговой аттестации)</b>
Документ, удостоверяющий личность аттестуемого	Не важен	Может потребоваться	Важно
Наблюдение за процессом оценивания	Не требуется	Требуется частично	Требуется
Оценка надежности и валидности результатов оценивания	Не требуется	Требуется частично	Требуется

Целевые ориентиры в процессе *конструирования инструментария и выбора релевантных технологий оценивания* для итоговой аттестации выпускников можно наглядно представить в форме направленной прямой линии с реперными точками, соответствующими целям оценивания. Направление прямой на рис. 4 иллюстрирует нарастание трудностей и требуемого сопутствующего профессионализма специалистов в разработке оценочного инструментария для этого аттестационного процесса (рис. 4).

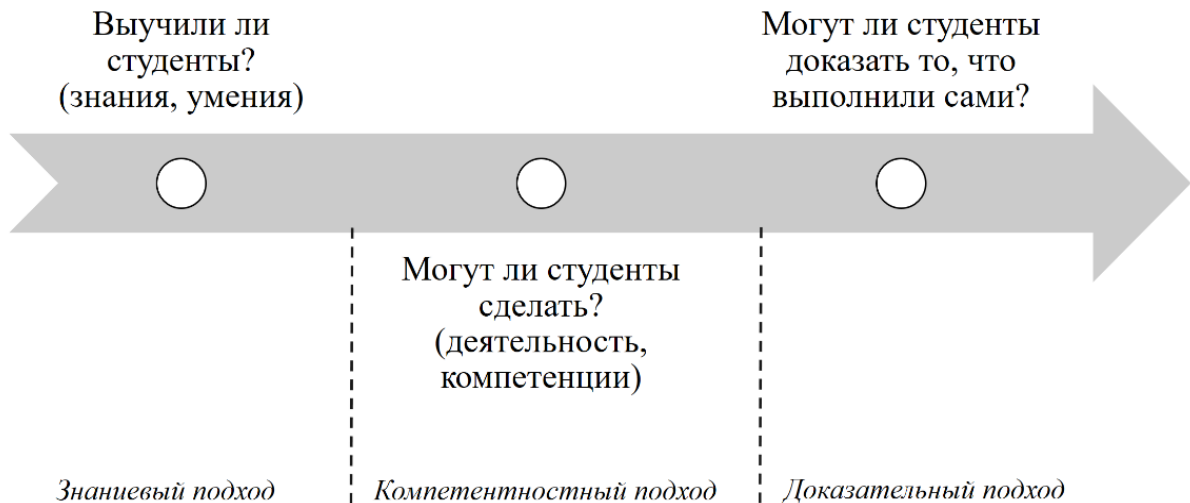


Рис 4. Целевые ориентиры в процессе конструирования инструментария для итоговой аттестации

Первую точку, расположенную слева на рис. 4, можно условно считать оцениванием, направленным на выявление уровня тех знаний и умений, которые необходимы студентам для формирования профессиональных компетенций. Ключевым требованием, выполнение которого необходимо для обеспечения качества оценок знаний и умений при итоговой аттестации, является установление связи оцениваемых знаний и умений с профессиональными стандартами. Определение набора знаний и умений, подлежащих оцениванию, должно основываться не на учебных планах и рабочих программах дисциплин, а на профессиональных стандартах. Это замечание справедливо в том случае, если содержание рабочих программ не учитывает требования документов сферы труда. Логическая цепочка, которой должен руководствоваться разработчик инструментария для аттестации при планировании его содержания, имеет вид, представленный на рис. 5.

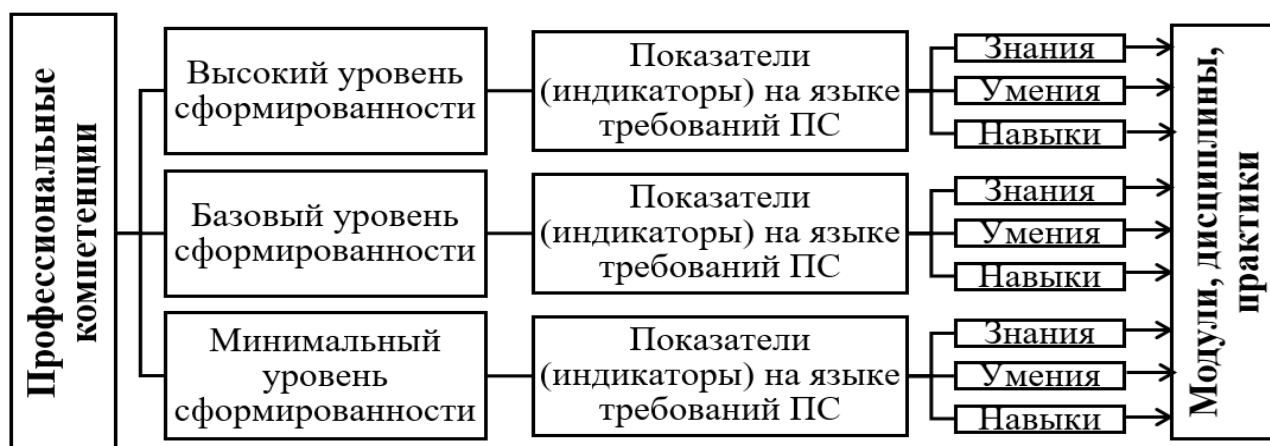


Рис 5. Логика планирования содержания аттестационного инструментария при оценивании знаний и умений, лежащих в основе формирования компетенций

Логика этой схемы продиктована не только нормативными и инструктивными документами, в русле которых можно предложить, чтобы целевые индикаторы профессиональных компетенций в образовательных программах были описаны не только с помощью трудовых функций профессиональных стандартов и соответствующих трудовых действий, но и с учетом опыта практической подготовки оценочного инструментария в сотрудничестве с работодателями [146].

Вторая точка в центре на рис. 4 соответствует задаче оценивания уровня сформированности компетенций, поэтому она предполагает выявление того, как студенты способны применить полученные знания и умения в квазипрофессиональных ситуациях [336]. В качестве оценочных средств для решения второй задачи необходимо использовать ситуационные задания, которые можно отождествить с множественными (многомерными) оценочными кейсами [105, 135].

Крайняя точка справа на рис. 4 иллюстрирует идеи доказательного подхода в оценочных процессах, когда испытуемый должен подтвердить, что данный им ответ не является результатом случайного выбора, подсказки или интуиции, а представляет обоснованное применение собственных знаний и умений для решения в ситуациях, имитирующих профессиональные задачи [85]. На языке



компетентностного подхода доказательства дают основания для выводов о том, что у аттестуемого выпускника действительно сформирована компетенция или совокупность компетенций. Таким образом, опосредованно доказательный подход способствует росту надежности и валидности результатов образовательных измерений.

Ключевой идеей доказательного подхода в оценивании являются радикальные различия между данными и доказательствами [85]. Роль данных в оценочных процессах отводится наблюдаемым результатам испытуемых, полученным с помощью оценочного инструментария. Доказательства – это конкретные, эмпирические свидетельства, позволяющие сделать обоснованные выводы о результатах испытуемых и приблизиться к их истинным баллам [138]. В случае, если подтверждается гипотеза, то данные становятся доказательствами. Само оценивание, согласно S. W. Messick, рассматривается как процесс рассуждений, который идет от анализа наблюдаемых проявлений реакции аттестуемых студентов на предложенные им задания к получению выводов об их знаниях или компетенциях [5, 85, 138].

С развитием доказательного подхода в образовании исследователи и практики связывают повышение качества образования. В первую очередь, речь идет об эффективности тех или иных технологий обучения, о принципиально иной обратной связи с обучающимися, развитии у них аргументации, логического мышления [219]. Оценочные процедуры, строящиеся на доказательном подходе, дают ответы на следующие вопросы:

- какими аргументами или контраргументами руководствовался аттестуемый, отвечая на задания закрытого типа (с одними или несколькими правильными ответами);

- почему происходит неверный выбор;

- каким образом при выполнении оценочных заданий, нацеленных на решение квазипрофессиональных и профессиональных задач, происходит переход от демонстрации теоретических знаний к проявлению способов практической деятельности;

– можно ли обобщить результаты выполнения аттестуемым отдельных заданий оценочного инструментария и сделать выводы о его способностях и компетентности, и аналогично для всей однородной выборки аттестуемых?

В целом доказательный подход ориентирован на создание нового поколения оценочного инструментария и может выступать основой для его научной организации, поскольку обеспечивает координацию работы различных специалистов по созданию подобного инструментария, в частности статистиков, авторов заданий, администраторов процессов предъявления измерителей и проектировщиков их интерфейса [138]. Применение в итоговой аттестации инструментария, включающего тесты и множественные кейсы, в сочетании с доказательным подходом позволяет добиться повышения качества оценок студентов с позиций роста их надежности, валидности и аутентичности. Об этом свидетельствует опыт работы многочисленных тестовых служб, в том числе опыт аккредитации специалистов здравоохранения в России, в проектировании и апробации которой принимал участие и автор диссертации, где множественные интерактивные кейсы (ситуационные задания) выполняются испытуемыми на третьей стадии аккредитации с 2018 года [105, 227].

Вопросы, связанные с выбором инструментария и технологий оценивания, следует рассматривать при соблюдении нескольких условий:

- получение надежных, валидных и сопоставимых результатов итоговой аттестации;
- сочетание количественного и качественного уровней образовательных измерений при одновременном измерении нескольких латентных признаков – компетенций;
- цикличности оценочных процедур: конструирование оценочных заданий и их калибровка, анализ соответствия целям (содержательная валидность), обеспечение информационной безопасности, обнаружение студенческого мошенничества;
- повышение мотивации к выполнению заданий в процессе аттестации за счет создания ситуации успеха для каждого аттестуемого выпускника.

Последнее условие, связанное со снижением психологической нагрузки на аттестуемых студентов, создание им ситуации, когда предъявляемые в процессе аттестационной процедуры задания оказываются трудными, но посильными для выполнения, – вполне возможно, если опираться на современную теорию тестов IRT. Полученные ранее автором диссертации результаты исследования применения адаптивного тестирования, строящегося на математических моделях IRT, в промежуточной аттестации студентов при дистанционной форме ее организации подтвердили состоятельность гуманистического и одновременно технологического подходов к организации педагогического контроля [138].

Отметим наиболее значимые факторы среды, определившие наш исследовательский интерес. К их числу относятся:

- стремление к повышению объективности, ответственности и обоснованности в принятии различных управленческих решений по степени их важности;

- развитие научных подходов и взглядов на организацию оценочных процедур в условиях актуальной трактовки качества результатов образования в рамках компетентностного подхода;

- существенное расширение возможностей компьютеров в обработке большого массива данных, совершенствование алгоритмов машинного обучения.

Именно эти факторы позволили автору настоящего исследования предложить иной подход к организации аттестационных процедур. Опираясь на сформулированные выше целевые ориентиры современного оценивания, учитывая логику построения содержания образовательных программ и этапы подготовки студентов к конкретным видам профессиональной деятельности, и, самое главное, значимость оценивания с высокими ставками в условиях компетентностного подхода, автор диссертации предлагает и обосновывает применение многостадийного оценивания уровня сформированности компетентности выпускников при итоговой аттестации.

Сама идея оценочной процедуры в несколько этапов, по результатам которой должны приниматься ответственные и обоснованные управленческие

решения, интуитивно понятна как теоретикам, так и практикам. Действительно, благодаря сочетанию различных свойств и возможностей оценочных средств и многоэтапности самой процедуры можно повысить надежность и валидность результатов, оценить различные способы деятельности выпускников и сделать аргументированные выводы с минимальными классификационными ошибками. Однако это будет возможным, если будет должным образом разработано методологическое обоснование, описана вся технология многостадийного оценивания и показана экономическая целесообразность.

Вместе с тем в практике аккредитации специалистов здравоохранения сейчас используется модель многостадийных измерений, обоснование которой выполнялось в рамках указанного выше проекта «Развитие системы оценки качества профессионального образования на основе создания и внедрения механизмов сертификации квалификаций специалистов и выпускников образовательных учреждений с учетом интеграции требований федерального государственного образовательного стандарта и профессиональных стандартов». Модель была предложена В.И. Звонниковым и М.Б. Чельшковой и рассматривалась в работах [199, 219]. Многостадийность оценивания фрагментарно описывалась в диссертационном исследовании М. В. Гуськовой в контексте изучения процесса эвалюации в управлении качеством образования [55], а также в кандидатской диссертации Т. Н. Малаховой «Проектирование инструментария для аккредитации выпускников медицинских вузов» (2021 г.).

Близким по замыслу также является подход группы исследователей (Н. А. Бочарова, С. А. Писарева, М. Ю. Пучков, В. И. Снегурова, А. П. Тряпицына), выполнявших проект «Организация и проведение мероприятий по уровневой оценке компетенций учителей русского языка и математики в 2017 году», в рамках которого уточнялись «имеющиеся подходы к построению научно обоснованной и практически целесообразной системы аттестации учителей, а также построению национальной системы учительского роста» [120, с. 170]. Авторы этой концепции, в частности, вводят три уровня оценивания и берут в качестве содержательной основы для диагностической работы требования ФГОС

начального общего, основного общего, среднего общего образования и профессионального стандарта педагога. На первом уровне оценивается действие по образцу – умение решать задачи профессиональной деятельности с заранее определенным условием, на втором – умение решать задачи профессиональной деятельности в изменяющейся ситуации, отражающей реальные не запланированные условия, и способность выбирать разнообразные средства их решения; на третьем уровне – умение решать задачи профессиональной деятельности, действуя в ситуации, требующей не только учета развивающихся и усложняющихся контекстов (условий) реальной деятельности, вариативность средств решения, но и использование новых ресурсов для их решения [120]. Такая процедура не предполагает выставления оценок, как в рамках итоговой аттестации в высшем образовании. Педагоги последовательно выполняют три части диагностической работы. При этом, если учитель получает менее 80 % от максимального количества баллов на первом этапе оценки предметной компетенции, то он не допускается к выполнению следующих частей диагностической работы [171].

Отметим, что в перечисленных исследованиях, как и в других отечественных работах по смежной проблематике, связанной с образовательными измерениями и оцениванием в профессиональном образовании [8, 16, 38, 46, 50, 60, 67, 70, 218 и др.], комплексно не рассматривались вопросы многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании с представлением соответствующей методологии и соответствующей технологии. Такая задача ставится в настоящем исследовании и предлагается ее решение, обеспечивающее получение надежных, валидных и аутентичных результатов для обоснованных классификационных выводов об уровне сформированности компетентности аттестуемых выпускников.

### **1.3. Понятийно-терминологический аппарат многостадийного оценивания при итоговой аттестации в высшем образовании**

Вопрос, связанный с определением понятийного аппарата, является первоочередным при обосновании способов организации и построения многостадийного оценивания при итоговой аттестации. Структуру и классификационную основу понятийного аппарата можно провести, рассматривая базовую группу родовых понятий и терминов и переходя к конкретным терминам, отражающим специфику многостадийного оценивания и его обоснование для итоговой аттестации выпускников в высшем образовании. При этом особо отметим следующее. В процессе эволюции одна группа терминов в контрольно-оценочной деятельности в образовании была определена в результате фундаментальных дидактических работ или закреплена нормативно. Другая группа, составляющая систему базовых понятий исследования, интуитивно понимаемая многими субъектами образовательного процесса, нуждается как в уточнении, дополнении, переосмыслении, так и в операционализированном определении с учетом последних научно-технологических достижений в области образовательных измерений и оценивания.

Развитию понятийного аппарата контрольно-оценочной деятельности в образовании во многом способствует пересмотр сложившихся традиционных подходов к ее организации, поиск новых форм и методов педагогического контроля и оценивания, отвечающих на современные вызовы. Ключевая идея исследования состоит в выделении совокупности базовых определений, которые включают в себя признаки, отделяющие многостадийное оценивание от других смежных понятий.

Эта процедура должна выстраиваться:

- на понимании природы компетенций и их связи с будущей профессиональной деятельностью студентов,
- на изменении подходов к оцениванию этих сложных конструктов, организационно-педагогических условий при сочетании количественных и качественных уровней образовательных измерений,

– на стремлении к повышению их эффективности при обеспечении высокой надежности, валидности и аутентичности результатов оценивания.

Разработка понятийного аппарата должна вестись с соблюдением большого охвата и полноты рассматриваемых категорий и вводимых терминов, при этом стараясь избегать противоречий при их использовании. Поэтому в данном исследовании выбран путь обратный обобщению, т. е. не от первопонятий к производным, обобщенным терминам, а напротив – от общих определений терминов и понятий к последующей детализации и развертыванию в ряд конкретизирующих, уточняющих и нацеленных на отдельные существенные признаки оценочного инструментария и многостадийного оценивания. Именно этот путь представляется автору наиболее предпочтительным, чтобы уйти от уточняющей конкретизации, которой изобилуют многие исследования в области оценивания и образовательных измерений. В целом же сложность рассматриваемого объекта предопределяет трудность его описания с помощью ограниченного числа признаков и характеристик, что также свойственно теоретической деятельности при создании понятийного аппарата в других науках [262].

В своих работах М. Б. Челышкова при определении понятийного поля в контрольно-оценочной деятельности предлагает выделять сложные определения в отдельный класс и использовать специальный прием, состоящий во введении обобщающих слов для обозначения группы признаков или характеристик [262]. При этом требуются разъяснения введенных обобщающих слов после определяемого понятия. В этой логике и ведется формирование понятийно-терминологического аппарата через рассмотрение исходных, основных и производных понятий, которые должны характеризоваться полнотой, функциональностью и корректностью.

Применительно к оценке качества образования и определению уровня подготовленности студентов и выпускников следует выделить *аккредитацию*, *сертификацию* и *аттестацию*. Эти три термина имеют отношение как к сфере образования, так и к сфере труда, т. е. будущей профессиональной деятельности обучающихся, и непосредственно связаны с оцениванием.

Под *аккредитацией* (лат. *accredo* – доверять) понимается процедура официального подтверждения соответствия объекта установленным показателям и критериям. Эти показатели, как правило, для сферы образования связаны с образовательными стандартами. В настоящее время процедура государственной аккредитации претерпевает определенный этап модернизации, когда на смену разовой процедуре (один раз в 6 лет для уровней профессионального образования) приходит аккредитационный мониторинг, предусматривающий систематическое стандартизированное наблюдение за выполнением организациями, осуществляющими образовательную деятельность, аккредитационных показателей. К числу утвержденных показателей аккредитационного мониторинга для высшего образования<sup>1</sup>, непосредственно связанных с предметом нашего исследования, относятся:

- средний балл ЕГЭ обучающихся, принятых по его результатам на обучение по образовательным программам;
- доля обучающихся, успешно завершивших обучение по образовательным программам высшего образования, от общей численности обучающихся, поступивших на обучение по соответствующей образовательной программе;
- наличие внутренней системы оценки качества образования.

В сфере труда этот термин с 2016 года официально закреплен для процедуры аккредитации специалистов здравоохранения, где имеет расширенную трактовку в отличие от термина «сертификация» [199]. Так, в нормативных документах системы здравоохранения аккредитацией называется процедура, при которой определяется соответствие готовности лица, имеющего медицинское образование, к осуществлению профессиональной деятельности по определенной специальности<sup>2</sup>.

---

<sup>1</sup> Приказ Рособрнадзора № 660, Минпросвещения России № 306, Минобрнауки России № 448 от 24.04.2023 «Об осуществлении Федеральной службой по надзору в сфере образования и науки, Министерством просвещения Российской Федерации и Министерством науки и высшего образования Российской Федерации аккредитационного мониторинга системы образования» (URL: [www.consultant.ru](http://www.consultant.ru), дата обращения: 24.08.2023).

<sup>2</sup> Федеральный закон от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» (ст. 69. Право на осуществление медицинской деятельности и фармацевтической деятельности).



Близкий по смыслу к аккредитации термин *сертификация* (лат. *certum* – верно + лат. *facere* – делать), под которым понимается подтверждение соответствия объектов требованиям технических регламентов, положениям стандартов, своду правил или условиям договоров, осуществляемое специальным органом по сертификации. Объектами сертификации чаще всего выступают услуги, продукция, системы менеджмента и персонал.

Рассматриваемое в данном исследовании в качестве одного из основных понятие *аттестация* (от лат. *attestatio* – свидетельство) – в образовании обозначает процесс подтверждения квалификации студента и выпускника путем выявления соответствия его уровня компетентности требованиям образовательной программы, разработанной на основе образовательных стандартов и с учетом требований профессиональных стандартов (при их наличии). В сфере труда этот термин может использоваться в деятельности организации, когда на основании подтверждения уровня подготовки, мастерства или квалификации работника ему выдают аттестат, а отказ в выдаче аттестата является одним из оснований расторжения трудовых отношений по инициативе работодателя<sup>3</sup>.

В отличие от терминов «аккредитация» и «сертификация» термин «аттестация» используется только в отношении физического лица. Конечно, представленная трактовка термина «аттестация» относится больше к итоговой, чем промежуточной аттестации, но как в первом, так и во втором случаях – принимаются решения в отношении обучающегося студента о его дальнейшей образовательной траектории. Отличие этих видов аттестаций будет приведено ниже в классификационной схеме оценивания. Здесь важно сделать акцент на *итоговой аттестации* – оценочной процедуре определения уровня сформированности компетентности выпускников и присвоения им квалификации, которая проводится после завершения освоения образовательной программы и должна основываться на принципах объективности и независимости.

---

<sup>3</sup> «Трудовой кодекс Российской Федерации» от 30.12.2001 № 197-ФЗ.

Термин *оценивание* (от лат. *assidere* – сидеть с), затрагивающий практически все сферы человеческой деятельности, в образовании связан с процессом формирования оценки, которая впоследствии может быть переведена в отметку, для принятия дидактических и управленческих решений. Такие решения варьируются от оценки уровня достижений обучающегося до соответствия образовательной программы или образовательной организации устанавливаемым требованиям, а иногда и всей системы образования. Следует также отличать другой термин, связанный с оценкой, – *эвалюация*. В отечественных исследованиях впервые он был введен и рассмотрен авторами коллективной монографии «Педагогическое образование в университете: контекстно-биографический подход» еще в 2001 году [202]. Впоследствии исследование *эвалюации* получило развитие в научных трудах В. И. Звонникова, М. В. Гуськовой, В. К. Загвоздкина, Н. Н. Найденовой [55, 78, 88, 172]. В указанных работах подчеркивалось, что прямой перевод англоязычного термина «*evaluation*» как оценивание не тождественен переводу другого англоязычного термина – «*assessment*». Последний действительно следует относить к образовательному оцениванию, он связан непосредственно с аттестационными процедурами, тогда, как термин «эвалюация», введенный в научный оборот В. И. Звонниковым [83], следует рассматривать гораздо шире, считая эвалюацию интегрированной категорией оценочно-аналитической деятельности в управлении. В содержание этого термина вкладывается весь спектр теоретико-методологических и практических работ по системному исследованию ценности и позитивности воздействия проектов и программ при достижении поставленных целей [54]. Эвалюация может проводиться в различных сферах, не только в образовании, но и в здравоохранении, строительстве, государственном управлении, социальной политике, позволяя оценивать эффекты и результаты от реализации тех или иных программ, включая проекты национального уровня, и их влияние на повышение и улучшение качества жизни [198].

В данном исследовании *оценивание* рассматривается именно в контексте аттестационных процедур в образовании, т.е. это аналог *assessment*. По сути,

оценивание проводится, прежде всего, с обучающимся и для обучающегося. Образовательное оценивание отличается от любого другого оценивания, поскольку в центре этого процесса, как правило, находится обучающийся (ребенок, подросток, взрослый человек), а технологии и содержание отражают степень развития общества и его культурные традиции. В условиях итоговой аттестации студентов *оценивание* мы рассматриваем как процесс формирования оценки результатов освоения образовательной программы и уровня сформированности компетенций, в котором интегрируются и представляются в определенной шкале (шкалах) данные, полученные в результате стандартизированной процедуры образовательного измерения с использованием различных измерителей.

Обращение к теории измерений вызвано стремлением исследователей и практиков получить объективную информацию о реальном уровне результатов обучения как отдельного обучающегося, так и в целом о системе образования в разрезе образовательных программ, образовательных организаций, групп обучающихся, регионов и страны. Теория образовательных измерений лежит в основе международных сравнительных исследований качества образования, поскольку в этих процедурах требуются специальные методы шкалирования и выравнивания результатов измерения, способы генерализации (распространения выводов на всю генеральную совокупность исследуемой выборки) получаемых данных.

Понятно, что в текущем или рубежном контроле, при формирующем или диагностическом оценивании можно обойтись традиционными оценочными средствами и сложившимися практиками в контрольно-оценочной деятельности. Но на уровне промежуточной аттестации (переход из класса в класс, с курса на курс), итоговой аттестации (выдача документов об образовании – аттестатов об общем образовании, дипломов о высшем образовании), а также допуска к профессиональной деятельности (сертификация квалификаций, аккредитация специалистов) ужесточаются требования к качеству оценочного инструментария и процедурам их применения. По этой причине в оценочных процедурах аккредитации, сертификации и аттестации не используют авторские оценочные

средства, разработка которых не опирается на соответствующий научный аппарат и не включает этапы апробации и обоснование качества инструментария. Такой научный аппарат, на котором предлагается выстраивать оценочные процедуры в итоговой аттестации в данном исследовании, состоит из бипарадигмального подхода в образовательных измерениях и теории тестов, концептуальное обоснование которых представлено в следующей главе.

Определяя понятийно-терминологическое поле исследования, следует обратить внимание на использование такой категории, как «образовательные измерения». В отечественной педагогической литературе встречается несколько достаточно распространенных названий – «педагогические измерения» и «психометрика». Первая категория – «педагогические измерения» – сложилась в нашей стране еще в 80-х годах XX века, на новом этапе развития тестовых процедур оценивания учебных достижений обучающихся, и необходимости обоснования качества разрабатываемых педагогических тестов. Исследователи, опираясь на зарубежные разработки, стали переводить «*educational measurements*» как «педагогические измерения», отделяя тем самым измерения в социальных науках в целом и подчеркивая их место в дидактике, развивая строгие формализованные методы в педагогическом контроле и оценке. Столь же не вполне удачным применительно к задачам образовательного оценивания представляется употребление категории «психометрика», связанной в первую очередь с психодиагностикой, разработкой психологических тестов и с измерениями психологических конструктов.

Вместе с тем, основой как для психометрики, так и для педагогических измерений выступает методология, включающая в себя классическую теорию тестов, современную теорию тестов и теорию генерализации, позволяющую обобщать результаты измерений, полученные на определенной выборке испытуемых на всю генеральную совокупность. Используемый в названиях указанных теорий термин «*тест*» трактуется сегодня в широком смысле – как любой измеритель, отвечающий требованиям со стороны теории.

Наряду с указанными категориями, в теории и практике дидактики активно используется также термин «*диагностика*» (от греч. *δια-γνωστικός*, *dia* – между, через, после; *gnosis* – знание, познание, понятие; *diagnostikos* – способность распознавать), близкий, но не тождественный термину «измерения». Существенное различие состоит в том, что диагностика признает любые методы и средства контроля и оценивания, даже если они не отвечают научным критериям качества измерений. В. И. Звонников отмечает, что «в рамках диагностических процедур не разработано никакого аппарата для оценки надежности и валидности эмпирических данных, да к этому в диагностике обычно и не стремятся» [83, с. 192]. Диагностика может рассматриваться как самостоятельная теория распознавания, основными задачами которой являются исследование закономерности диагностического процесса и изучение методологии и методик диагностирования [9]. Различают психологическую и педагогическую диагностику. Е. А. Михайлычев рассматривает педагогическую диагностику как процесс установления и изучения признаков, характеризующих состояние различных элементов педагогической системы, условий ее реализации, для совершения прогнозирования возможных отклонений и нарушений [169]. Как утверждает Е. К. Артищева, «педагогическая диагностика отождествляется с практикой или процедурой выявления качества образовательного процесса, со способом получения информации о ходе учебного процесса и его субъектах, с функциями, задачами и целями диагностики либо только ее структурой, диагностической деятельностью» [9, с. 36]. При этом «педагогическая диагностика как педагогическое явление включает в себя психологическую диагностику и дидактический контроль, а также педагогический мониторинг и педагогическую прогностику» [там же, с. 88]. Таким образом, педагогическая диагностика рассматривается как достаточно широкая категория, а ее методы классифицируются по степени формализованности, отражают регламентацию структуры диагностического обследования и обработки его результатов. Многие в педагогической диагностике основываются на профессиональном опыте и компетентности диагноста. Для сбора данных в диагностике широко

распространены субъективные процедуры, наблюдения, при осуществлении которых каждый выбирает свою совокупность методов и средств. При диагностике для разных объектов одного и того же ряда нередко анализируются и интерпретируются данные разной природы, полученные на основе субъективных оценок и в процессе наблюдения и потому, в силу значительной ошибки измерения, не адекватные состоянию исследуемых объектов. Очевидно, что выводы, сделанные на основе таких оценок, не способствуют выявлению качества объектов оценивания и не дают полного представления о состоянии объектов, исключая возможность их корректного сравнения [88].

Также следует отметить, что в понятийно-терминологическом поле ряда исследований, рассматривающих различные аспекты качества образования, – управление, оценка, измерение, обеспечение – рассматривается категория «*квалиметрия*», а в качестве одного из методологических оснований выступает квалиметрический подход [46, 70, 76, 130, 225, 243, 250 и др.]. Действительно, в 70-х годах прошлого века научное направление, занимающееся изучением и разработкой принципов и методов количественной оценки качества и определенное как *квалиметрия*, получило достаточно широкое развитие [1]. Несмотря на то, что в данной научной области рассматривались вопросы измерения и количественной оценки качества продукции в экономическом и промышленном секторах, многие подходы и методы были перенесены и в образование. Это было предопределено временем и задачами, стоявшими перед всей системой образования в конце 90-х годов XX века и начала XXI века [128, 184, 211, 234, 248]. Сегодня, анализируя современные исследования, посвященные качеству образования и его оценке, можно констатировать снижение квалиметрической терминологии, хотя стоит согласиться, что в основе квалиметрического подхода в образовании лежат образовательные измерения [88, 243].

Опираясь на широкую трактовку самого понятия «образование», включающего в себя процессы обучения, воспитания и развития личности, автор исследования предлагает использовать в терминологическом аппарате дидактики профессионального образования и контрольно-оценочной деятельности категорию

«образовательные измерения». Данная категория в полной мере не только соответствует англоязычному аналогу области научного знания, связанного с теориями о поведении человека, развитием объективных методов и подходов к оценке процесса и результата, но и отражает суть решения сегодняшних и стоящих в перспективе задач в области образовательного оценивания. Под *образовательными измерениями* предлагается понимать научно-практическое междисциплинарное направление, включающее обоснование всех взаимосвязанных аспектов разработки инструментария, его применения для сбора, обработки, анализа и интерпретации результатов измерений в целях развития самого инструментария и оценивания аттестуемых, получения надежных, валидных и аутентичных результатов для принятия обоснованных решений различного уровня управления в системе образования. Междисциплинарность образовательных измерений определяется используемым аппаратом различных наук, на которых выстраиваются непротиворечащие друг другу процедуры. Образовательные измерения – это категория, базирующаяся на положениях психологии, педагогики, теории управления, социологии и математико-статистические методах.

В качестве доказательного фундамента предложенного определения можно провести сравнение традиционного оценивания (включая диагностику), базирующегося, как правило, на традициях, и оценивание, в основе которого лежат образовательные измерения.

*Таблица 3*

**Сравнение традиционного оценивания и оценивания на основе образовательных измерений**

<b>Традиционное оценивание, включая диагностику</b>	<b>Оценивание на основе образовательных измерений</b>
Оценочное средство создается на основе опыта, его качество определяется интуитивно	Оценочное средство создается на основе определенной технологии, его надежность и валидность оценивается

<b>Традиционное оценивание, включая диагностику</b>	<b>Оценивание на основе образовательных измерений</b>
	численными методами и подтверждается эмпирическими данными
Нет возможности определить ошибку оценивания и принять решение о степени доверия к его результатам	Есть возможность определить ошибку измерения и принять решение о степени доверия к результатам измерения
Нет возможности проанализировать степень адекватности полученных оценок концептуально выбранному предмету оценочного процесса	Есть возможность проанализировать степень адекватности полученных оценок концептуально выбранному предмету измерения – величину конструктивной валидности
Результаты не сопоставимы	Можно обеспечить сопоставимость результатов
Результаты только качественные	Результаты количественные и качественные

В первом случае при традиционном оценивании инструментарий – оценочное средство – создается на основе эмпирического опыта его автора, а качество такого средства оценивается интуитивно. Во втором случае при образовательных измерениях оценочное средство создается на основе технологии, его надежность и валидность оценивается численными методами на базе эмпирических данных, полученных путем апробации оценочного средства на выборке испытуемых.

Зачастую в повседневной деятельности любой педагог встает перед выбором, какую отметку выставить студенту – «отлично» или «хорошо», «хорошо» или «удовлетворительно». В случае традиционного оценивания нет возможности определить ошибку оценки и принять решение о степени доверия к результатам



оценивания. Иначе выглядит ситуация в случае обращения к образовательным измерениям, когда оценку величины ошибки можно получить с помощью математико-статистического аппарата и на этом основании уже принять или отвергнуть полученные результаты. Важное преимущество оценивания на основе образовательных измерений по сравнению с традиционным оцениванием – это возможность определить, действительно ли оценивается то, что планировалось. В случае экзаменов высокой значимости это преимущество многое определяет. Оценивая результаты обучения привычными методами, контрольные задания подбираются на основе интуиции. При измерениях используются методы факторного и корреляционного анализа, позволяющие оценить степень адекватности полученных оценок концептуально выбранному предмету измерения. При традиционном оценивании невозможно говорить о сопоставимости выставляемых оценок не только у различных преподавателей по одним и тем же дисциплинам, но даже и у одного педагога. Этому есть несколько причин: различные критерии оценивания, собственные оценочные суждения и эталоны для сравнения, субъективное отношение к обучающемуся, усталость и т. д. В случае образовательных измерений сопоставимость обеспечивается специальными технологиями конструирования измерителей, шкалирования результатов их выполнения и математическим аппаратом выравнивания, позволяющим перевести результаты испытуемых в единую метрическую шкалу. И, наконец, благодаря образовательным измерениям можно перевести оценки в количественную интервальную шкалу.

Все эти неоспоримые преимущества образовательных измерений особенно важны в связи с задачей оценивания уровня сформированности компетентности при итоговой аттестации и служат теми самыми основаниями, обуславливающими выбор строгих методов получения обоснованных и сопоставимых оценок.

Во избежание ситуаций недооценки (ложно отрицательной классификации) или переоценки (ложно положительной классификации) способностей выпускников на аттестационной процедуре требуется введение в нее понятия *критериального балла* (или порогового балла, минимального стандарта оценивания

– всё это можно считать синонимичными понятиями). Оно связано с одним из подходов к интерпретации результатов выполнения оценочных заданий (измерителей) в образовательных измерениях – *критериально-ориентированным*. Такой подход, в отличие от нормативно-ориентированного, широко распространен в процедурах сертификации и лицензирования в сферах профессионального отбора кадров, когда вводится, по Эбелю, стандартная оценка содержания и требуется соотнести результаты прохождения испытуемым оценочной процедуры не с результатами других испытуемых, а с определенной содержательной областью или видами деятельности (областью поведения) [5, 125, 302]. Термин «критериально-ориентированное измерение» был впервые введен Р. Гласером в 1963 году [307]. В различных источниках можно встретить другие альтернативные термины – предметно-ориентированное, содержательно-ориентированное измерение и т.п. Это не меняет главной сути такого подхода – интерпретации выполнения оценочных заданий с точки зрения того, *что* каждый аттестуемый может продемонстрировать и *как* сделать, а не то, как он выглядит на фоне других аттестуемых.

Для целей итоговой аттестации в высшем образовании критериально-ориентированный подход представляется наиболее перспективным, поскольку он обеспечивает:

- 1) получение итоговой оценки уровня сформированности компетентности выпускника;
- 2) поэтапное оценивание уровня компетентности за счет определения порогового балла на каждой стадии;
- 3) проведение анализа качества подготовки специалистов по каждой образовательной программе, выпускники которой завершают ее освоение, для последующих корректирующих мероприятий по улучшению содержания или условий ее реализации.

Следует выделить несколько ключевых задач при таком подходе к проведению и интерпретации результатов оценивания. Первая задача – это разработка спецификации для оценочного инструментария (можно также встретить

другое значение этого процесса – раскрытие цели оценивания). *Спецификация* – это документ, в котором содержится информация о целях, задачах, плане и структуре оценочного инструментария (измерителя – в терминологии образовательных измерений), а также указаны основные требования к правилам проведения оценивания. Вторая задача – установление критериального балла или стандарта выполнения, применяемого для оптимального отбора между аттестованными или не аттестованными по результатам выполнения совокупности оценочных заданий. Решение данных задач предложено в третьей и четвертой главах диссертации.

Базовым понятием как в контрольно-оценочной деятельности, так и в аттестационных процедурах является оценочный инструментарий (измеритель), который включает в себя два компонента – само оценочное средство и шкалу, на которой фиксируются оценки измеряемого конструкта (баллы).

Под *оценочным средством* понимается совокупность заданий стандартизированной формы, обладающих необходимыми системообразующими статистическими характеристиками и обеспечивающих получение надежных, валидных и аутентичных результатов измеряемого конструкта. Сами задания могут включать в себя вопросы, задачи, упражнения, учебные или квазипрофессиональные проблемы и ситуации. Для целей итоговой аттестации следует использовать только такие задания, которые прошли апробацию на представительной выборке испытуемых и получили статистические оценки своих характеристик – *трудность* и *дифференцирующая способность*. Под *трудностью задания* понимается величина, показывающая долю испытуемых из выборки, ответивших правильно на данное задание. *Дифференцирующая способность задания* – статистическая характеристика, показывающая насколько данное задание способно разграничить испытуемых с высоким и низким уровнем подготовки.

Второй компонент оценочного инструментария – *шкала*, на которой откладываются оценки измеряемой переменной (или переменных – в случае многомерного измерения), является целью измерения. Эти оценки могут относиться как к количественному, так и к качественному уровню измерений. На

шкале, как правило, откладываются производные баллы, получаемые в результате преобразования первичных баллов. *Процедура шкалирования* направлена на расширение возможностей интерпретации и сопоставимости оценок, но не связана с повышением точности измерений. Методика шкалирования результатов многостадийного оценивания при итоговой аттестации представлена в четвертой главе диссертации.

В рамках данного исследования требуется рассмотрение еще одного ключевого понятия – *конструкт или переменная измерения*, понимаемый как гипотетический концепт, с помощью которого можно описать поведение человека. Термин «переменная» в образовательных измерениях используется не случайно, указывая на наличие ошибки измерения, величина которой изменяется и зависит от многих условий – качества измерителя (оценочного инструментария), условий его использования и проведения самой аттестационной процедуры. Как отмечают Л. Крокер и Дж. Алгина, существование таких конструктов не может быть подтверждено абсолютно, на практике это означает, что нужны наблюдение и оценка поведения человека [125]. Роль такого конструкта или переменной (переменных) для оценивания в условиях итоговой аттестации отводится совокупности освоенных обучающимся компетенций в соответствии с требованиями образовательной программы и стандарта. Эти переменные носят скрытый характер и недоступны для непосредственного наблюдения. В этой связи измерению подлежит то, что доступно для наблюдения, так называемые признаки проявления. Их определение должно происходить в два этапа. Первый этап – экспертный, когда аргументируется и обосновывается, почему то или иное наблюдаемое поведение будет свидетельствовать о наличии скрытого качества. Второй этап – статистический, предполагающий аналитическую работу после результатов измерений. Процесс представления связи между конструктом (переменными измерения) и наблюдаемым поведением (измеряемыми характеристиками переменных) называется *операционализацией*. Работа по операционализации компетенций является значимой не только на этапе оценивания достигнутых результатов в процессе обучения, но и на этапе

целеполагания и определения содержания обучения. Обобщенно этот процесс представлен на рис. 6.

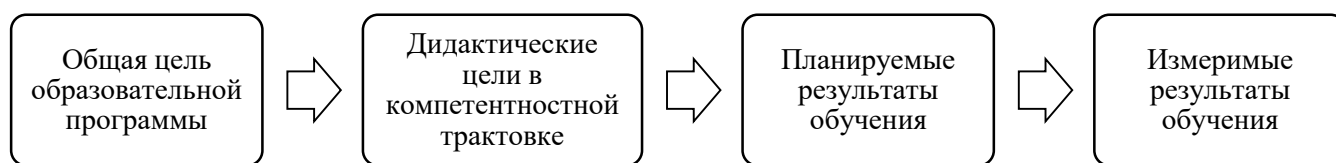


Рис. 6. Процесс операционализации

Интересуемое в рамках данной работы понятие «компетентность» впервые исследовал клинический психолог Роберт Уайт (R. White). Его статья 1959 года «Motivation reconsidered: The concept of competence» («Пересмотр мотивации: концепция компетентности»), опубликованная в журнале *Psychological Review*, считается самой известной работой в этом направлении. В ней Уайт утверждал, что компетентность является фундаментальным человеческим стремлением, которое мотивирует людей. Он предположил, что компетентность и действия, которые заставляют нас чувствовать себя компетентными, являются мотивирующими, поскольку компетентность рассматривается как путь к успеху. Позднее, следуя этому теоретическому подходу, в зарубежных странах стали разрабатывать оценочный инструментарий, позволяющий предсказывать эффективность в деятельности и жизненные успехи после окончания обучения с высокой прогностической валидностью [87].

Позже *компетентность* стала обсуждаться в связи со способностью преуспеть в своей роли в организации, т. е. как профессиональная компетентность. Исследователи выясняли, что делает человека компетентным в работе и как повысить его компетентность [37, 230, 292].

Подробный анализ компетентностного подхода для профессиональной управленческой деятельности приводится в трудах Л. Спенсера и С. Спенсера [230]. Согласно их определению, *компетенции* рассматриваются как базовые качества индивидуума, имеющие причинные связи с эффективностью его деятельности после окончания обучения. Таким образом, уже на понятийном

уровне в этом определении закладываются три ключевые идеи компетентностного подхода, одна из которых связана с отсроченным характером проявления компетенций после окончания обучения, вторая – с причинным характером связей с эффективностью профессиональной деятельности, а третья – с базовым характером компетенций, позволяющим рассматривать их как глубоко лежащие устойчивые поведенческие характеристики человеческой личности, прогнозирующие эффективность профессиональной деятельности на основе освоенных знаний, умений и навыков.

Следуя логике Спенсеров, можно утверждать, что компетенции имеют металатентную природу и потому крайне трудно поддаются любым попыткам формирования и оценивания [230]. По их мнению, компетенции предопределяют мыследеятельность и практические действия человека в разных ситуациях. Также компетенции связаны с мотивацией, физиологией, психологией личности, ее самоустановками. Поэтому все это целесообразно представить в виде окружностей, как на рисунке 7.

Для лучшего понимания того, как происходит формирование и оценивание компетенций следует рассматривать движение с внутреннего круга на внешний уровень и обратно. В этой визуализации и раскрывается сама идея проявления компетенций при выполнении определенных заданий, когда можно наблюдать за результатами. И, возвращаясь во внутренний круг, получать надежные и валидные оценки об уровне сформированности компетенций. Можно видеть, что переходы с одного уровня на другой будут подвержены влиянию различных факторов, смещающих оценку и увеличивая тем самым ошибку. Вот почему так важно использовать такой оценочный инструментарий, который может гарантировать измерение заявленных переменных (наблюдаемых признаков), а по получаемым оценкам обоснованно делать выводы об успешной профессиональной деятельности выпускников.

К настоящему времени в отечественной дидактике сложилось относительно общее понимание терминов «компетентность» и «компетенция», хотя в самом начале их нормативного введения было порождено много определений,

классификаций, что приводило к еще большему протесту большей части педагогических коллективов в высшей школе, разрабатывающих необходимую документацию по ФГОС (рабочие программы, программы формирования компетенций, паспорта компетенций, фонды оценочных средств и т. п.). Мы придерживаемся той позиции, что *компетенции* являются операционализированными составными характеристиками интегрального понятия «*компетентность*», указывая статус при классификации аттестуемых выпускников на основе оценок уровня их сформированности [158]. Тогда *компетентность* – это интегративное понятие, которое распадается на спектр дискретных компетенций. Оно характеризует в целом способности обучающихся, формируется на базе освоения ряда связанных знаний и умений и позволяет эффективно выполнять профессиональную деятельность по отношению к стандартам, существующим в ней [87]. Таким образом, вслед за В. И. Звонниковым и М. Б. Челышковой примем, что *компетентность* – это интегральная характеристика потенциальных способностей обучающегося, которая распадается на спектр формируемых в процессе обучения, проявляющихся в профессиональной деятельности и оцениваемых компетенций [87].

Опираясь на исследования указанных авторов, в данной работе будем рассматривать *компетенцию* как способность применять знания, умения, навыки и личностные качества для успешной деятельности в различных профессиональных либо жизненных ситуациях, а *компетентность* – как уровень владения выпускником совокупностью компетенций, отражающих способность к применению знаний, умений, способов деятельности (практического опыта) и сформированных на их основе компетенций для успешного выполнения трудовых функций в определенной профессиональной области [87].

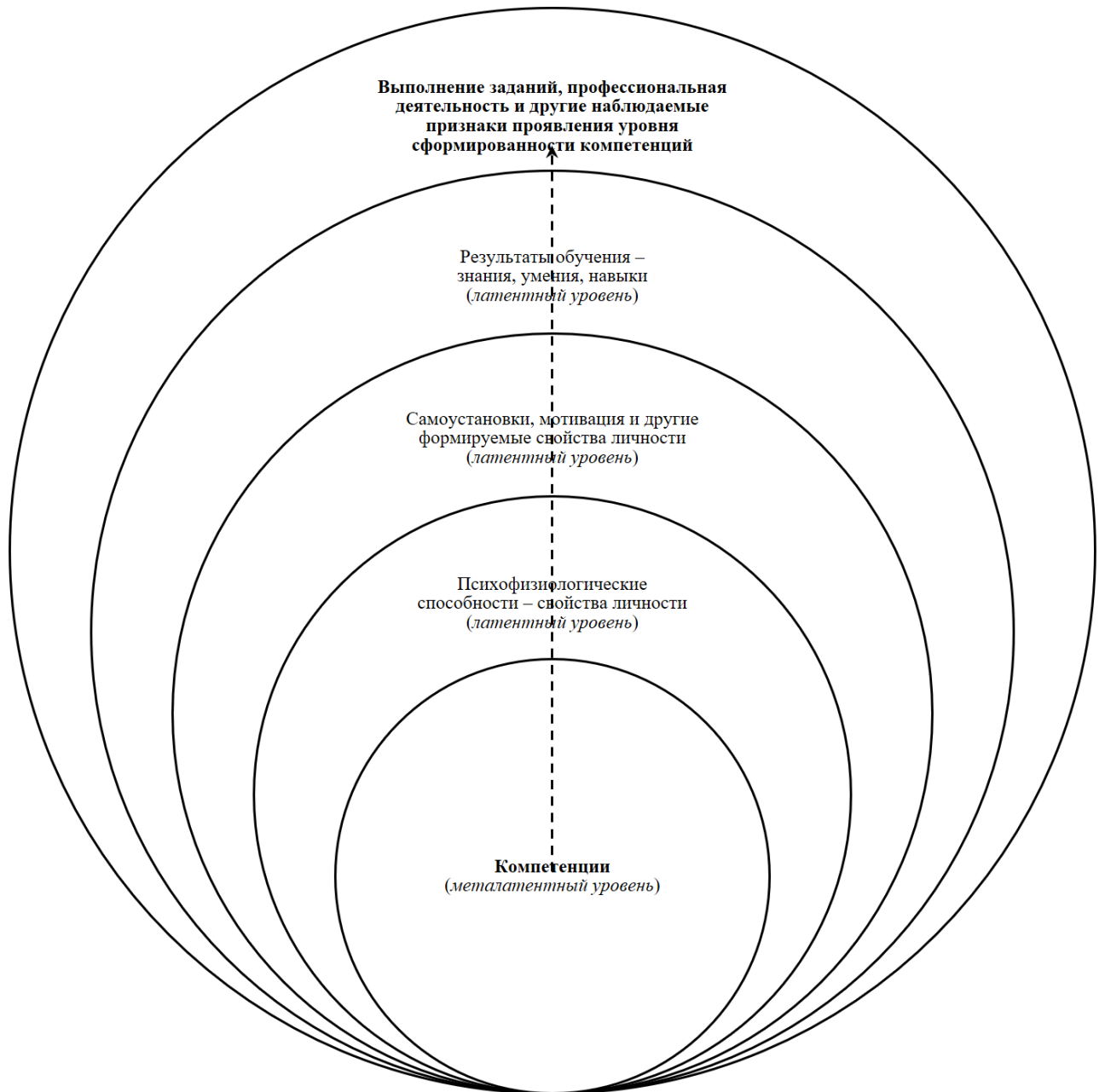
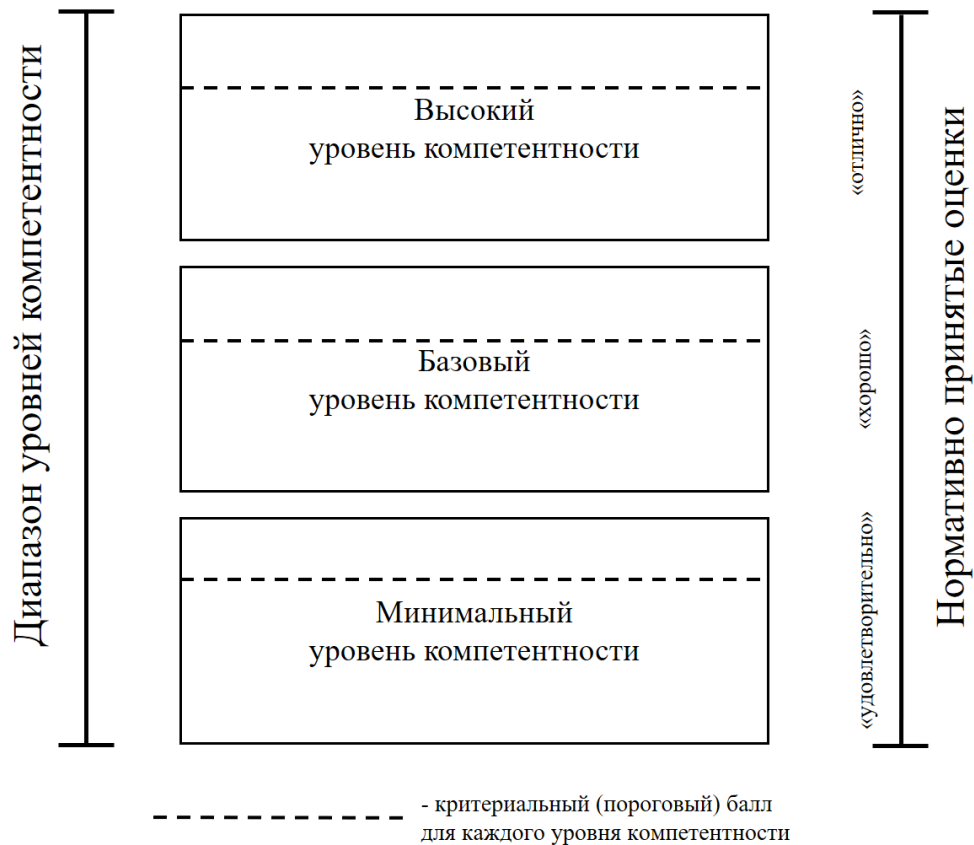


Рис. 7. Визуализация взаимосвязи характеристик обучающегося

В силу того, что компетенции являются способностями, то их не принято оценивать в виде определенного числа баллов на шкале. В профессиональных службах оценивания и исследованиях по данному вопросу отмечается, что необходимо определить и ввести диапазоны компетентности, внутри которых будут располагаться результаты аттестуемых, полученные при оценивании их уровня компетентности. Определяя количество диапазонов, следует опираться не только на строго дихотомический подход (компетентен / некомпетентен), но также



учитывать сложившиеся практики и нормативные требования в системе высшего образования. В данном исследовании предлагается ввести три диапазона уровня компетентности, выделив минимальный, базовый и высокий уровни. Это согласуется с принятой в отечественной системе образования и установленной нормативным документом пятибалльной шкалой оценивания. В действительности эта шкала является четырехбалльной, поскольку предусматривает по результатам аттестационных испытаний выставление одной из четырех оценок – «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно». В случае получения оценки «неудовлетворительно», выпускник считается не прошедшим итоговую аттестацию, а следовательно – не освоившим образовательную программу. Таким образом, предлагаемые три уровня компетентности выпускника (минимальный, базовый, высокий) будут соотноситься с принятой градацией оценок (удовлетворительно, хорошо, отлично). Можно считать, что предлагаемая вертикальная шкала уровня компетентности легко интерпретируется, задается в понятных для аттестуемых студентов и выпускников категориях, исключая абстрактные оценочные прилагательные (например, превосходный, хороший, плохой и др.), и дифференцированно соотносится с традиционной системой оценивания, принятой в отечественной высшей школе (рис. 8). Не упрощая процедуру шкалирования результатов оценивания, которая подробно описана в четвертой главе диссертации, отметим, что выделению уровней сформированности компетентности предшествует работа по описанию признаков проявления подлежащих измерению компетенций. При этом, когда число компетенций достаточно велико, их объединяют в кластер, детализируя уровневые признаки проявления. Эта технология представлена в третьей главе диссертации.



*Рис. 8.* Вертикальная шкала для диапазона уровней компетентности, соотнесенная с нормативно принятыми оценками

На рис. 8 показаны следующие уровни компетентности: минимальный, базовый и высокий. Они соотнесены с оценками «удовлетворительно», «хорошо» и «отлично» соответственно. Продвижение по этой шкале от минимального к высокому уровню компетентности предполагает выполнение аттестуемым оценочных заданий соответствующей трудности: чем выше уровень компетентности, тем труднее совокупность заданий. Для реализации такой аттестационной процедуры в исследовании вводится понятие многостадийное оценивание, которое представляет собой определенную последовательность взаимосвязанных и взаимообусловленных научно обоснованных этапов и процедур создания, предъявления и оценки результатов выполнения оценочных заданий аттестационного инструментария. Такое оценивание, проводимое в несколько

этапов (стадий), отвечает требованиям компетентного подхода, а также профессиональных стандартов, в которых трудовые функции специалистов описаны через необходимые знания, умения и трудовые действия.

### **Выводы по первой главе**

Становление и развитие подходов к организации и содержанию контрольно-оценочной деятельности в отечественной системе образования менялись по мере развития самого общества, укладов, технологий, а также под влиянием различных политических, социальных и экономических факторов. Эта деятельность всегда привлекает к себе гораздо больший общественный интерес, чем методы и содержание обучения, поскольку напрямую касается каждого обучающегося, его семьи и общества в целом, характеризуя состояние системы образования. Сведение результатов оценивания и аттестации только к получению высоких оценок, которые затем некорректным образом преобразуются и используются в разных социальных и политических контекстах, приводит к искажениям реальной действительности. Этот путь прошла и отечественная система образования, отменяя и вновь вводя отметки в школе. Участие России в международных сравнительных исследованиях качества образования привело к необходимости пересмотра многих аспектов в дидактике, включая контроль и оценку.

Высшая школа отличалась большей свободой при определении методов контроля и критериев выставления оценок, что зачастую приводило к еще большей субъективности и несопоставимости, когда ставились вопросы о качестве подготовки специалистов. Введение государственных образовательных стандартов в системе высшего образования, а также новые социально-экономические отношения в обществе привели в начале XXI века к обсуждению назревших вопросов о качестве образования и соответствия молодых специалистов требованиям рынка труда. Позднее, во втором десятилетии этого века, в целях сближения двух сфер – образования и труда – начался процесс разработки

внедрения профессиональных стандартов, которые вузы должны учитывать при проектировании и реализации образовательных программ. Новый этап взаимодействия представителей работодателей и высшей школы привел к появлению новых образовательных стандартов третьего поколения и развитию системы независимой оценки квалификаций.

Компетентностный подход, положенный в основу третьего поколения государственных образовательных стандартов для подготовки специалистов, задал необходимость пересмотра основных категорий – компетентности как образовательного результата освоения образовательной программы, содержание (чему обучать?), технологии (как?) и способов получения обратной связи и оценки этих результатов. Изменился язык описания результатов обучения в компетентностной трактовке. Это не только привычные для отечественной дидактики знания, умения и навыки обучающихся, но и овладение практическим опытом для эффективной профессиональной деятельности. При этом уровень сформированных образовательных результатов (компетенций) студента и компетентности выпускника должен определяться и оцениваться не в сравнении его с другими студентами и выпускниками, не ориентируясь на гипотетический средний уровень, а соотноситься с определенной содержательной областью. Такой подход в оценивании называется критериально-ориентированным, и именно он берется за основу при проектировании аттестационных процедур высокой значимости и разработке современного оценочного инструментария для этого.

Новое понимание результатов освоения образовательных программ не отрицает базовой триады – знаний, умений, владений, а напротив на их основе и с учетом психофизиологических, личностных, ценностных установок и особенностей, они интегрируются в понятие *компетентность*, обуславливающее эффективную профессиональную деятельность. Этот постулат необходимо учитывать, отвечая на основной вопрос при итоговой аттестации выпускников в высшем образовании: каким образом должно быть выстроено оценивание, чтобы одновременно можно было измерять большое количество заявленных в образовательных программах компетенций, при этом обеспечивая независимость

этой процедуры, а также сопоставимость получаемых оценок? Для ответа на этот вопрос в данной работе предлагается переход от оценивания устных или письменных ответов аттестуемых студентов на 2–3 случайно полученных вопроса экзаменационных билетов к многостадийному оцениванию, предусматривающему применение релевантного оценочного инструментария на нескольких этапах аттестационной процедуры. Многостадийное оценивание способствует повышению надежности и валидности результатов при измерении таких сложных конструктов как компетенции за счет применения различных оценочных средств на разных уровнях измерений. Такой оценочный инструментарий и оценивание в несколько этапов должны основываться на определенной методологии, а ее реализация на практике проводится по определенной технологии, которые представлены в следующих главах диссертации.

## **ГЛАВА 2. МЕТОДОЛОГИЯ МНОГОСТАДИЙНОГО ОЦЕНИВАНИЯ УРОВНЯ КОМПЕТЕНТНОСТИ ВЫПУСКНИКОВ ПРИ ИТОГОВОЙ АТТЕСТАЦИИ В ВЫСШЕМ ОБРАЗОВАНИИ**

### **2.1. Цели и функции многостадийного оценивания при итоговой аттестации**

Определяя цели многостадийного оценивания, следует учитывать контекст, в котором оно изучается, а именно – итоговую аттестацию выпускников, завершивших освоение основной профессиональной образовательной программы. В этой связи на многостадийное оценивание будут распространяться общие и специальные цели аттестации, связанные со всеми присущими особенностями содержательного, организационно-педагогического и управленческого характера.

К числу общих целей, имеющих отношение и к системе подготовке специалистов с высшим образованием, и к сфере труда, и к научно-технологическому развитию, следует отнести:

- повышение уровня и качества подготовки специалистов;
- реализация государственной политики в области подготовки кадров для различных отраслей экономики;
- соблюдение установленных регулятором на нормативном уровне требований к качеству высшего образования, отраженных в федеральных государственных образовательных стандартах или самостоятельно установленных образовательных стандартах;
- сопоставимость дипломов о высшем образовании и квалификаций выпускников и молодых специалистов, полученных в различных образовательных организациях по одним и тем же или родственным направлениям подготовки и профилям;
- обеспечение преемственности требований к качеству высшего образования на всех его уровнях – бакалавриата, специалитета, магистратуры, аспирантуры;

- снижение доли субъективизма и связанной с ним несправедливости традиционной аттестационной процедуры;

- предоставление всем субъектам и пользователям системы образования и сферы труда объективной информации о реальном уровне качества подготовки специалистов и компетентности выпускников;

- совершенствование аттестационных процедур, в т. ч. итоговой аттестации, и системы оценивания на основе образовательных измерений.

Современная ситуация в высшем образовании характеризуется совершенствованием системы подготовки и переподготовки кадров по актуальным для экономики и общества научно-технологическим направлениям, стандартизацией образовательных результатов (универсальные, общепрофессиональные и профессиональные компетенции) с одновременной вариативностью профилей обучения, разнообразием учебных планов и способов решения поставленных дидактических задач, цифровой трансформацией, применением дистанционных образовательных технологий, искусственного интеллекта, нейросетей и больших данных [51, 132, 190, 193, 244 и др.]. Это позволяет сформулировать к названным выше общим целям ряд специальных. Перечисленные характеристики проистекают из стратегических и программных документов и концепций – Указа Президента РФ от 28.02.2024 г. № 145 «О Стратегии научно-технологического развития Российской Федерации», постановления Правительства РФ от 11.10.2023 г. № 1678 «Об утверждении Правил применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ», распоряжения Правительства РФ от 24.06.2022 г. № 1688-р «Об утверждении Концепции подготовки педагогических кадров для системы образования на период до 2030 года» и др.

На многостадийное оценивание распространяются следующие специальные цели:

- достижение объективности и независимости в аттестационных процедурах;

- повышение обоснованности решений при итоговой аттестации;
- обеспечение эффективности аттестационных процедур за счет методов оптимизации количества оценочных заданий и их трудности, что приводит к точности получаемых оценок об уровне компетентности аттестуемых выпускников;
- создание ситуации успеха для каждого аттестуемого выпускника благодаря дифференциации и индивидуализации процедуры оценивания;
- обеспечение благоприятной психологической атмосферы для прохождения аттестационной процедуры высокой значимости, снижающей невербальное психологическое давление экзаменаторов.

В числе функций, которые традиционно связывают с контрольно-оценочной деятельностью, следует выделить только те, которые соответствуют целям многостадийного оценивания именно при итоговой аттестации выпускников в высшем образовании. Присущие педагогическому контролю и оцениванию диагностическая, обучающая, развивающая, корректирующая функции могут быть реализованы в условиях итоговой аттестации, если речь идет о процессе подготовки или самоподготовки к ее успешному прохождению. Действительно, для выявления пробелов, саморегуляции, содержательной и психологической готовности к итоговой аттестации реализация указанных функций будет этому способствовать.

В случае итогового оценивания, по результатам которого делаются экспертные заключения об уровне сформированности компетентности каждого выпускника, его готовности к профессиональной деятельности, а также принимаются решения о выставлении оценок сдачи аттестационных испытаний, присвоении квалификации и выдаче документа о соответствующем образовании, выделяются следующие функции – собственно *оценочная*, *информационная*, *сравнительная* и *прогностическая*. Следует отметить, что содержательно эти функции педагогического контроля и оценивания на современном этапе развития науки и технологий трактуются в залоге новых открывшихся возможностей работы с получаемыми данными.



Полноценная реализация оценочной функции в итоговой аттестации способствует справедливости и доверию к получаемым оценкам, на снижение которых влияет субъективность тех, кто проводит оценивание: педагогов, работодателей, административных работников. Именно развитие научной области образовательных измерений, реализация на практике положений теории тестов, и, в частности, современной теории тестов – IRT, обоснование которой приводится ниже в данной главе, обеспечивают достижение специфической объективности и сопоставимости оценок по результатам прохождения выпускниками аттестационных испытаний. Специфическая, или инвариантная, объективность – это одна из трех характеристик объективности вместе с процедурной и классической, которые традиционно рассматриваются, когда речь идет об образовательных измерениях [83, 266, 279, 298, 299 и другие]. Подробнее объективность образовательных измерений и связанная с ним надежность результатов представлены в третьей главе. Применительно к оценочной функции многостадийного оценивания следует отметить, что достижение эффекта инвариантности при оценивании уровня подготовленности аттестуемых и трудности заданий измерителя связано с применением IRT. Согласно основным положениям современной теории тестов, инвариантная объективность характеризуется тем, что на оценку каждого аттестуемого не влияют оценки других аттестуемых и особенности оценочного инструментария, проявляющиеся в подборе по трудности заданий [88].

Выполнение оценочной функции современного многостадийного оценивания обладает определенными особенностями, поскольку оценка по его результатам:

- не сравнивает аттестуемых друг с другом за счет дифференциации, переходящей в индивидуализацию, процедур предъявления заданий, окончания аттестационного испытания и предъявления результатов;

- предоставляет каждому выпускнику возможность видеть продвижение по уровням сформированности своей компетентности (от минимального к базовому до высокого уровня сформированности компетентности) благодаря поэтапному

процессу выполнения оценочных заданий и возможности в этой связи качественной интерпретации получаемой оценки;

– носит объективный и оперативный характер за счет реализации самой процедуры на основе алгоритмов ИРТ и предъявления результатов выполнения заданий на каждом этапе оценивания.

Реализация информационной функции многостадийного оценивания связана со сбором данных по каждой стадии оценивания в процессе аттестационного испытания (сдачи государственного экзамена, защиты выпускной квалификационной работы) выпускниками для анализа результатов и их корректной интерпретации. Анализ данных прохождения аттестационных испытаний необходим, в первую очередь, для сопоставления со стандартами образования – соответствие результатов освоения выпускниками образовательных программ требованиям ФГОС ВО. Кроме того, для обоснованного принятия административно-управленческих решений, например, таких, как корректировка содержания образовательной программы, изменения учебного плана, отдельных рабочих программ или методического обеспечения дисциплин (учебных модулей), практик и пр., эти данные должны рассматриваться в совокупности с информацией, получаемой из дополнительных сведений и источников. К таковым могут относиться – баллы ЕГЭ, на основании которых были зачислены абитуриенты для освоения образовательной программы высшего образования; квалификационный уровень, качественный и количественный состав научно-педагогических работников, участвующих в реализации образовательной программы; материально-техническое обеспечение образовательной организации; трудоустройство выпускников прошлых лет и их становление и развитие в той профессиональной деятельности, к которой они готовились. Другими словами, в целях мониторинга качества высшего образования и подготовки специалистов реализация многостадийного оценивания при итоговой аттестации создает информационную основу для анализа данных с использованием различных методов, пользователями результатов которого могут выступать все участники образовательного процесса, а также заинтересованные стороны.

Сравнительная функция многостадийного оценивания проявляется, в первую очередь, на уровнях управления высшим образованием по одинаковым или сходным профилям образовательных программ. Итоговая аттестация в высшем образовании не должна предусматривать сравнение выпускников друг с другом, но должна обеспечивать сопоставление результатов освоения образовательных программ путем корректного анализа входных и выходных данных об уровне подготовки обучающихся. В качестве выходных данных при реализации сравнительной функции могут использоваться результаты корректно проведенных образовательных измерений, начиная с использования стандартизированного оценочного инструментария, репрезентативной выборки выпускников по одинаковым направлениям подготовки и образовательным программам, обеспечения надежности и валидности и заканчивая шкалированием и интерпретацией результатов, дополненных ответами на профессионально разработанные анкеты, которые выявляют существенные факторы в подготовке специалистов. Сравнительная функция часто не принимается ни теми, кто учит, ни теми, кого учат, что только отчасти справедливо, хотя сами критики не учитывают ряд научных оснований. В действительности сравнивать допустимо только то, что измеряется одним и тем же инструментарием (измерителем), а результаты измерения переносятся на одну и ту же шкалу. Сравнить количество выпускников, получивших различные оценки по привычной 5-балльной шкале в двух или более различных вузах, которые реализовывали образовательные программы по одинаковым направлениям подготовки, будет некорректно. Это обусловлено и характером порядковой шкалы качественного уровня измерений, и отсутствием единых критериев оценивания результатов выполнения аттестационных испытаний, а также субъективностью экспертов – членов экзаменационных комиссий в этих образовательных организациях. В целом, для выполнения сравнительной функции многостадийного оценивания при итоговой аттестации должны выбираться такие группы объектов, которые гомогенны по своим характеристикам. Начавшийся в конце первого десятилетия XXI века в системе высшего образования проект ФЭПО (федеральный интернет-экзамен в

сфере профессионального образования) можно отнести к случаю выполнения информационной и сравнительной функций оценивания в целом, но подчеркнем, что там не учитывались факторы, оказывающие влияние на смещение результатов сравнения. К числу таких факторов можно отнести как территориальные и социально-демографические особенности, так и содержательные аспекты формирования заявленных для оценивания результатов обучения по конкретным учебным дисциплинам и модулям образовательных программ.

Роль прогностической функции многостадийного оценивания сегодня значительно возрастает. Это обусловлено не только новыми возможностями информационного обеспечения оперативной обработки больших данных и развития вероятностных моделей предсказания наметившихся позитивных или негативных тенденций в образовании. По результатам выполнения оценочных заданий разного уровня и формы по определенным алгоритмам и с соблюдением всех организационно-педагогических условий аттестационных процедур можно с высокой долей вероятности определить и дальнейшее развитие образовательных программ, их востребованность абитуриентами, а также успех в профессиональной деятельности выпускников при их трудоустройстве по специальности. На основе таких прогнозных значений, которые проводятся математико-статистическими методами, можно принимать тактические и стратегические решения по многим направлениям развития высшего образования, его содержательных и организационных аспектов.

Описанные функции многостадийного оценивания при итоговой аттестации, как правило, реализуются в сочетании друг с другом при определенной ведущей роли одной из них. Это зависит от того, какая из целей многостадийного оценивания является приоритетной и каков контекст рассмотрения результатов итоговой аттестации.

## 2.2. Принципы многостадийного оценивания

В широком смысле принципы выступают нормативами, предписывающими характер той или иной деятельности, а также являются центральными понятиями, распространяющимися на все рассматриваемые процессы [183]. Введение в процедуры итоговой аттестации в высшем образовании строгих формализованных методов оценивания преимущественно обусловлено тем, что со стороны государства и профессионального сообщества, представляющего рынок труда, выдвигается новый «социальный заказ» к подготовке современных специалистов. Поэтому научную организацию многостадийного оценивания следует выстраивать на ряде основополагающих принципов, охватывающих весь процесс оценивания: от определения компетенций и компетентности как подлежащих измерению конструктов до анализа данных и интерпретации результатов с целью принятия обоснованных классификационных решений. Соблюдение таких принципов будет способствовать повышению эффективности самих аттестационных процедур (испытаний).

Помимо общих, ставших уже классическими, принципов оценивания при итоговой аттестации, которые распространяются и на многостадийное оценивание, таких как *объективность, независимость, научность, релевантность, достоверность, справедливость и эффективность*, предлагается дополнить данный перечень рядом специальных принципов. К их числу относятся – *дифференциация, индивидуализация, последовательность, критериальность, открытость, сочетание количественного и качественного уровней образовательных измерений, кумулятивность*.

Первые два принципа – объективности и независимости – являются нормативно установленными в статье 59 «Итоговая аттестация» 273-ФЗ «Об образовании в Российской Федерации» [188].

*Принцип объективности* в многостадийном оценивании, строящемся на образовательных измерениях, связан с неизбежным существованием ошибки измерения. Согласно классическому пониманию, величина ошибки измерения

определяется как разность между наблюдаемым и истинным баллом аттестуемого. Тем самым этот принцип предполагает минимизацию такой ошибки, которая может носить характер случайной или систематической. К появлению ошибок систематического характера приводят некачественно разработанный оценочный инструментарий и задания в нем, процедуры его применения, нарушения методик сбора данных и их интерпретации. Случайные ошибки измерения возникают в результате поведения аттестуемых или экзаменаторов. Различные действия как тех, так и других во время аттестационной процедуры будут приводить к снижению объективности, а значит, снижению надежности результатов и повышению величины случайной ошибки. Ее минимизация может достигаться различными методами – от полного исключения экзаменаторов из процесса и перевод аттестационной процедуры в компьютерный формат с автоматизированной обработкой результатов до применения методов коллегиальной (экспертной) оценки, строгого соблюдения разработанных инструкций по оцениванию результатов выполнения заданий и других. В целом число факторов, которые влияют на увеличение ошибки измерения, достаточно велико даже в тех случаях, когда применяются специальные математические методы и модели измерения IRT.

*Принцип независимости* в условиях проведения итоговой аттестации рассматривается, прежде всего, с позиции, кто оценивает, делает выводы и принимает решения. Речь идет о подходах к формированию аттестационной комиссии, которую возглавляет независимый по отношению к образовательной организации председатель. Его кандидатура, как правило, утверждается учредителем по представлению самой организации. Также в состав аттестационной комиссии должны входить представители работодателей, заинтересованные в беспристрастной оценке уровня компетентности выпускников и отборе их для профессиональной деятельности. Другими словами, независимость обеспечивается тем, что уровень компетентности выпускника оценивают не те, кто его обучал и формировал компетенции, а внешние по отношению к аттестуемому и образовательной организации эксперты.

*Принцип научности* многостадийного оценивания предписывает организацию и проведение его строго в соответствии с выбранной основой – образовательными измерениями. Эти основы предусматривают выполнение ряда обязательных требований разработки и применения оценочного инструментария для аттестационной процедуры. Согласно принятым правилам, выделяют пять таких этапов [87]:

1) этап целеполагания, заключающийся в выборе подлежащего измерению конструкта, определения выборки аттестуемых и содержательной области, описывающей этот конструкт;

2) этап конструирования заданий инструментария и обоснование их качества;

3) этап апробации всех разработанных заданий инструментария с последующей математико-статистической обработкой данных и корректировкой заданий;

4) этап применения инструментария для оценивания аттестуемых и шкалирование результатов;

5) этап интерпретации полученных результатов выполнения заданий экзаменуемыми в соответствии с установленными критериальными баллами (стандарты прохождения аттестационных испытаний) и принятие классификационного решения в отношении каждого аттестуемого выпускника (аттестован – не аттестован).

*Принцип релевантности* нацелен на согласованность получаемых на основе многостадийного оценивания результатов об уровне сформированности компетентности выпускников с заявленными целями реализуемой образовательной программы. Интерпретация результатов аттестации согласно данному принципу рассматривается по отношению к различным уровням принятия обоснованных решений. Так, если итоговая аттестация проводится по единой методологии и технологии для всех образовательных программ из общей укрупненной группы специальностей и направлений или общего направления подготовки, то данный принцип нацеливает на решение задач управления качеством подготовки специалистов по определенным профилям. Как отмечалось

выше, это способствует созданию единого образовательного пространства и выработке общего понимания целей и результатов высшего образования. Принцип релевантности означает сквозную согласованность с общими направлениями развития образования, устанавливаемыми нормами, правилами на уровне государства, и с потребностями конкретных университетов и структурных подразделений, которые отвечают за содержание подготовки, формирование единых образовательных результатов (универсальных и общепрофессиональных компетенций) и вырабатывают синхронизированные с работодателями формулировки профессиональных компетенций в связке с профессиональными стандартами.

*Принцип достоверности* регламентирует получение неискаженной информации о результатах итоговой аттестации выпускников и уровня сформированности их компетентности. Этим результатам должны доверять не только те, кто непосредственно участвует в аттестационных процедурах, но и другие пользователи, включая профессиональные и общественные объединения, органы управления и надзора за образованием. Достоверность связана с выбором таких технологий, а вместе с ними и методов, инструментария оценивания, которые не вызывают противоречивых мнений и позволяют получать объективные оценки.

*Принцип справедливости* в оценивании один из достаточно сложно реализуемых. Часто, хотя и неверно, справедливость отождествляют с принципом равенства обучающихся на получение беспристрастной и непредвзятой оценки. Данный принцип предписывает получение по результатам выполнения аттестационного испытания оценки, которая не случайным образом на условиях везения или невезения будет получена аттестуемым выпускником, а напротив – соответствовать всем тем усилиям при освоении образовательной программы, которые прикладывал студент на протяжении всего периода обучения. В такой ситуации многостадийное оценивание будет в полной мере содействовать реализации данного принципа.

*Принцип эффективности* регламентирует выбор таких технологий многостадийного оценивания на основе образовательных измерений, которые



обеспечивают полноценную реализацию всех функций при итоговой аттестации и при этом оптимизируют затрачиваемые ресурсы (временные, финансовые, трудовые).

В настоящее время указанные принципы, часть которых в полной мере, а другая частично реализуется в аттестационных процедурах, следует дополнить рядом специальных, на которых базируется многостадийное оценивание, повышая его состоятельность в оценивании высокой значимости.

*Принцип дифференциации* относится как к самим аттестуемым, так и к установлению и определению уровней компетентности выпускников, завершивших освоение образовательной программы. В случае дифференциации аттестуемых предписывается их разделение в начале аттестационного испытания на первой стадии оценивания, когда каждому из них предъявляется отдельный набор заданий, подобранный в предписанном интервале трудности и оптимальный по уровню подготовленности каждого испытуемого. Компетентность как интегративное понятие, включающее в себя совокупность необходимых к освоению обучающимся универсальных, общепрофессиональных и профессиональных компетенций, также предусматривает ее сформированность у каждого выпускника на определенном уровне. Поэтому такая дифференциация должна выполняться и быть заранее известна всем участникам итоговой аттестации.

*Принцип индивидуализации* тесно связан с принципом дифференциации и предполагает, что процесс оценивания, проходящий в несколько стадий, за счет разделения всех аттестуемых по уровню их подготовленности реализуется в дальнейшем в индивидуальном режиме. Это означает, что выпускники выполняют задания, которые предъявляются им с учетом правильности ответов на предыдущие задания. Другими словами, первый этап оценивания – это компьютерный формат предъявления заданий по специальным алгоритмам, учитывающим успешность их выполнения и уровень подготовленности каждого аттестуемого. Оценки, полученные в результате выполнения испытаний первого этапа, определяют переход к следующему этапу оценивания, уровень трудности

которого выше и соответствует другому уровню компетентности. И при этом в качестве обязательного условия учитывается готовность выпускника переходить к следующей стадии оценивания.

*Принцип последовательности* предполагает поэтапный переход аттестуемого в режиме многостадийного оценивания – от выполнения обязательных заданий, соответствующих минимальному уровню компетентности, до высокого уровня компетентности, на котором необходимо доказать степень своей готовности к решению профессиональных задач.

*Принцип критериальности* означает обязательное наличие стандарта выполнения (порогового балла) или критериального балла (критерия) для каждого этапа оценивания, позволяющего отнести аттестуемого выпускника к одной из двух категорий – аттестован или не аттестован – для каждого уровня компетентности. Этот принцип способствует преодолению пороговых баллов каждым аттестуемым благодаря четко определенным границам аттестации для каждого этапа оценивания.

*Принцип открытости* должен пониматься всеми участниками аттестационных процедур однозначно. Он предполагает право каждого выпускника заранее ознакомиться с правилами проведения многостадийного оценивания, пройти пробное оценивание, знать спецификацию (план) оценочного инструментария для аттестационного испытания, критерии определения уровня компетентности и правил перевода (шкалирование) получаемых оценок в традиционную пятибалльную шкалу отметок.

*Принцип сочетания количественного и качественного уровней образовательных измерений* предписывает при организации многостадийного оценивания выбор определенного дизайна и схемы бипарадигмальных измерений, корректной обработки получаемых данных при использовании различного инструментария на разных уровнях измерений для их последующей интерпретации и принятии обоснованных классификационных решений по каждому аттестуемому выпускнику.

*Принцип кумулятивности* в многостадийном оценивании предполагает накопление информации о выполнении аттестуемым оценочных заданий на каждом этапе, соответствующем определенному уровню компетентности, и ее вклад в общую итоговую оценку по результатам прохождения всего аттестационного испытания. Данный принцип содействует обоснованности решений при аттестации в условиях постепенного нарастания трудности каждой стадии оценивания.

Соблюдение указанных принципов при организации многостадийного оценивания на основе образовательных измерений при итоговой аттестации существенно изменяет форму и содержание аттестационных испытаний, минимизируя основания для получения неправдоподобных и противоречивых результатов об уровне компетентности выпускников.

### **2.3. Бипарадигмальный подход и теория тестов как методологические основания многостадийного оценивания уровня компетентности выпускников при итоговой аттестации**

Выбор бипарадигмального подхода и теории тестов в качестве методологических оснований многостадийного оценивания компетентности у выпускников, завершающих освоение образовательных программ высшего образования, обусловлен их доказанной научной состоятельностью в ряде работ отечественных и зарубежных исследователей [53, 59, 86, 155, 173, 219, 261, 272, 276, 291, 293, 318, 336 и др.]. Вместе с тем сама возможность обоснованного и эффективного сочетания различных уровней образовательных измерений в аттестационных процедурах высокой значимости и широкое применение в практике общего образования теории тестов в контексте итоговой аттестации именно в высшем образовании рассматривались до этого фрагментарным образом и не в единстве. Данное исследование призвано устранить этот пробел, обогащая и развивая методологический аппарат педагогической науки, теорию отечественной

дидактики профессионального образования и контрольно-оценочную деятельность в высшем образовании.

Сочетание количественного и качественного уровней в образовательных измерениях становится реализуемым на практике благодаря самой постановке задачи оценивания результатов обучения в компетентностной трактовке. В соответствии с данным выше определением многостадийного оценивания, на каждой стадии аттестационного испытания используются различные виды измерителей (оценочных средств), которые соответствуют либо количественному, либо качественному уровню образовательных измерений. Для итоговой аттестации определено три уровня компетентности – минимальный, базовый и высокий, что согласуется с нормативным требованием выставления по результатам прохождения аттестационного испытания оценок «удовлетворительно», «хорошо» и «отлично» соответственно. Первый и второй этапы оценивания проводятся на количественном уровне измерений, третий этап – на качественном уровне. Такая последовательность использования различных измерителей, во-первых, не противоречит положениям теории измерений, когда данные количественной шкалы преобразуют в качественные оценочные суждения, и, во-вторых, выставление аттестуемым выпускникам оценок («неудовлетворительно», «удовлетворительно», «хорошо», «отлично») соответствует порядковой шкале качественного уровня измерений.

Идея отказа от монометода оценивания и переход к использованию совокупности методов и средств измерения в отечественной педагогической науке обсуждалась многократно, но доказательно была обоснована в исследованиях В. И. Звонникова для целей педагогических измерений в управлении качеством общего образования на основе анализа результатов ЕГЭ [88]. Именно в его работах, которые включали обоснование перехода в XXI веке от позитивизма к прагматизму в образовательных измерениях, был введен в научный оборот и обоснован бипарадигмальный подход. Приоритет при таком подходе «следует отдавать не единственному источнику получения информации, результатам стандартизированного оценивания, например, тестирования, а нескольким

источникам данных» [83, с. 129]. Суть заключается в понимании различий между количественными (интервальная шкала и шкала отношений) и качественными (номинальная шкала и порядковая шкала) результатами оценивания. К сожалению, несмотря на введение многих инноваций в контрольно-оценочной деятельности в образовании – от компьютерного оценивания и широкого использования разнообразных тестов до рейтингования обучающихся и образовательных организаций – нарушаются корректные правила и математические операции в работе с данными [54, 254]. Такое непонимание проистекает ввиду отсутствия культуры образовательных измерений, устоявшихся норм и правил работы с результатами оценочных процедур и принятой профессиональным сообществом методологии, теории и технологии.

Бипарадигмальный подход к образовательным измерениям в контрольно-оценочных процедурах выступает компромиссным решением для сторонников строгих количественных методов, стремящихся к исключению любых экспертных и субъективных оценок, и представителей качественных методов оценивания, при которых допускается нестрогая трактовка подлежащих измерению конструктов.

К количественному уровню образовательных измерений относятся такие данные, которые получаются в результате применения в аттестационных процедурах методов и инструментария, позволяющих проводить с ними математические операции и представлять их на интервальной шкале или шкале отношений. Кроме этого, количественный уровень образовательных измерений реализуется в автоматизированном режиме и так же проводится обработка результатов выполнения аттестуемыми оценочных заданий, т. е. без участия экзаменаторов.

Качественный уровень предполагает получение данных без широкого применения математико-статистических методов, а результаты измерения представляются в номинальной и порядковой шкалах. Такие данные получаются в результате экспертизы, интервью, собеседований, наблюдений, анализа портфолио аттестуемого, включающего отобранные им самим различные подтверждающие его достижения документы и материалы. В ходе такого экспертного оценивания те,

которые участвуют в аттестационной процедуре, – председатель и члены экзаменационной комиссии – тратят значительные временные ресурсы и собственные усилия, и в итоге получают субъективные оценки. Говорить в этом случае о надежности и валидности результатов просто не приходится, но получить коллегиальную согласованную экспертную оценку возможно, если применяется, например, метод Дельфи при принятии решения.

В таблице 4 представлены типы принятых в метрологии шкал и показана связь со свойствами, которыми они характеризуются. Основными свойствами шкал измерений являются:

1. Идентифицируемость – возможность присвоения числовых значений каждой переменной в наборе данных. К таким значениям не могут быть применены арифметические операции, потому что они служат только для идентификации, но не описания.

2. Величина (магнитуда) – это размерность шкалы измерения, где значения могут быть упорядочены от наименьшего к наибольшему.

3. Равенство интервалов – наличие стандартизированной шкалы, что означает, что разность между двумя любыми соседними уровнями шкалы одинакова. Упорядоченность шкалы не гарантирует равенство интервалов.

4. Абсолютный ноль – естественное и однозначное присутствие нулевой точки, изменение которой невозможно. Данная точка характеризует отсутствие измеряемого признака.

Наиболее общая классификация, предложенная С. Стивенсоном, включает четыре уровня измерений и фиксирует присущие им свойства [350]. Согласно такой классификации различают шкалы *качественные*: шкала наименований (или классификаций) и порядковая шкала; и *количественные*: интервальная шкала и шкала отношений.

Качественные шкалы иногда называют неметрическими, дискретными (концептуальными), а количественные – метрическими, непрерывными (материальными).

Соотнесение свойств с типами шкал измерений

Свойства	Тип шкалы			
	Номина- льная	Поряд- ковая	Интер- вальная	Отно- шений
Идентифицируемость	+	+	+	+
Величина (магнитуда)		+	+	+
Равенство интервалов			+	+
Абсолютный ноль				+

Для каждого уровня измерений существуют группы допустимых преобразований и операций с различными математическими и статистическими величинами, характеризующими измеряемые признаки.

*Качественные шкалы.* На качественном уровне отнесение эмпирических объектов измерения к различным классам проводится по признаку эквивалентности (шкала наименований, или номинальная шкала) или по признаку упорядочения внутри эквивалентных объектов одного класса (порядковая шкала). Для построения шкалы наименований и порядковой шкалы в основном применяются экспертные методы, при которых оценки на шкале считаются достоверными, если они признаны большинством экспертов. Примером номинальной шкалы в аттестационных процедурах может быть дихотомическая шкала, состоящая из результатов зачетной сессии, когда все студенты делятся на две группы – получивших и не получивших зачет.

Порядковые шкалы используются в образовании в тех случаях, когда педагогический контроль и оценивание осуществляется традиционными способами без применения специально разработанного инструментария.

Например, порядковой является привычная четырехбалльная шкала оценок, которую иногда неоправданно называют пятибалльной.

К недостаткам качественных шкал относят ограниченность сферы их применения и низкую точность измерения. Числа или символы, приписываемые объектам путем экспертного оценивания, субъективны и носят исключительно условный характер. Их нельзя суммировать или проводить вычитание, деление, умножение. Порядковая шкала является основой для построения таких шкал, как шкала Лайкерта, Гуттмана и др. [266].

*Количественные шкалы.* К количественным шкалам относятся интервальная шкала и шкала отношений. Процесс их построения основывается на измерениях, поэтому представленные в них оценки характеристик объектов отличаются более высокой объективностью по сравнению с оценками в качественных шкалах и поддаются определенным математическим операциям. Интервальная шкала используется для упорядочения объектов, свойства которых удовлетворяют отношениям эквивалентности, порядка и аддитивности. В ней определено расстояние между объектами и предусмотрена общая для всех объектов единица измерения, а началом отсчета является условно выбранная нулевая точка. Благодаря существованию единицы измерения в интервальной шкале возможны все арифметические действия над числами, кроме операции деления в силу отсутствия абсолютного нуля.

Шкала отношений описывает свойства объектов, удовлетворяющие отношениям эквивалентности, порядка, аддитивности и пропорциональности. Последнее свойство появляется благодаря существованию в этой шкале однозначного естественно определенного критерия нулевого проявления измеряемого свойства – абсолютного нуля. Другими словами, шкала отношений является интервальной шкалой с естественным, а не условным началом отсчета, что расширяет возможности преобразований чисел, приписанных объектам. По сравнению со всеми ранее рассмотренными шкалами эта шкала обеспечивает самый высокий уровень измерений, но реализовать ее в образовании невозможно в силу отсутствия абсолютного нуля.



Сообразно уровням измерений и допустимым операциям на соответствующих им шкалах можно рассматривать две схемы – последовательного и параллельного дизайна бипарадигмального подхода к образовательным измерениям в итоговой аттестации [83].

Последовательный дизайн может проводиться в двух вариантах и оба они зависят от заявленных целей, а также условий, в которых проводятся. На рисунке 9 представлены оба варианта последовательного дизайна. В силу того, что такой дизайн предусматривает применение сначала одной группы методов, а затем другой, то вся процедура образовательных измерений будет разбиваться на несколько этапов. Эти этапы будут соответствовать либо количественному уровню измерений, либо качественному. Сообразно этим уровням выбирается определенный инструментарий, результаты его применения будут выступать условием перехода на следующий этап. В качестве примера последовательного дизайна можно рассматривать следующую процедуру: сначала устное собеседование, что относится к качественному уровню измерений, а по его результатам допуск к прохождению стандартизированному оцениванию без участия человека (например, компьютерное тестирование или выполнение заданий на тренажерах). Аналогичный пример можно привести в обратную сторону – от количественному к качественному уровню измерений, что будет ближе к реальной практической ситуации для аттестационных процедур с высокими ставками. Действительно, оказывается целесообразнее провести стандартизированную оценочную процедуру, соответствующую минимально необходимому уровню компетентности и исключаящую какие-либо экспертные суждения, а затем для получения выводов об уровне компетентности испытуемых проводить собеседование. Итоговая оценка будет составляться на основе прохождения обоих этапов различного уровня образовательных измерений.

Например, при критериально-ориентированном подходе к образовательным измерениям на первом этапе можно принимать решение о переходе испытуемых ко второму этапу только в тех случаях, когда их результаты превышают пороговый балл. В другой ситуации испытуемых допускают ко второму этапу без всяких

условий и подсчитывают суммарный балл, включающий результаты по обоим этапам.

При выборе параллельного дизайна в образовательных измерениях соответствующие действия количественного и качественного характера проводятся одновременно, после чего следует операция объединения получаемых результатов. На рисунке 10 показана схема параллельного дизайна реализации бипарадигмального подхода. В этом случае количественные и качественные измерения проводятся одновременно, а итоговый балл аттестуемых получается путем суммированием оценок по всем правильно выполненным заданиям. На практике параллельный дизайн мало распространен в силу больших временных затрат, привлечения экспертов для проверки выполнения заданий.

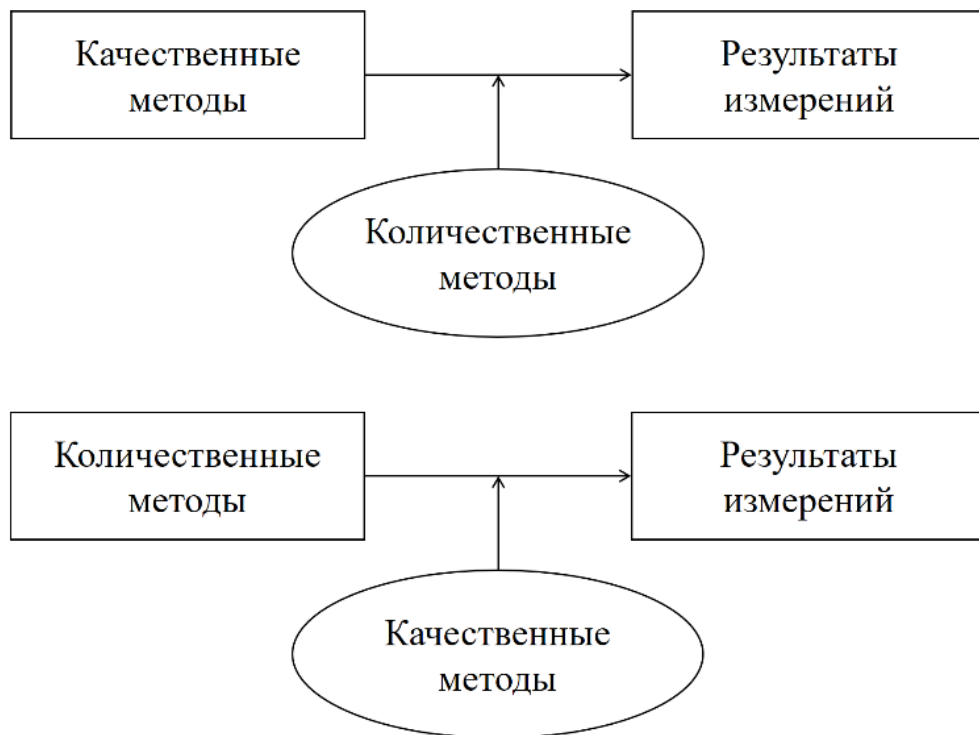
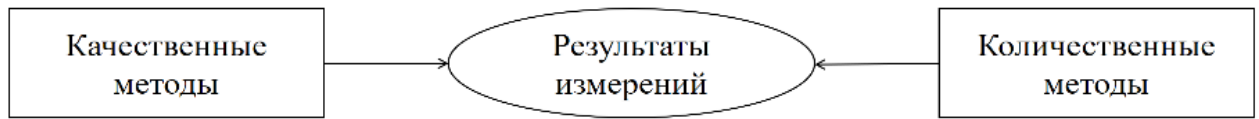


Рис. 9. Схема последовательного дизайна бипарадигмального подхода к образовательным измерениям



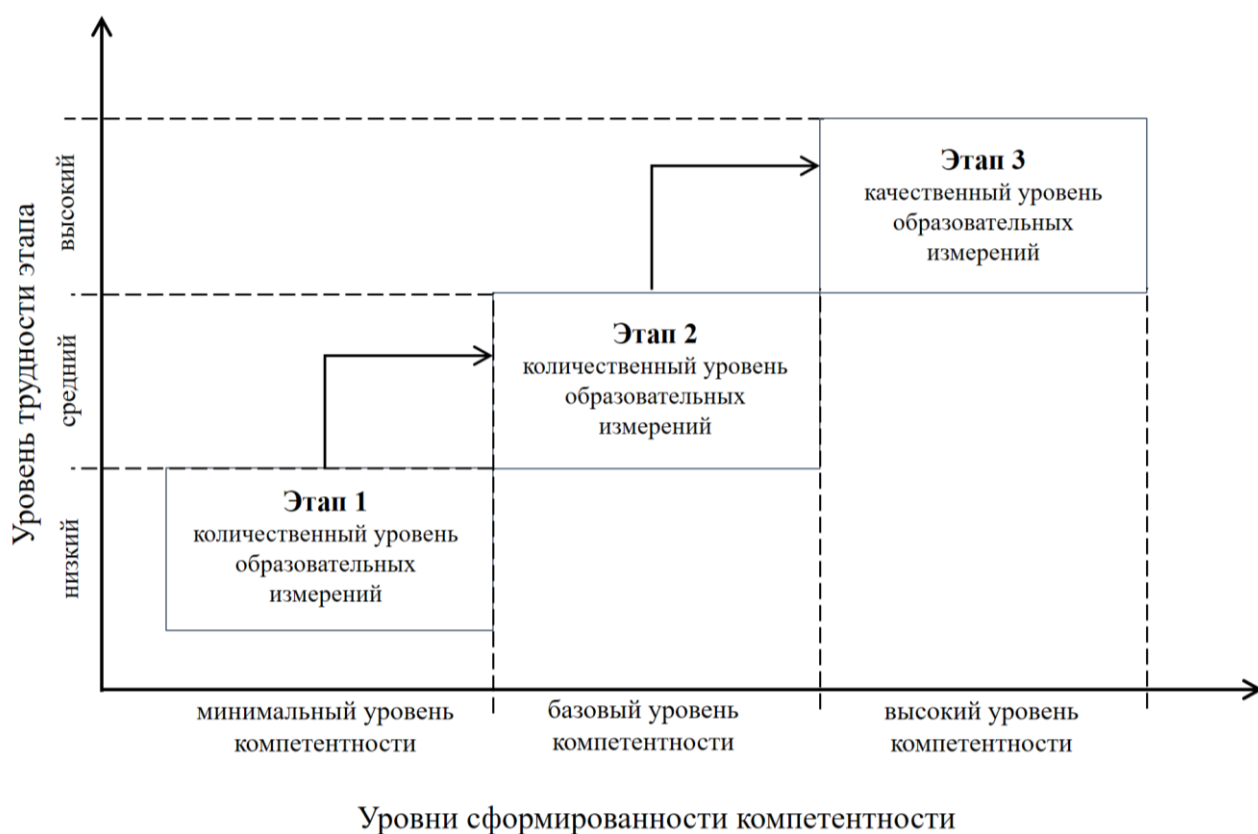
*Рис. 10.* Схема параллельного дизайна бипарадигмального подхода к образовательным измерениям

Развитие бипарадигмального подхода привело к смене приоритетов в выборе средств и процедур оценивания при аттестации. Стандартизированные тесты с выбором ответов, обеспечивающие количественный уровень измерений благодаря автоматизированной проверке, стали применять в сочетании с другими оценочными средствами, включающими различные формы заданий и средства измерения, для проверки выполнения которых требуются качественные оценки экспертов [88]. При интерпретации количественных результатов измерения стали привлекать дополнительную информацию качественного характера, содержащую данные о предыстории развития обучающегося, результаты его анкетирования, различные биографические материалы или контекстные данные, описывающие условия его проживания, работы или обучения.

На представленной основе этого подхода и предлагается выстраивать многостадийное оценивание при итоговой аттестации, поскольку именно возможность сочетать разнообразные формы заданий, включение компетентностно-ориентированных и ситуационных заданий отвечает современному пониманию компетентности. Но, принимая во внимание специфику итоговой аттестации и самих аттестационных испытаний, таких как государственный экзамен, и сложность оцениваемых конструктов – совокупность компетенций в составе компетентности, следует уточнить и конкретизировать схемы последовательного дизайна бипарадигмального подхода. Для этого целесообразно взять систему координат на плоскости, расположив на вертикальной

оси уровень трудности этапа измерения, а по горизонтальной оси заранее согласованные уровни компетентности (рис. 11).

Обобщенное представление аттестационной процедуры – государственного экзамена представлено на рис. 11. Ключевая идея бипарадигмального подхода при реализации многостадийного оценивания заключается в последовательном сочетании двух уровней образовательных измерений – сначала количественного, а затем качественного по определенным правилам обработки результатов каждого этапа и с использованием соответствующих методов и инструментария. Методы и инструментарий как часть методологии многостадийного оценивания представлены в следующем параграфе данной главы, а сама технология реализации описывается в третьей главе диссертации.



*Рис. 11.* Обобщенное представление многостадийного оценивания при последовательном дизайне бипарадигмального подхода

*Вторая составляющая научной организации* многостадийного оценивания – это теория тестов, выступающая общей теоретической рамкой процесса разработки инструментария для измерений в педагогических и психологических науках. Исторически развитие теории тестов связывают с развитием психологии, чему посвящено немало очерков и исследований [5, 109, 125, 165, 268, 272, 279, 299, 301, 313, 330, 333 и др.]. Наиболее полным образом теория тестов и ее приложение к решению задач объективного оценивания в образовании представлены в работах отечественных исследователей В. И. Звонникова, Е. Ю. Кардановой, А. А. Маслака, Ю.М. Неймана, М. Б. Челышковой и др. [83, 87, 114, 165, 180, 219, 263 и др.] и зарубежных авторов Дж. Алгины, Д. Андриха, Д. Вайса, Л. Крокер, В. Линдена, Р. Линн, Ф. Лорда, Б. Райта, Т. Халадайна, Р. Хэмбелтона [272, 274, 294, 314, 316, 329, 331, 339, 365 и др.].

В теории тестов работают с конструктом как гипотетическим концептом, теоретической широкой категорией, объясняющей поведение человека. Конструкт не доступен для непосредственного наблюдения, т. е. является латентной характеристикой, но обнаруживается в поведении человека. Степень, в которой конструкт может характеризовать человека, может быть оценена на основе наблюдений за его поведением [125]. В числе ключевых вопросов, на решение которых нацелена теория тестов, можно выделить следующие:

1. Обоснование и выбор конструкта, подлежащего исследованию, и его операционализация, которая может быть описана разными исследователями через различные виды поведения испытуемого. Это обстоятельство приводит к многообразию подходов к измерению конструктов. Ровно такая ситуация наблюдается сегодня в системе высшего образования, когда, с одной стороны, стандарты задают единые для всех образовательные результаты в форме компетенций, а с другой – фиксируется различная трактовка в операционализации этих компетенций и их оценивании.

2. Определение необходимого и достаточного количества оценочных заданий и их содержания, входящих в систему измерителей, с помощью которой

можно получить надежные и валидные результаты при измерении исследуемого конструкта.

3. Вычисление величины ошибки измерения, неизбежно присутствующей при любых оценочных процедурах и оказывающей влияние на итоговый результат, и способы ее минимизации.

4. Выбор единицы измерения и шкалы, на которой будут зафиксированы результаты измерения конструкта, а также интерпретация значений этих результатов.

5. Связь исследуемого конструкта внутри теоретической системы с другими конструктами. В случае измерения компетенции это означает необходимость определения ее не только в измеряемых наблюдаемых признаках поведения, но и установление логических связей с другими формируемыми компетенциями, образующими интегральную характеристику выпускника – компетентность.

В состав теории тестов входят *классическая теория тестов*, *современная теория тестов* и теория генерализации, позволяющая проводить обобщение получаемых на выборке испытуемых результатов измерений на всю генеральную совокупность. При этом используемое словосочетание «теория тестов» несколько не ограничивает рассматриваемые в этих теориях вопросы, так как тест понимается широко как измеритель (оценочное средство), состоящий из совокупности заданий в стандартизированной форме, которые обладают системообразующими статистическими характеристиками и обеспечивают получение надежных и валидных оценок измеряемого конструкта.

Классическая теория тестов представляет собой связанный с психологическими и образовательными измерениями раздел науки, составляющий концептуальную основу понимания истинного балла, надежности, процедуры для ее оценки и валидности результатов измерения. Истоки научного развития этой теории связывают с антропометрикой и именем Ф. Гальтона (которого принято считать ее основателем), применявшего коэффициент корреляции и регрессию в своих исследованиях человека [110, 125]. Наряду с именем Гальтона большой вклад в развитие теории тестов внес и Ч. Спирмен, который исследовал

интеллектуальные способности учащихся и развивал корреляционные процедуры, введя факторный анализ. К классикам развития теории тестов относят Дж. Кэттэла, предложившего термин «ментальное тестирование» и утверждавшего необходимость тестирования на репрезентативных группах испытуемых с целью получения достоверных оценок. Основы объективного изучения поведения также заложил Э. Торндайк, чья работа «Введение в теорию ментальных и социальных измерений» (1904) относится к числу первых книг по теории тестов, в которой рассматривались вопросы измерений в психологии и образовании. Работы Дж. Гилфорда, Г. Гулликсена, Р. Линна, Ф. Лорда, М. Новика и др. внесли вклад в развитие классической теории тестов, развивая обоснованную Спирменом модель о содержании ошибочного компонента в измерениях характеристик людей.

В развитии классической теории тестов можно выделить два периода. Первый приходится на начало XX века, когда сформировались теоретические положения об измерениях, надежности и валидности, которые основывались на концепции параллельных измерений и теории корреляции. На этой основе конструировались тесты достижений, разрабатывались методы анализа качества заданий, шкалирования результатов их выполнения. Для стандартизации процедуры тестирования научно обосновывались методы формирования репрезентативной выборки. Второй период относится к 60-м годам XX века, когда происходило широкое использование стандартизированного тестирования с заданиями закрытой формы (преимущественно с выбором ответов), допускающих автоматизированную проверку их выполнения.

В основе классической теории тестов лежит концептуальное понятие истинного балла, который рассматривается как идеализированная константа испытуемого в гипотетической генеральной совокупности заданий бесконечного теста. Такое понимание истинного балла носит концептуальный характер, что вызывает сложности при переходе от теоретического определения к эмпирическим референтам, и этот переход по-разному представлен в классической и современной теории. В классической теории тестов, как отмечает М. Б. Чельшкова, идея перехода состоит в построении уравнения линейной регрессии, но в силу низкой

эффективности не применяется в современной практике образовательных измерений [263]. Вместе с истинным баллом рассматривается еще одно понятие – ошибка измерения, которая определяется как статистическая величина, отражающая степень отклонения наблюдаемого балла от истинного балла испытуемого. Дисперсию результатов измерений можно трактовать как результат влияния факторов двух видов, один из которых связан с изменчивостью самой измеряемой переменной, а другой – с группой случайных погрешностей, обусловленных нестабильностью измерительной процедуры [262]. Величина ошибок измерения, порождаемых группой случайных факторов, является предметом оценивания в процессе исследования надежности измерений [262, 340, 357]. Существование ошибки измерения закладывается на аксиоматическом уровне в классической теории, логический фундамент которой строится на пяти утверждениях [97, 298, 335].

Первое утверждение заключается в том, что наблюдаемый результат  $X$  рассматривается как композиция двух компонентов – истинной оценки  $T$  и случайной ошибки измерения  $\varepsilon$ :

$$X = T + \varepsilon \quad (1)$$

Очевидно, что одной из целей при измерениях является минимизация ошибки измерения, приближающая наблюдаемый балл к истинному.

Второе утверждение классической теории связано с понятием «параллельных форм теста». Несколько форм считаются параллельными, если они разработаны в соответствии с единой спецификацией, содержат одинаковое количество заданий попарно равной трудности и каждый испытуемый получит одну и ту же оценку при выполнении двух форм теста. Согласно этому постулату истинный балл испытуемого не изменяется при использовании различных параллельных форм теста и ковариации результатов тестирования должны быть одинаковы.

Третье утверждение строится на гипотетическом предположении о существовании результатов бесконечного числа измерений одного и того же испытуемого с помощью параллельных форм теста. При этом среднее значение



ошибок измерения истинного балла испытуемого стремится к нулю при бесконечном числе испытаний.

Четвертое положение основано на идее измерения бесконечного числа испытуемых с помощью параллельных форм теста. В этом случае среднее значение ошибок измерения баллов будет стремиться к нулю при бесконечном числе испытуемых.

Пятое утверждение говорит о том, что истинный балл испытуемого величина постоянная, не изменяется при использовании параллельных форм теста.

Эти утверждения выражаются фундаментальным соотношением классической теории тестов:

$$\sigma_X^2 = \sigma_T^2 + \sigma_\varepsilon^2, \quad (2)$$

где  $\sigma_X^2$  – дисперсия наблюдаемых баллов  $X$  генеральной совокупности испытуемых,  $\sigma_T^2$  – дисперсия истинных баллов  $T$ ,  $\sigma_\varepsilon^2$  – дисперсия ошибок измерения  $\varepsilon$ .

Перечисленные утверждения в классической теории дополняются двумя предположениями. Первое связано с допущением о непрерывном распределении истинных баллов в генеральной совокупности в противоположность объективно существующему дискретному характеру распределения наблюдаемых баллов тестируемой выборки. Второе предположение заключается в том, что распределение наблюдаемых баллов, истинных баллов и ошибки измерения подчинены нормальному закону.

Если равенство (2) представить в следующем виде:

$$\frac{\sigma_T^2}{\sigma_X^2} = 1 - \frac{\sigma_\varepsilon^2}{\sigma_X^2}, \quad (3)$$

то из этого соотношения (3) получают формулу для оценивания надежности результатов измерений, где  $\frac{\sigma_\varepsilon^2}{\sigma_X^2}$  – понимается как среднее арифметическое дисперсий ошибок измерений различных испытуемых из генеральной совокупности, поскольку ошибка при оценке истинного балла будет меняться для различных испытуемых из группы. Чем ближе  $\sigma_X^2$  к  $\sigma_T^2$ , тем выше корреляция между множеством наблюдаемых баллов  $X$  и множеством истинных баллов  $T$ , и,

следовательно, надежнее результаты по тесту. В классической теории под *надежностью* понимают характеристику теста, отражающую точность измерений, а также устойчивость результатов к действию случайных факторов, с неизбежностью приводящих к ошибкам измерений [263]. Тест считается надежным, если он обеспечивает высокую точность измерений, а также дает при повторном выполнении на той же выборке близкие результаты при условии того, что подготовка испытуемых не изменилась за время до повторного выполнения теста. Отношение  $\sigma_T^2/\sigma_X^2$  рассматривают как концептуальное определение коэффициента надежности, обозначаемого символом  $r_H$  и имеющего вид:

$$r_H = 1 - \frac{\sigma_\varepsilon^2}{\sigma_X^2}. \quad (4)$$

Коэффициент надежности (4) носит основополагающий характер, но является теоретическим понятием, поскольку по эмпирическим результатам выполнения теста или параллельных форм тестов, которых в практике не существует, нельзя определить дисперсию ошибочных компонентов измерения  $\sigma_\varepsilon^2$ . Поэтому используют различные методы оценивания коэффициента надежности по получаемым оценкам [87, 125, 263]. К их числу относятся:

- процедуры, предусматривающие предъявление теста два раза одной и той же выборке испытуемых;
- процедуры с однократным предъявлением теста;
- методики, связанные с вычислением ковариаций заданий теста.

Для повышения надежности руководствуются следующими положениями:

- 1) повышение внутренней согласованности теста, которая обеспечивается за счет содержательной однородности заданий при отборе их в тест или субтест. Это означает, что следует отдавать предпочтение заданиям, которые имеют наибольшую корреляцию с другими заданиями;
- 2) увеличение длины теста, что согласуется с концептуальным понятием истинного балла испытуемого и выполнении им бесконечного числа заданий, при котором наблюдаемый балл будет приближаться к истинному, а ошибка измерения уменьшаться. В практике осуществления оценочной деятельности, конечно, это

недостижимо, так как по мере увеличения числа заданий у испытуемого снижается мотивация к их выполнению, возрастает утомляемость и, тем самым, увеличивается ошибка измерения. Поэтому следует принимать во внимание дисперсию тестовых баллов, их распределение, физиологические особенности испытуемых и их возраст, оптимальное время для прохождения аттестационной процедуры, форму используемых оценочных заданий.

Указанные положения относятся не только к вопросам обеспечения надежности измерений, но подводят к другому важному вопросу о контроле заданий по трудности, которые по своему назначению и содержанию нацелены на измерение выбранной переменной, для которой они и разрабатывались [147]. Ответы на эти вопросы связаны с исследованиями валидности. Выделение одного критерия качества тестов для оценивания только надежности или только валидности, нередко наблюдаемые на практике при разработке оценочного инструментария, не состоятельны в методологическом плане, поскольку оба критерия – надежность и валидность – представляют совокупность необходимых и достаточных условий единого процесса обоснования качества заданий и получаемых при их использовании результатов. *Валидность* трактуется как характеристика адекватности эмпирических результатов поставленной цели измерения [5, 12, 147, 263, 364]. Отождествление валидности непосредственно с измерителем считается некорректным, так как исследование валидности необходимо для понимания того, какие выводы следует сделать по полученным результатам выполнения оценочных заданий.

Валидность также относится к методологическим аспектам образовательных измерений и имеет множество аспектов рассмотрения. Выявление валидности предусматривает анализ различных аспектов соответствия целям измерения путем соотнесения получаемых результатов с различными критериями, независимо описывающими вне ситуации оценивания все то, что измеряется. Высокая корреляция результатов измерений с такими критериями свидетельствует о высокой валидности результатов измерения. Поскольку критериев может быть несколько, то различают несколько видов валидности и методов ее исследования

[125, 219, 325]. Для исследования выделяют три вида валидности – содержательную, конструктивную и прогностическую.

Понятие «конструктивная валидность» было введено в 1954 г. в Технических рекомендациях для психологических тестов и диагностических методик, послуживших основой для современных стандартов качества и справедливости в образовательном тестировании [302]. Считается, что конструктивная валидность является родовым понятием по отношению к содержательной и прогностической валидности, поскольку она связана с концептуальным выбором конструкта, который подлежит измерению, и с рассмотрением меры его проявления в наблюдаемых результатах. Д. Кэмпбелл утверждал, что для доказательства конструктивной валидности необходимо показать, что тест имеет высокие корреляции с переменными, которые он должен измерять, и не имеет значимых корреляций с переменными, которые не обозначены как цель измерения [83].

Исследование содержательной валидности связано с обоснованием выбранного для измерения конструкта (или конструктов в случае многомерных измерений) и содержания для разработки оценочного инструментария. Другими словами, определяя данный вид валидности, содержательную валидность следует рассматривать как степень релевантного и репрезентативного отражения выбранной переменной в оценочных заданиях теста. Для обеспечения содержательной валидности составляется спецификация – особый документ, отражающий цели, структуру и наполнение оценочного инструментария, – которая подлежит компетентной оценке экспертов. Такая независимая экспертиза спецификации необходима для подтверждения адекватности, полноты и правильных пропорций предлагаемых заданий тому содержанию, которое подлежит оцениванию.

Прогностическая валидность связана с ситуацией, когда необходимо получить выводы относительно последующей деятельности испытуемого. Таким образом, прогностическая валидность показывает меру вероятности прогноза будущей успешности в обучении или профессиональной деятельности испытуемого по результатам его оценивания [5, 87, 125].

Качество результатов измерений, которые связаны с классическими понятиями надежности и валидности, не может быть определено один раз. Различные виды валидности обеспечивают различные выводы о результатах, так же, как и оценка надежности результатов измерений зависит от ряда факторов, меняющихся при различных ситуациях применения оценочного инструментария. Поэтому говорить об аутентичности результатов оценивания можно только после того, как проведено измерение и оценена надежность и валидность. Описание методик определения надежности и содержательной валидности результатов при итоговой аттестации в рамках данного исследования рассматривается в третьей главе.

Классическую теорию, несмотря на существующие ограничения и недостатки, следует рассматривать как важную часть при проектировании аттестационных процедур и на начальном этапе разработки и апробации оценочного инструментария, поскольку ее математико-статистический аппарат доступен без специальной подготовки многим исследователям и практикам. Вместе с тем, следует отметить недостатки этой теории. К их числу относятся [262]:

1) неустойчивость статистических оценок характеристик заданий и самого теста. Это означает, что задания можно использовать только на выборках, которые близки по своим свойствам тем выборкам испытуемых, на которых они разрабатывались и апробировались. Оценка трудности заданий зависит от уровня подготовленности конкретной группы испытуемых;

2) отсутствие возможности сравнить результаты испытуемых, поскольку получаемые оценки в рамках этой теории зависят от количества и трудности заданий определенного варианта теста либо параллельным вариантам. Оценки испытуемых выражаются в порядковой шкале, которая не обеспечивает сопоставимость ввиду отсутствия единицы измерения и точки отсчета при шкалировании результатов;

3) неизбежное влияние на результаты измерений ошибок (случайных и систематических), методы минимизации которых мало эффективны в рамках классической теории тестов;

4) отсутствие прогностической информации, по которой можно предсказать вероятность успеха испытуемых определенного уровня подготовленности при выполнении различных по трудности заданий;

5) невозможность получения дифференцированной ошибки измерения, поскольку в классической теории тестов она оказывается одинаковой для всех испытуемых при однократном выполнении заданий теста.

В силу стремления исследователей преодолеть эти недостатки, а также благодаря открывшимся возможностям, связанным с развитием информационно-компьютерных технологий, вычислительных мощностей, к 80-м годам XX века сформировался аппарат *современной теории тестов*. Строго говоря, речь идет о теории Item Response Theory (IRT), дословный перевод ее названия на русский язык не имеет никакого смысла. Однако в зарубежной литературе IRT ставится в один ряд с названием Modern Test Theory, тем самым подчеркивая следующий этап после развития классической (традиционной) теории тестов (Classical Test Theory, СТТ). В данном исследовании будет использоваться несколько названий – современная теория тестов и IRT. При этом под IRT, являющейся частью общей теории латентно-структурного анализа, будем понимать теорию методов исследования и интерпретации наблюдаемых (эмпирических) результатов выполнения теста для получения оценок латентных параметров испытуемых и заданий теста. Также ряд исследователей определяют IRT как научную основу для конструирования и применения педагогических тестов [158, 263] или как теорию моделирования и параметризации тестов [114, 165, 180].

Начало становления IRT относят к 40–50-м годам XX века, связывая с работами П. Лазарсфельда, К. Ричардсона, Д. Лоули и Л. Такера [263, 301, 333, 346]. Позднее в 60-х годах XX века в исследованиях Г. Раша, Ф. Лорда, А. Бирнбаума были предложены математико-статистические модели измерения IRT [284, 330, 345]. В настоящее время возможности и аппарат IRT широко используется в различных оценочных процедурах, включая сравнительные исследования качества образования, ЕГЭ, систему аккредитации специалистов в области здравоохранения, адаптивное тестирование, измерение уровня

компетентности, актуальных навыков XXI века и другие [53, 59, 64, 82, 135, 138, 140, 161, 197, 262, 291, 297, 308, 312, 315, 336, 358, 360].

Первоначально в IRT вводится основное предположение о существовании функциональной взаимосвязи между наблюдаемыми результатами измерения и латентными характеристиками испытуемых, выступающими в роли конструкта. Эти латентные характеристики в соответствии с определенными целями измерения могут рассматриваться как способности испытуемых, освоенные компетенции или уровни сформированности компетентности. Функциональную связь между наблюдаемыми результатами и латентными параметрами отображают в форме математических моделей IRT, которые представлены в Приложении 1. Для аналитического задания этой взаимосвязи наблюдаемые результаты связывают с вероятностью правильного ответа испытуемого на различные по трудности задания. Сам по себе латентный параметр появляется в результате концептуализации оцениваемой истинной величины измеряемого качества испытуемого и в этом смысле может быть отождествлен с истинными компонентами измерения классической теории тестов. Но в отличие от классической теории, где индивидуальный балл испытуемого рассматривается как постоянное число, в IRT латентный параметр трактуется как некоторая переменная. Начальная оценка параметра получается непосредственно из эмпирических данных тестирования. Переменный характер измеряемой величины указывает на возможность последовательного приближения к объективным оценкам параметра с помощью тех или иных итерационных методов.

Помимо предположения о связи между наблюдаемыми и латентными переменными в IRT вводится еще одно предположение, которое связано с понятием локальной независимости, означающей отсутствие вероятностной связи между ответами испытуемого с уровнем подготовленности, обозначаемым символом  $\theta$ , на различные по трудности задания. Иными словами, в рамках предположения о локальной независимости можно утверждать, что при фиксированном значении параметра  $\theta$  нет никакой связи между вероятностями правильных ответов на различные задания. На практике, конечно, ответы

испытуемых на задания связаны между собой и тем сильнее, чем выше корреляция между заданиями и больше пересечение содержательных элементов в различных заданиях измерителя. Положение о локальной независимости носит формальный характер, но оно играет роль при использовании математического аппарата ИРТ. Так оно позволяет рассматривать вероятность выполнения совокупности заданий как произведение вероятностей выполнения отдельных заданий и реализовать на основе теоремы об умножении вероятностей возможности для получения оценок наибольшего правдоподобия параметров – уровня подготовленности  $\theta$  и уровня трудности задания  $\beta$  [263].

Третье предположение ИРТ, как и в классической теории, связано с непрерывным распределением истинных баллов в генеральной совокупности испытуемых и о нормальном законе распределения наблюдаемых баллов, оценок параметров испытуемых и трудности заданий, а также ошибок измерения.

Первоначально аппарат теории ИРТ был создан исключительно для одномерных измерений, когда каждому испытуемому ставилось в соответствие только одно значение латентного параметра. Требование одномерности впоследствии было снято в рамках разработанной многомерной современной теории тестов [90, 299, 329, 348]. Аналогично шел процесс развития математических моделей для различной формы представления наблюдаемых баллов по заданиям измерителя. На смену моделям, предназначенным для дихотомического оценивания результатов выполнения заданий (за верный ответ на задание присваивается 1, за неверный ответ или пропущенное задание – 0), пришли политомические модели ИРТ, когда за ответы на каждое задание можно выставить различные баллы (0, 1, 2 и т. д.) [114, 279, 329]. Благодаря этим двум нововведениям появилась возможность использовать ИРТ для различных оценочных средств, в роли которых сегодня могут выступать не только стандартизованные тесты с выбором ответов, но и ситуационные задания практического характера [105, 206, 351].



К преимуществам современной теории тестов, реализуемым при условии того, что эмпирические результаты измерений соответствуют требованиям ее моделей, относят [263]:

- инвариантность оценок параметра испытуемых, обеспечиваемую благодаря их независимости от трудности заданий измерителя;
- устойчивость оценок параметра трудности заданий, обеспечиваемую благодаря их инвариантности от свойств выборки испытуемых; благодаря устойчивости оценок проводится калибровка заданий и создаются структурированные банки для компьютерной генерации параллельных вариантов измерителей;
- представление оценок параметров испытуемых и трудности заданий измерителя в одной и той же шкале, имеющей свойства интервальной;
- наличие дифференцированных ошибок измерения для оценок параметров испытуемых и трудности заданий измерителя;
- возможность предсказать вероятность правильного выполнения задания измерителя любым испытуемым до его предъявления;
- возможность оценить эффективность различных по трудности заданий для измерения испытуемых, отличающихся по уровню подготовки;
- возможность построения аппарата, основанного на информационной функции теста, для достижения планируемой надежности измерений при моделировании измерителя из банка калиброванных заданий;
- возможность построения адаптивных алгоритмов путем оптимизации трудности заданий в процессе их предъявления испытуемым, обеспечиваемой благодаря представлению оценок параметров испытуемых и трудности заданий измерителя в одной и той же шкале;
- возможность построения специальной функции для анализа дифференцированного функционирования заданий и исключения тех из них, которые дискриминируют отдельные группы испытуемых по различным признакам (гендерным, языковым и т.д.).

При обсуждении подходов к оцениванию параметров испытуемых в IRT принято подчеркивать их латентный характер, который принципиально отличает оценки параметров в IRT от наблюдаемых баллов, получаемых с помощью методов классической теории тестов. Латентность означает, что в IRT совершается переход от наблюдаемых (сырых) баллов порядкового характера к точечным оценкам в метрической шкале, которые могут быть приняты за истинные баллы испытуемых. В основе перехода лежит определенный математический аппарат, использующий метод максимального правдоподобия и описанный в следующей главе. Переход от наблюдаемых баллов к оценкам латентных параметров осуществляется методами последовательного приближения на основе специальных итерационных процессов, где в качестве начальных значений выбирают наблюдаемые оценки латентного параметра, полученные непосредственно из эмпирических результатов измерения.

Вероятностные модели IRT основаны на утверждении о существовании функциональной связи между наблюдаемыми результатами измерения и латентными параметрами испытуемых [314]. При установлении вида этой связи важно понимать, что первопричиной являются латентные параметры. Если говорить точнее, то взаимодействие двух множеств значений латентных параметров испытуемых  $\theta$  и трудности заданий  $\beta$  порождает наблюдаемые результаты измерения  $x_{ij}$ , где индексы  $i$  и  $j$  обозначают номера испытуемых и заданий измерителя соответственно [263].

На практике же ставится задача, обратная теоретическим положениям – по ответам испытуемых на задания измерителя оценить значения латентных параметров  $\theta$  и  $\beta$ . Для ее решения необходимо ответить на два вопроса. Первый вопрос касается выбора соотношения между латентными параметрами  $\theta$  и  $\beta$ , отражающего их взаимодействие. Второй вопрос связан с определением математической модели, описывающей связь между латентными параметрами  $\theta$  и  $\beta$  и наблюдаемыми результатами измерения  $x_{ij}$ .

Ответ на первый вопрос дал датский математик Г. Раш (G. Rasch), рассмотревший оптимальное соотношение между  $\theta$  и  $\beta$  и предложивший форму связи этих параметров в виде разности:  $\theta - \beta$  [345]. Введение разности для оценок

уровня подготовленности и трудности заданий предусматривает существование единой интервальной шкалы с единицей измерения, получившей в работах зарубежных исследователей название «логит» («*logit*» – от сокращенного английского *log-odds unit* – единица логарифмических шансов). Из анализа разности параметров  $\theta_i$  и  $\beta_j$  в абсолютном значении, т. е. по модулю  $|\theta_i - \beta_j|$ , можно вычислить расстояние, на котором находится испытуемый  $i$  с оценкой уровня подготовленности  $\theta_i$  от задания с трудностью  $\beta_j$ . Эта информация позволяет выбирать для оценочной процедуры такие задания, которые были бы оптимальны по трудности для испытуемых. Если рассматриваемая разность по модулю велика и отрицательна, то данное задание не предназначено для оценивания испытуемого с параметром  $\theta_i$ . В другом случае, когда получаются большие положительные значения этой разности, такие задания подходят для критериально-ориентированного подхода к образовательным измерениям, но являются малоинформативными для нормативно-ориентированной интерпретации результатов.

Определение математической модели, описывающей взаимосвязь между эмпирическими результатами измерения и значениями латентных параметров  $\theta$  и  $\beta$ , является центральным в IRT для ответа на второй вопрос. Относительная инвариантность значений латентных переменных от конкретного измерения, определенная устойчивость частот появлений их значений послужили основанием для использования понятия вероятности события как меры возможности его появления. В качестве такого события исследователи выбрали правильный ответ  $i$ -го испытуемого на  $j$ -е задание теста. Условную вероятность такого правильного ответа обозначают символом  $P_{ij}$  и выражают с помощью различных математических моделей, которые записываются как функции одной переменной в случае одномерных измерений. Можно рассматривать условную вероятность правильного выполнения  $i$ -м испытуемым с уровнем подготовленности  $\theta_i$  различных по трудности заданий, считая  $\theta_i$  параметром  $i$ -го испытуемого, а  $\beta$  – независимой переменной. Тогда условная вероятность  $P_i$  будет функцией латентной переменной  $\beta$ :

$$P_i \{x_{ij} = 1 | \theta_i\} = f(\theta_i - \beta_j), i = 1, 2, \dots, N. \quad (5)$$

Аналогично вводится условная вероятность правильного выполнения  $j$ -го задания трудностью  $\beta_j$  различными испытуемыми. В данном случае независимой переменной является  $\theta$ , а  $\beta_j$  – параметр, определяющий трудность  $j$ -го задания измерителя. Тогда

$$P_j \{x_{ij} = 1 | \beta_j\} = F(\theta - \beta_j), j = 1, 2, \dots, n, \quad (6)$$

где

$$x_{ij} = \begin{cases} 1, & \text{если ответ } i\text{-го испытуемого на } j\text{-е задание верный,} \\ 0, & \text{если ответ } i\text{-го испытуемого на } j\text{-е задание неверный,} \end{cases}$$

$N$  – число испытуемых,  $n$  – количество заданий в тесте.

Если подставить в функцию  $P_i(\theta)$  значение переменной  $\theta = \theta_i$  или в функцию  $P_j(\beta)$  значение  $\beta = \beta_j$ , то получится выражение для вероятности  $P_{ij}$ , значения которой можно охарактеризовать следующим образом:

$$P_{ij} \rightarrow 1, \text{ когда } \theta_i = \beta_j \gg 0,$$

$$P_{ij} \rightarrow 0, \text{ когда } \theta_i - \beta_j \ll 0,$$

$$P_{ij} = 1/2 \text{ при } \theta_i = \beta_j.$$

В теории ИРТ функции  $f(\beta)$  и  $F(\theta)$  получили название Item Response Functions (IRF). Графиком первой функции является убывающая индивидуальная кривая испытуемого (рис. 13), а вторая возрастающая функция – это характеристическая кривая задания S-образной формы (рис. 13).

При выборе вида функций  $P_i$  и  $P_j$  учитываются обстоятельства как эмпирического, так и математического характера. Подробный анализ оснований для такого выбора обоснован в работе Ф. Лорда [330].

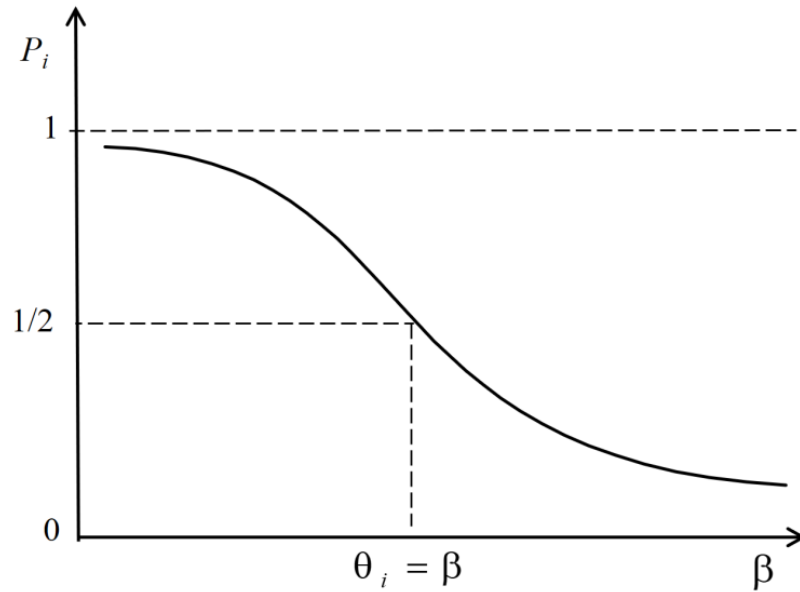


Рис. 12. График функции  $P_i=f(\beta)$  (индивидуальная кривая испытуемого)

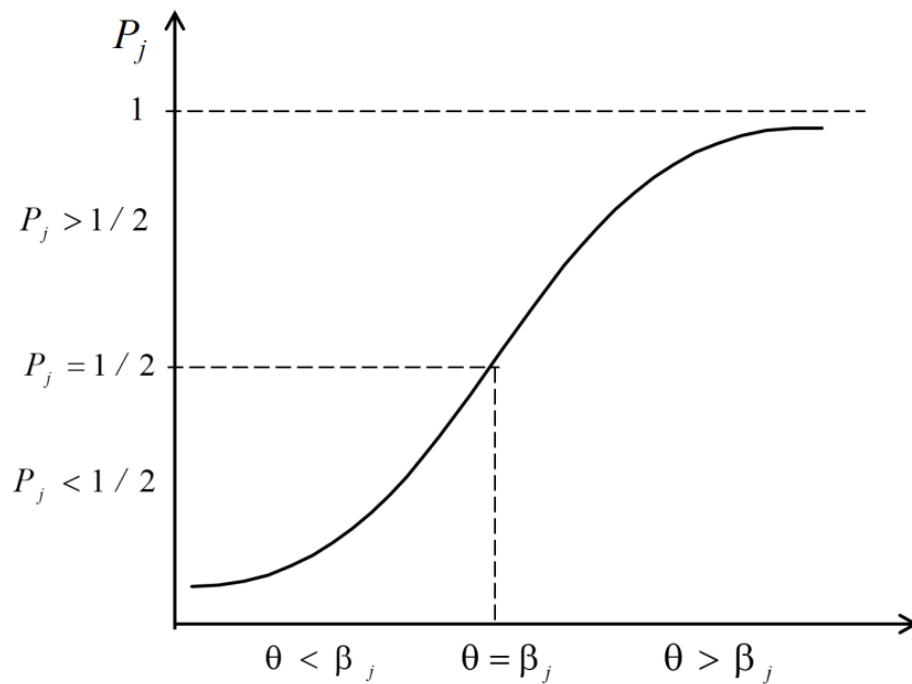


Рис. 13. График функции  $P_j=F(\theta)$  (характеристическая кривая задания)

Описание логистических функций приводится в различной литературе по современной теории тестов [67, 114, 165, 263, 279, 286, 314, 330 и др.]. Так, М. Б. Челышкова указывает, что в предположении о существовании нормального распределения оценок латентных параметров  $\theta$  и  $\beta$  таких функций предлагаются две [263]. Одна из них, обычно обозначаемая  $\Psi(x)$ , относится к семейству

логистических кривых, другая  $\Phi(x)$  – является интегральной функцией нормированного нормального распределения. Поскольку для одних и тех же значений  $x$  ординаты точек графиков функций  $\Phi(x)$  и  $\Psi(1,7x)$  отличаются друг от друга достаточно мало, то в том, что их две, нет противоречия. Для всех  $x$ , принадлежащих области определения этих функций,  $|\Phi(x) - \Psi(1,7x)| < 0,01$ . Аргумент в пользу выбора логистической функции связан с относительной простотой ее аналитического задания, удобной при оценивании параметров  $\theta$  и  $\beta$ , поэтому в практических приложениях предпочтение обычно отдают функции  $\Psi(1,7x)$ . Число используемых параметров, входящих в аналитическое задание функций, служит основанием для подразделения семейства IRF на отдельные виды [263]. Подробно эти модели представлены в Приложении 1.

Следует отметить, что в IRT рассматриваются не только дихотомические задания, в результате выполнения которых испытуемый получает 1 балл, но и полиномические – с несколькими категориями выполнения заданий. Как правило, в ситуации дихотомического оценивания выполнения заданий применяют модели Раша, Бирнбаума и трехпараметрическую модель. Для случая с полиномическими оценками по заданиям измерителя используются модели с градуированными ответами, модель частичного кредитования с произвольными промежуточными категориями выполнения заданий, модель с фиксированными промежуточными категориями выполнения заданий.

В сравнении с классической теорией тестов IRT существенным образом продвигает подходы к объективным методам оценивания, используя в своем арсенале, в первую очередь, обширный математический аппарат. Широкое использование IRT в практике образовательного оценивания стало возможным благодаря вычислительным мощностям, современному программному обеспечению и компьютерной технике, поскольку реализация алгоритмов IRT не мыслима без этого. Вместе с тем, необходимо заметить, что при освоении вопросов теории тестов, получение первого опыта разработки авторских оценочных заданий и измерителей, апробации и обработки результатов, конечно, следует начинать с

классической теории тестов. Этот тезис будет обоснован в опытно-экспериментальной части работы.

Третья составляющая, которая входит в состав рассматриваемой теории тестов, является теория генерализации (Г-теория). Ее обоснованием занимались Л. Кронбах, Г. Глезер, Н. Раджарасам, Р. Бреннан и др. Г-теория отвечает на вопрос, в какой мере результаты измерений, проведенных на выборочной совокупности испытуемых, могут быть распространены на всю генеральную совокупность. Эта теория связана с оцениванием устойчивости результатов измерений в контексте их надежности, гомогенности и параллельности вариантов теста [287].

В основе концепции Г-теории лежит представление о генеральной оценке [83, 287, 294]. В рамках классической теории тестов исследовались различные аспекты надежности: оценка устойчивости измерений, параллельности форм теста, его внутренней согласованности, однородности и др. под влиянием случайных ошибок. При этом систематические ошибки, всегда присутствующие в измерениях, оставались без внимания. В Г-теории рассматривается концепция генеральной оценки в отличие от классического истинного балла и анализируются как случайные, так и систематические ошибочные компоненты. Результаты измерений комплексно анализируются для выявления влияния различных условий измерения, которые имеют специально название – фасеты [88, 287]. Фасеты формируются по одному или нескольким факторам: это могут быть различные множества и совокупности испытуемых, заданий, экзаменаторов, условий, показателей и т. д. При анализе генерализуемости данных составляют дизайн данных с определенной группировкой фасетов, подбор которых зависит от целей измерения. Благодаря этому становится возможным использование результатов измерений в практике и для принятия управленческих решений.

В основе оценивания генерализуемости данных измерений лежит дисперсионный анализ. За коэффициент генерализации чаще всего принимают отношение дисперсии генеральной оценки к дисперсии наблюдаемой оценки. Вид формулы для коэффициента генерализации зависит от специфических

особенностей фасетного дизайна. Первые шаги в генерализации данных измерения включают обычные начальные процедуры для анализа вариаций ANOVA (Analysis of Variance), которые сводятся к формированию фасетов, определению модели исследования, сумм квадратов степеней свободы и среднего квадратического для каждого из фасетов в дизайне исследования [88, 287].

При построении фасетных дизайнов для генерализации используют круговые диаграммы Венна, в которых средняя площадь для главного компонента фасета представлена с помощью круга, взаимодействие фасетов выражено пересечением или вложением кругов меньшей площади [55, 88]. Обращение к диаграммам при массовых оценочных процедурах позволяет визуализировать влияние всех факторов на результаты измерений и выстроить дизайн, способствующий обоснованной генерализации и интерпретации ее результатов при принятии управленческих решений в аттестационных процедурах.

Как отмечает в своем исследовании В.И. Звонников, теория генерализации обеспечивает научную основу для оценивания надежности широкого диапазона компонентов измерений: измерителей и инструкций к ним, рейтинговых шкал и уровневых шкал, моделей измерения, процедур проведения оценивания и обработки их результатов, поведения экзаменаторов и экспертов, а также других компонентов, влияющих на ошибку измерения [83]. Применение положений Г-теории целесообразно на федеральном уровне, например, обобщая результаты итоговой аттестации выпускников по одним и тем же направлениям подготовки с учетом направленности (профиля) образовательных программ соответствующего уровня высшего образования. Это позволит говорить об обоснованных управленческих решениях на уровне образовательной политики страны, но применять теорию генерализации к нескольким образовательным организациям или внутри нее не имеет никакого смысла. Применение этой теории тесно связано с управленческими решениями, и они будут тем точнее, чем корректнее будут использоваться для этого данные. Поэтому результаты, получаемые на количественном уровне измерений на основе IRT, будут наиболее релевантными,



что подтверждается соответствующими исследованиями отечественных ученых [55, 88, 243].

Таким образом, представленные бипарадигмальный подход с последовательным дизайном проведения измерений и теория тестов, обеспечивающая реализацию принципов многостадийного оценивания при итоговой аттестации, служат научными основаниями для организации аттестационных процедур с высоким уровнем значимости. Научная организация предусматривает определение и выбор методов и инструментария, соответствующего теоретико-методологическим основаниям, которые представлены в следующем параграфе.

#### **2.4. Методы и инструментарий многостадийного оценивания уровня компетентности выпускников при итоговой аттестации**

Получение данных, на основании которых принимаются различного уровня решения, в случае данного исследования – выводы об уровне компетентности выпускников при итоговой аттестации, связано не только с теоретическими основами организации (*как?*) и содержания (*что?*) аттестационных процедур, но с методами и инструментарием (*с помощью чего?*) проведения оценивания.

Определяя методы для многостадийного оценивания в условиях итоговой аттестации, необходимо учитывать следующие методологические основания:

1) компетентностный подход к проектированию содержания, результатов освоения образовательной программы и оценочного инструментария для аттестационных процедур;

2) бипарадигмальный подход с последовательным дизайном образовательных измерений и соотносимый с уровневой оценкой компетентности;

3) положения теории тестов, в частности, классической теории для начального этапа конструирования оценочных заданий с первичной апробацией и

современной теории тестов для полномасштабной реализации аттестационных процедур с целью обеспечения надежных и валидных результатов;

4) критериально-ориентированный подход для интерпретации получаемых результатов аттестационных процедур в отношении каждого аттестуемого студента или выпускника на каждой стадии оценивания.

Первое основание предусматривает отбор таких методов и сообразно им разработку такого инструментария, какие будут отвечать согласованному пониманию компетенций и компетентности, их уровневого оценивания на разных этапах освоения обучающимися образовательных программ.

Исходя из уровневости оценивания компетентности, аттестационные процедуры должны предусматривать использование различного оценочного инструментария и соответствующих шкал, на которые размещается результат измерений и корректно интерпретируется. Для этого при разработке и использовании авторам и организаторам аттестационных процедур необходимы положения теории тестов, определяющие допустимость и уместность использования того или иного оценочного инструментария.

Критериально-ориентированный подход, имеющий своей целью интерпретировать результаты аттестуемых по отношению к содержательной области, включенной в измеритель, и разделить их на аттестованных и не аттестованных согласно заранее установленному критериальному баллу (стандарту выполнения каждого этапа многостадийного оценивания), также задает правила отбора и включения в оценочный инструментарий заданий определенной трудности. Методики установления критериального балла и его эмпирическая валидизация представлены в главе диссертации об опытно-экспериментальной работе, но, рассматривая методы и средства как составляющие методологии многостадийного оценивания, следует дать пояснения и замечания относительного самого подхода.

При критериально-ориентированном подходе инструментарий должен с максимальной полнотой отразить все то, что можно принять за полный объем планируемых к освоению результатов образовательной программы. Доля

правильно выполненных заданий рассматривается как уровень освоения общего объема (то, что можно принять за 100 %). При итоговой аттестации встает вопрос, что следует принять за полный объем. Строго говоря, это должен быть весь перечень компетенций (универсальных, общепрофессиональных и профессиональных), число которых часто превосходит 30 разнообразных позиций, но оценить их в рамках одного аттестационного испытания невозможно. К этому выводу приводит не только здравый смысл, но и практические попытки многих университетских коллективов на начальном этапе реализации стандартов третьего поколения и оценивания компетенций. Позднее появились подходы, связанные с кластеризацией и паспортизацией компетенций, которые предусматривали укрупнение их по смыслу и по содержанию, что положительным образом сказалось на понимании процессов оценивания компетенций [73, 261]. С другой стороны, в действующих нормативных документах, регламентирующих порядок проведения итоговой аттестации по программам высшего образования, предусматривается две формы аттестационных испытаний – государственный экзамен и защита выпускной квалификационной работы. При этом государственный экзамен проводится по одной или нескольким дисциплинам (модулям) образовательной программы, результаты освоения которых имеют *определяющее значение для профессиональной деятельности* выпускников [189]. Таким образом, на уровне регулятора закладывается норма, согласующаяся со стремлением получить в рамках этого аттестационного испытания информацию об уровне компетентности выпускника и при этом соотнести знания, умения и опыт практической деятельности с требованиями, заложенными во ФГОС и профессиональных стандартах. В качестве результатов, имеющих определяющее значение для профессиональной деятельности, в первую очередь отдается предпочтение профессиональным компетенциям и тем дисциплинам и модулям, которые их формируют. Поэтому за содержательную область, подлежащую операционализации и конструированию для нее измерителей, выбираются соответствующие дисциплины учебного плана образовательной программы, которые иногда называют «ядерными» или базовыми. Содержание таких

измерителей должно включать весь запланированный к освоению обучающимися обязательный или необходимый объем. Основную часть заданий измерителей для аттестации делают значительно легче средней трудности, особенно при планировании процента аттестуемых, которые не пройдут за пороговый балл. Типичное распределение заданий по трудности представлено на рис. 14, где изображена классическая ситуация для аттестационного испытания. Если процент не аттестованных не должен быть больше 10 % и пороговый балл установлен на уровне 70 % (т.е. выпускники, которые выполнили меньше 70 % заданий, становятся неаттестованными), то в измеритель включают не менее 70 % легких заданий, с которыми справится 90 % выпускников [156].

Другими словами, каждый этап многостадийного оценивания носит критериально-ориентированный характер и должен иметь свой критериальный или пороговый балл, преодоление которого свидетельствует о соответствии выпускника определенному уровню компетентности.

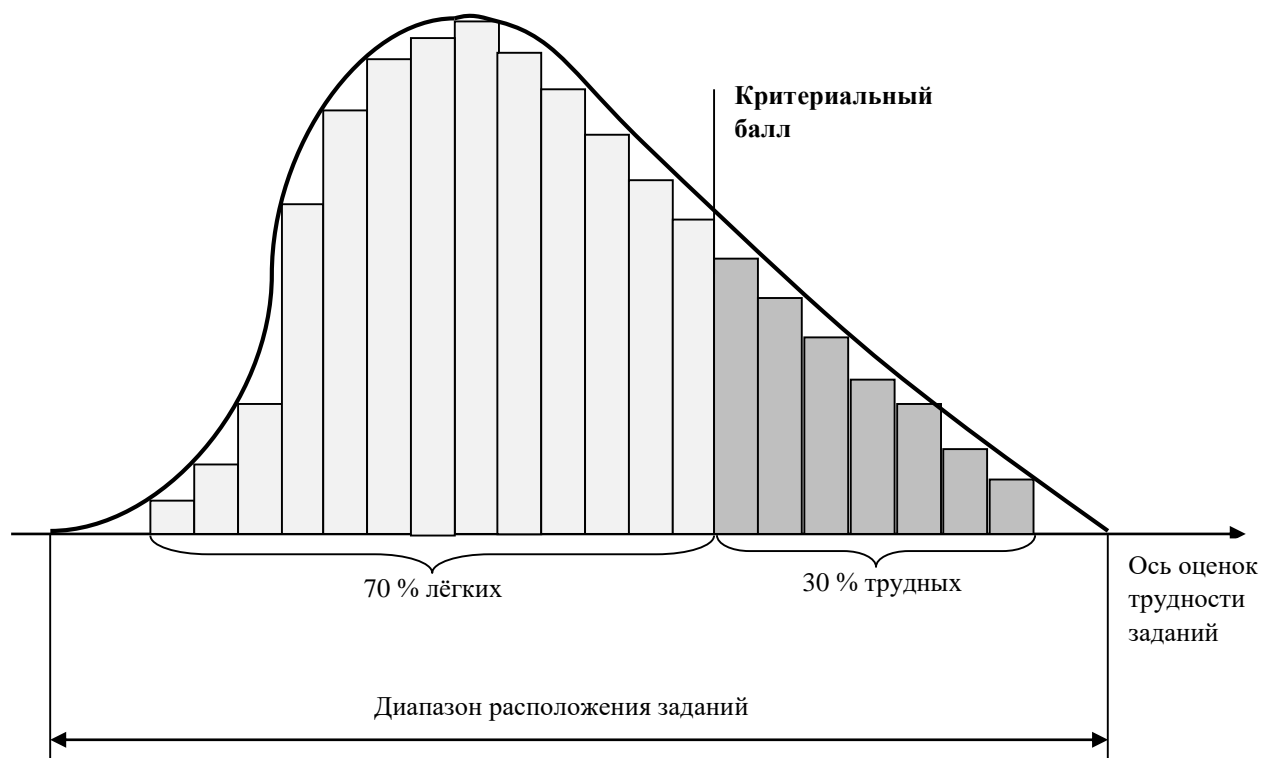


Рис. 14. Распределение заданий по трудности при критериально-ориентированном подходе

Дифференциация методов и инструментария для итоговой аттестации может быть проведена по различным основаниям. В случае нашего исследования, в

котором обоснован выбор в пользу бипарадигмального подхода к образовательным измерениям в условиях компетентностной трактовки результатов освоения выпускниками образовательных программ, такая дифференциация проводится сообразно уровням измерений. Первая группа методов и средств обеспечивает количественный уровень измерений, вторая – качественный (рис. 15).



Рис. 15. Представление методов и инструментария для многостадийного оценивания в соответствии с уровнями образовательных измерений

К группе, соответствующей количественному уровню измерений и таким же шкалам, относятся стандартизированные оценочные задания, проверка правильности выполнения которых может быть автоматизирована. К таким заданиям относятся задания закрытой формы (с выбором одного или нескольких правильных ответов, на установление соответствия или правильной последовательности, ситуационные задания с высокой степенью детерминированности контекста). Задания такого уровня, как правило, носят

репродуктивный и поисковый характер. В терминологии данного исследования – это задания минимального и базового уровней компетентности, т. е. при условии их успешного выполнения соответствующие оценкам «удовлетворительно» и «хорошо». Как отмечалось выше, при проведении обследований качества образования, условий реализации образовательных программ и удовлетворенности обучающихся, профессорско-преподавательского состава и работодателей целесообразно дополнять оценки выпускников по итоговой аттестации результатами стандартизированного анкетирования. Этот метод также относится к количественному уровню, а сами анкеты как инструментальный требуют при их разработке проведения апробации, доказательства надежности и валидности. При стремлении получения максимально разносторонней информации о результатах реализации образовательных программ в целях повышения качества образования и подготовки специалистов сочетание методов тестирования и анкетирования, конечно, будет приносить положительный эффект.

Вторая группа качественного уровня измерений включает такие методы и инструментальный, реализация которых осуществляется с участием экспертов. В первую очередь, речь идет об оценке результатов выполнения заданий качественного уровня измерений. В условиях итоговой аттестации такую оценку необходимо проводить не одному человеку, а группе экспертов. В этом и есть одно из предназначений создания экзаменационных комиссий для итоговой аттестации, когда классификационные решения и общая оценка уровня подготовленности студента и компетентности выпускника принимаются групповым методом, коллегиально. К качественному уровню следует отнести задания практических умений, кейс-измерители, ситуационные задания с низкой степенью детерминированности контекста, структурированное интервью, собеседование по профессиональной проблеме, портфолио.

Вне зависимости от уровней измерения можно сформулировать ряд требований к аттестационным процедурам:

– методы и инструментальный должны обеспечивать информацией, на каком уровне и в каком объеме освоена образовательная программа и компетенции;

– оценочные задания должны быть релевантны измеряемым образовательным результатам и быть согласованы с требованиями образовательных и профессиональных стандартов;

– по полученным результатам можно корректно зафиксировать прирост и динамику освоения образовательной программы на протяжении всего периода обучения;

– механизмы применения методов и средств оценивания не должны приводить к неоднозначным оценкам и нарушениям принципов многостадийного оценивания.

Определенный для каждого уровня измерений инструментарий требует уточнений в контексте рассматриваемой в диссертации проблемы. Так, необходимо уточнить понятие «тест» в рамках критериально-ориентированного подхода. Ранее было определено, что под тестом в современном понимании рассматривают совокупность различных форм заданий (не только с выбором одного или нескольких правильных ответов), обладающих системообразующими свойствами и устойчивыми характеристиками (прежде всего, трудность и дифференцирующая способность). Заявленный критериально-ориентированный подход требует уточнений для данного понятия. Поэтому под *тестом* на количественном уровне образовательных измерений для многостадийного оценивания будем понимать систему таких заданий, которые предъявляются в определенной стратегии и по алгоритмам, обеспечивающим надежность и содержательную валидность результатов по отношению к установленным стандартам выполнения конкретного этапа оценивания. Такой тест в процедурах итоговой аттестации является междисциплинарным, поскольку для измерения заявляется сразу несколько компетенций. Часть профессиональных компетенций, как правило, соотносится с трудовыми функциями, при условии, если результаты освоения образовательной программы проектировались с учетом требований профессиональных стандартов. Аттестационный тест, который в многостадийном оценивании предъявляется аттестуемому на первом этапе для измерения уровня компетентности выпускника на минимальном уровне (оценка

«удовлетворительно»)), в силу гетерогенности состоит из тестлетов – однородных по содержанию субтестов, служащих измерению отдельного конструкта. Количество таких тестлетов будет соответствовать числу выделенных базовых дисциплин (модулей) учебного плана образовательной программы, результаты освоения которых имеют определяющее значение для профессиональной деятельности выпускников [160, 326]. Высокий уровень значимости аттестационных испытаний, в которых используется такой инструментарий, требует серьезного подхода к разработке оценочных заданий, их обязательной апробации желательно на репрезентативных выборках, что не всегда становится возможным в силу разных обстоятельств. Такая апробация должна завершаться коррекцией оценочных заданий, повышая содержательную валидность инструментария и приводя его к профессиональному уровню.

Вторая категория оценочных заданий, относящихся к количественному уровню измерений и предусматривающих в ходе их выполнения не просто воспроизводство конкретных знаний и демонстрацию умений (репродуктивный уровень), а решение задач проблемно-поискового характера – это *ситуационные или профессионально-ориентированные задания*. Этот инструментарий используется на втором этапе многостадийного оценивания, и уровень его трудности соответствует базовому уровню компетентности выпускника. Следует отметить, что ситуационные задания являются частным случаем кейсов, более широкой категории оценочных средств, но их отличие состоит в степени детерминированности контекста [105]. По сравнению с кейсами, допускающими изменение контекста со стороны аттестуемых, ситуационные задания строго стандартизированы, если речь идет об автоматизированной проверке, но также носят междисциплинарный характер. Как и в случае разработки теста, перед созданием ситуационного задания необходимо определить число и характер конструктов (компетенций), подлежащих измерению, сформировать проблемный контекст, который имеет аутентичный характер и актуальный для того вида профессиональной деятельности, к которой готовился выпускник при освоении образовательной программы. Основная часть задания содержит текст с ситуацией



и контекстом из реальной профессиональной деятельности. В общем виде ситуация представляет собой совокупность событий, связанных в целое одной проблемой. Эти события объединены причинно-следственными связями, способствующими или препятствующими разрешению противоречия. Ситуация содержит явные и латентные противоречия и характеризуется контекстом, под которым понимаются граничные условия, описывающие количественно или качественно особенности действий, компонентов и участников ситуации. Профессионально-ориентированное задание предполагает ответ на несколько вопросов, которые направлены на оценку измеряемых профессиональных компетенций [144].

Оценочные рубрики, т. е. правила, по которым следует проверять правильность выполнения задания, должны содержать одинаковое число градаций в одном измерителе. Как правило, таких градаций вводят не более четырех. Экспертизу ситуационных заданий желательно проводить специалисту, который не участвовал в его разработке. Этот оценочный инструментарий также должен проходить апробацию на представительной выборке, чтобы повысить качество структуры и содержания ситуационных заданий.

После того как аттестуемый подтверждает компетентность на базовом уровне, что соответствует оценке «хорошо», он может выходить на следующий этап оценивания – высокого уровня компетентности (оценка «отлично»). Этот уровень предусматривает структурированное интервью или свободное собеседование на заданную тему (поставленную проблему). Кроме того, целесообразно добавлять к этим качественным методам оценивания также портфолио, выступающее одновременно и методом, и средством оценивания.

Интервью следует отнести к научно обоснованному методу сбора данных, который проводится на качественном уровне. Приписывание каких-либо чисел ответам аттестуемых не делает этот метод количественным и не предусматривает использование метрической шкалы. Интервью классифицируют на неструктурированные (открытые) и структурированные (закрытые). В силу специфики аттестационных испытаний, ограниченности временных ресурсов, возможности открытых диалогов с выпускниками на защите выпускной

квалификационной работы, на аттестационной процедуре в форме государственного экзамена, строящегося на многостадийном оценивании, предпочтение следует отдать структурированному интервью. Такой метод предполагает последовательность вопросов – от общего к частному, соответствующих предмету обсуждения. Для представления данных интервьюирования применяют неметрические шкалы (например, Лайкерта). Конечно, проведение интервью является трудозатратным и может оказаться малоэффективным, повышая субъективность при принятии классификационных решений и выставления итоговой оценки выпускнику. Кроме того, сами эксперты, выполняющие функции экзаменаторов, должны быть готовы к проведению такого метода оценивания. В этой связи целесообразнее рассматривать структурированное собеседование с выпускником, который прошел два этапа оценивания и подтвердил соответствие уровню базовой компетентности. Такое собеседование проводится по определенной профессиональной проблеме, которую может предложить как сам аттестуемый выпускник, так и экзаменационная комиссия. Дополнять ответы аттестуемого призвано портфолио, которое формируется самим обучающимся на протяжении всего периода обучения по образовательной программе. Метод портфолио и оценка его достаточно проработаны как в научном, так и методическом плане [40, 72, 87, 154 и др.]. Его следует рассматривать как дополнительную аргументацию при подготовке выводов об уровне компетентности выпускника на высоком уровне и выставлении оценки «отлично». Принято считать, что портфолио (или можно встретить название – паспорт профессиональных достижений) включает такие работы, которые раскрывают достижения его владельца, свидетельства самостоятельной работы. В связи с большими трудозатратами на анализ и оценку материалов портфолио, выпускники могут представить его заблаговременно в аттестационную комиссию. Такой метод аутентичного оценивания благоприятным образом сказывается на аттестационной процедуре высокой значимости, снижая психологическое напряжение аттестуемых.

Обобщая представленные в этом параграфе методы и соответствующие им средства, можно дополнить предложенную в параграфе 1.3 первой главы вертикальную шкалу для диапазона уровней компетентности релевантным им (уровням) инструментарием (рис. 16).

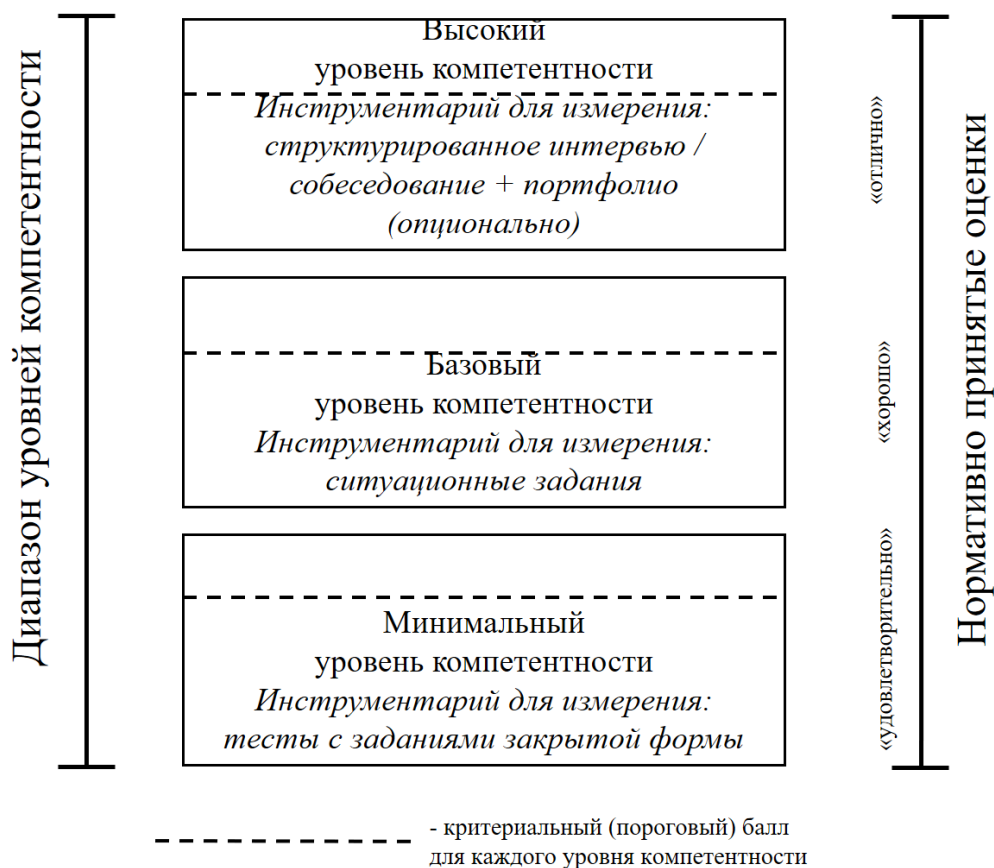


Рис. 16. Вертикальная шкала для диапазона уровней компетентности и инструментарий для измерения

## 2.5. Структурная модель реализации многостадийного оценивания уровня компетентности выпускников при итоговой аттестации

Термин «модель» широко используется в различных исследованиях и в практической деятельности, а в научной работе он необходим для представления или описания искусственно создаваемого образа изучаемого объекта или процесса. Другими словами, модель может рассматриваться как идеализированный результат

того, что исследуется, в данном случае многостадийное оценивание при итоговой аттестации выпускников. При этом, по А. М. Новикову, на основе ведущих научных идей должны появляться новые сочетания, механизмы связи между компонентами изучаемого процесса, показывающие способы организации практических действий и их результатов [182]. Выделяя познавательные и прагматичные модели, А. М. Новиков указывает, что последние «носят нормативный характер для дальнейшей деятельности, играют роль стандарта, образца, под который «подгоняется» в дальнейшем как сама деятельность, так и ее результаты» [183, с. 261]. При этом модель должна давать ответы на вопросы: Что будет? Как должно быть? И как добиться желаемого результата?

Конструируя такой прагматичный образец (модель) многостадийного оценивания при итоговой аттестации, необходимо учитывать требования согласованности, ясности и адекватности, которые предъявляются к моделям. В контексте рассматриваемого в нашем исследовании процесса требование согласованности должно выражаться в том, что многостадийное оценивание будет органично встроено в существующую систему итоговой аттестации в высшем образовании, не приводя ее функционирование к противоречиям. Ясность или простота модели заключается в том, что «чем проще модель, тем она ближе к моделируемой реальности и тем она удобнее для использования». Иными словами, для широкого применения многостадийного оценивания описываемая его модель должна быть понятно представлена и конкретизирована для большинства пользователей. Требование адекватности означает совпадение свойств модели и свойств моделируемого объекта или процесса. Степень этой адекватности можно подтвердить после реализации на практике модели многостадийного оценивания при анализе полученных результатов, соотносимых с заявленными целями.

В данном исследовании предлагается рассматривать модель многостадийного оценивания компетентности выпускников как способ представления аттестационной процедуры, включающей методологический, содержательный и технологический структурные компоненты, а также компонент принятия решений (рис. 17).

## СТРУКТУРНАЯ МОДЕЛЬ МНОГОСТАДИЙНОГО ОЦЕНИВАНИЯ УРОВНЯ КОМПЕТЕНТНОСТИ ВЫПУСКНИКОВ ПРИ ИТОГОВОЙ АТТЕСТАЦИИ В ВЫСШЕМ ОБРАЗОВАНИИ

**Цель:** получение надежных, валидных и аутентичных результатов итоговой аттестации выпускников

**Результат:** обоснованные классификационные решения об уровне компетентности выпускников при итоговой аттестации в высшем образовании



Рис. 17. Представление структурной модели многостадийного оценивания уровня компетентности выпускников при итоговой аттестации в высшем образовании

В представленной структурной модели многостадийного оценивания фиксируется целеполагание и ожидаемый от ее реализации на практике результат. Как и отмечалось выше, целью применения многостадийного оценивания в аттестационном испытании высокой значимости – государственном экзамене – является получение надежных (устойчивых), валидных (достоверных) и аутентичных (подлинных, соответствующих) результатов итоговой аттестации. Результатом достижения этой цели будут выставляемые аттестуемым оценки, которые соответствуют принятой пятибалльной шкале («отлично», «хорошо», «удовлетворительно», «неудовлетворительно»), на основе обоснованных классификационных решений об уровне компетентности выпускников. Такая классификация предполагает отнесение выпускника к одному из трех принятых уровней – минимальный, базовый или высокий, которые соотносятся с традиционными оценками. В случае, когда аттестуемый не достигает критериального (порогового) балла на первом этапе минимального уровня компетентности, он считается не аттестованным.

*Методологический компонент* модели включает:

1) компетентностный подход, концептуально определяющий цели, содержание и результаты профессионального образования, в том числе высшего, и задающий деятельностную направленность в подготовке будущих специалистов для сферы труда. Большое количество компетенций, которые определяются в образовательных стандартах и дополняются на основе профессиональных стандартов, с целью их измеримости требует проведения процедур кластеризации и операционализации (выражение их через признаки проявления в деятельностной форме);

2) бипарадигмальный подход к образовательным измерениям предусматривает сочетание процедур измерения компетенций на количественном и качественном уровнях при выборе определенного дизайна измерений. В случае итоговой аттестации предпочтение отдается последовательному дизайну – от количественного (строго стандартизированного) уровня и автоматизированной оценке результатов выполнения задания к качественному уровню (менее

стандартизированному) с привлечением экспертов для оценивания компетентности выпускника на высоком уровне;

3) теорию тестов, состоящую из классической и современной, а также теории генерализации. Включение положений теории тестов в методологический компонент связано с тем, что на аксиоматическом уровне вводится понятие ошибки измерения, неизбежно присутствующей в любых физических и нефизических измерениях и которую стремятся минимизировать, а также связанные понятия надежности и валидности результатов измерений и способы их повышения при использовании тестов в контрольно-оценочных мероприятиях.

*Содержательный компонент* представленной модели задает основу, на которую следует опираться, проектируя инструментарий для многостадийного оценивания. В роли такой основы выступают образовательные стандарты для высшего образования и профессиональные стандарты. Последние нужны для соотнесения требований сферы труда к выпускникам, которых готовят в сфере образования, и отражения этих требований прежде всего в результатах. Именно к профессиональным стандартам делается отсылка в образовательных стандартах в части определения профессиональных компетенций как одних из результатов освоения выпускником образовательной программы. Вместе с универсальными и общепрофессиональными компетенциями профессиональные компетенции образуют содержательную совокупность компетенций, которая подлежит тщательному анализу. При этом стандарты не должны ассоциироваться исключительно с содержанием образовательных программ, это полностью ответственность образовательной организации, реализующей такие программы. Стандарты, как отмечалось выше, это не стандарты содержания, а стандарты результатов. И если обратиться к классификации стандартов В. К. Загвоздкина, выделяющего максимальные, средние и минимальные стандарты, то действующие ФГОС ВО следует рассматривать как минимальные стандарты [78]. Другими словами, минимальные стандарты – это не минимум содержания, это фиксация обязательных требований, не исключая достижения более высоких результатов. То же самое в равной степени можно отнести и к профессиональным

стандартам, в которых на уровне формулировок заложена описанная концепция минимальных стандартов. Например, если рассматривать описание трудовых функций, то они представлены в категориях «трудовые действия», «необходимые умения», «необходимые знания». Данный компонент модели предусматривает декомпозицию компетенций, подлежащих измерениям на аттестационном испытании, и представление их в виде сущностных признаков проявления на установленных уровнях сформированности компетенций.

*Технологический компонент* модели, который описан в следующей главе, предусматривает выполнение определенных этапов работ, обеспечивающих получение надежных, валидных и аутентичных результатов многостадийного оценивания выпускников. Эти работы включают в себя разработку оценочного инструментария для аттестационного испытания в форме государственного экзамена, соблюдение условий по его реализации с обязательной апробацией и корректировкой, а также обеспечение надежности и валидности результатов. При итоговой аттестации на первый план выходит содержательная валидность, которая предусматривает серьезную экспертную деятельность после тщательной детализации инструментария и составления спецификации. Именно для конструирования оценочных заданий и необходимо проводить указанные выше работы по кластеризации и паспортизации компетенций, выделению базовых профессиональных дисциплин, содержание которых обеспечивает подготовку к выполнению трудовых функций и готовность к определенному виду профессиональной деятельности. В технологическую составляющую также входят регламент проведения самой аттестационной процедуры, требования к техническим и программным средствам, с помощью которых проводится предъявление заданий и автоматизированная оценка правильности их выполнения, а также методики шкалирования результатов. Шкалирование предусматривает перевод первичных результатов в производные показатели.

Из технологического компонента логически следует *компонент принятия решений*. Поскольку шкалированный результат, полученный аттестуемым, должен быть соотнесен с определенным стандартом выполнения каждой стадии (этапа)



оценивания – критериальным (пороговым) баллом. Определение такого стандарта связано с областью принятия обоснованных и ответственных решений, т.к. процедура итоговой аттестации – это процедура с высоким уровнем значимости, по результатам которой должны быть выставлены оценки по традиционной пятибалльной системе, выпускнику присвоена квалификация и выдан документ об образовании. При определении стандартов выполнения каждого этапа оценивания не может быть общего заданного критерия (порогового), поскольку существует специфика аттестации в каждом направлении подготовки и области профессиональной деятельности, к которой готовят обучающегося. Кроме того, процедуры внутренней и внешней оценки качества образования предусматривают обратную связь от разных субъектов (выпускники, работодатели, профессорско-преподавательский состав, администрация и др.), по результатам которой полученные данные должны стать основой для принятия и других решений – корректировки и улучшения образовательной программы, конкретных рабочих программ учебных дисциплин, практик, итоговой аттестации, условий реализации и пр. [39, 43, 149, 163].

Предлагаемая структурная модель многостадийного оценивания обладает свойствами полноты и системности, охватывая практически весь спектр задач, решаемых с помощью образовательных измерений для определения уровня компетентности выпускников при итоговой аттестации. Эта модель призвана восполнить существующий пробел в дидактике профессионального образования, который касается теоретико-методологического обоснования контроля и оценки в русле современных тенденций сочетания количественных и качественных методов измерений и отвечает вышеуказанным требованиям: ингерентности (согласованности), простоты (ясности) и адекватности [183]. Данная модель включает все основные структурные компоненты многостадийного оценивания при итоговой аттестации: методологические основания, содержание, технологию и получение выводов для принятия решений.

## Выводы по второй главе

Вопросы методологического обоснования получения надежных, валидных и аутентичных результатов аттестационных процедур для обоснованных классификационных решений об уровне сформированности компетентности выпускников предусматривают обоснование целей, функций, определения принципов, основ и подходов многостадийного оценивания, а также выделение методов и соответствующего инструментария для этого.

В число целей многостадийного оценивания входят как общие, носящие социально значимый характер для системы высшего образования и совершенствования системы подготовки кадров для экономики страны, так и специальные. Последние связаны с достижением объективности, обоснованности, нацелены на обеспечение повышения эффективности аттестационных процедур, создание ситуации успеха и благоприятной психологической атмосферы, снижающей невербальное давление экзаменаторов.

На многостадийное оценивание возлагаются оценочная, информационная, сравнительная и прогностическая функции. С одной стороны, эти функции традиционно присущи педагогическому контролю, но введение многостадийного оценивания в итоговую аттестацию с новыми возможностями вычислительных мощностей и обработки данных позволяет их реализовывать с большей эффективностью.

Принципы как центральная категория в методологическом аппарате многостадийного оценивания классифицируются по трем группам: нормативные (принципы объективности и независимости), общедидактические (принципы научности, релевантности, достоверности, справедливости) и специальные. К числу специальных принципов, повышающих состоятельность многостадийного оценивания, относятся: дифференциация, индивидуализация, последовательность, критериальность, открытость, сочетание количественного и качественного уровней образовательных измерения, кумулятивность.

Основу методологии многостадийного оценивания закладывают компетентностный подход к содержанию высшего образования и трактовке результатов освоения образовательных программ, бипарадигмальный подход к образовательным измерениям и теория тестов. Разработка инструментария и его использование в многостадийном оценивании требует опоры на специальные положения, обеспечивающие научное обоснование качества всех компонентов процесса, средств и результатов образовательных измерений. В обоснование качества сложность вносит латентный характер подлежащих измерению конструктов (компетенций). Поэтому интерпретация результатов измерений при итоговой аттестации без общения с выпускником, без предыстории его обучения может оказаться крайне опасной и привести к ложным выводам относительно уровня его компетентности. В этой связи необходимо сочетать количественный и качественный уровни измерений. Обработка, анализ и интерпретация результатов измерений различного уровня должна проводиться в строгом соответствии с правилами и группами допустимых операций, выделенных для каждого уровня измерения, с учетом специфики оцениваемых переменных. Невыполнение правил обработки данных может привести к несправедливой оценке выпускников и некорректным аттестационным выводам.

Положения обеих теорий – и классической, и современной – необходимы для реализации многостадийного оценивания. Классическая теория тестов обеспечивает проведение первичной апробации оценочного инструментария и начального этапа анализа результатов, а современная теория IRT существенным образом продвигает подходы в измерении латентных переменных за счет математического аппарата и предлагаемых моделей (однопараметрической, двух- и трехпараметрических, политомических). Современная теория тестов развивает положения классической теории, позволяя переходить на количественный уровень образовательных измерений. В ее основе лежат вероятностные модели, позволяющие предсказывать успех или неуспех выполнения аттестуемыми с разным уровнем подготовки тех или иных заданий различной трудности.

Предложенная структурная модель многостадийного оценивания обладает свойствами полноты и системности, охватывая практически весь спектр задач, решаемых с помощью образовательных измерений для определения уровня компетентности выпускников при итоговой аттестации. Эта модель призвана восполнить существующий пробел в дидактике профессионального образования, касающийся теоретико-методологического обоснования контроля и оценки в русле современных тенденций сочетания количественных и качественных методов измерений, и отвечает требованиям ингерентности (согласованности), простоты (ясности) и адекватности. Данная модель включает все основные структурные компоненты многостадийного оценивания при итоговой аттестации: методологические основания, содержание, технологию и получение выводов для принятия решений.

### **ГЛАВА 3. ТЕХНОЛОГИЯ МНОГОСТАДИЙНОГО ОЦЕНИВАНИЯ УРОВНЯ КОМПЕТЕНТНОСТИ ВЫПУСКНИКОВ ПРИ ИТОГОВОЙ АТТЕСТАЦИИ В ВЫСШЕМ ОБРАЗОВАНИИ**

#### **3.1. Разработка инструментария для многостадийного оценивания уровня компетентности выпускников**

Реализация представленной во второй главе методологии многостадийного оценивания предполагает наличие соответствующей технологии, обеспечивающей достижение поставленных целей итоговой аттестации. Понимая под технологией как ведущей формой организации деятельности, вслед за А. М. Новиковым, систему условий, форм, методов, средств и критериев решения поставленной задачи, в этой и следующей главах описана разработанная технология многостадийного оценивания уровня компетентности выпускников при итоговой аттестации [183]. Такая технология понимается как системная совокупность взаимосвязанных методик – от планирования всего процесса оценивания, проектирования и разработки оценочного инструментария для каждого этапа оценивания, формата реализации (компьютерного или бланкового, письменного или устного) до обеспечения надежности и валидности результатов, которые получаются при его использовании в аттестационном испытании, а также соблюдение организационно-педагогических условий многостадийного оценивания и процедуры установления критериальных (пороговых) баллов для каждого этапа. Соблюдение всей последовательности при реализации технологии многостадийного оценивания, включая корректное шкалирование и интерпретацию результатов, будет обеспечивать объективное определение уровня компетентности выпускников и принятие обоснованных классификационных решений об их аттестации. Такое понимание технологии многостадийного оценивания при итоговой аттестации делает его органической частью всей дидактической системы реализации образовательной программы – от замысла,

разработки компетентностной модели специалиста, проектирования результатов и соответствующего содержания через учебный план и рабочие программы дисциплин (модулей) и практик до оценки качества подготовки специалистов.

Исходя из принятой в данной работе уровневости оценивания компетентности, аттестационное испытание (государственный экзамен) должно предусматривать использование различного оценочного инструментария и соответствующих шкал, на которые будут помещаться результаты измерений и проводиться корректная интерпретация. Поэтому описание методики разработки инструментария следует вести в дедуктивной логике, представив сначала общие этапы создания оценочного инструментария, а затем провести детализацию от количественного уровня измерений к качественному.

Создание оценочного инструментария для процедуры высокой значимости предполагает более тщательную работу от авторов оценочных заданий и организаторов итоговой аттестации, чем это принято для текущего контроля и промежуточной аттестации. Эта работа не одного дня и даже месяца, поскольку само предназначение итоговой аттестации и ее содержание должны вобрать все значимое содержание образовательной программы, которое подлежит оцениванию через демонстрацию выпускником сформированных компетенций. Такие компетенции формируются, в первую очередь, базовыми профессиональными дисциплинами (модулями) и практиками, имеющими определяющее значение для будущей профессиональной деятельности выпускника. Поскольку таких дисциплин, как правило, 4–5 и больше, то и перечень проверяемых компетенций будет объемным. Помощью разработчикам заданий должна служить проведенная заранее паспортизация профессиональных компетенций, которые определяются каждым университетом самостоятельно, но с опорой на соответствующий профессиональный стандарт (при наличии). Анализ содержания различных образовательных программ показал, что такая работа по кластеризации и паспортизации компетенций проводилась большинством университетов при проектировании и разработке учебных планов и, как правило, в части профессиональных компетенций наблюдается большое разнообразие как в

описании индикаторов (признаков проявления) этих компетенций, так и в содержании, а также способах их формирования. Как отмечалось ранее, унификацию общего для всех содержательного ядра профессионального цикла и образовательных результатов возможно обеспечить при координирующей деятельности профессиональных сообществ, роль которых призваны выполнять федеральные учебно-методические объединения и другие заинтересованные в консолидации усилий по повышению качества образования структуры. В обобщенном виде сама идея описания компетенций, сопряженных с квалификационными требованиями или трудовыми функциями соответствующих профессиональных стандартов, через признаки их проявления может быть представлена в табличной форме следующим образом (табл. 5).

Таблица 5

**Обобщенная форма представления компетенций  
через индикаторы (признаки их проявления) на разных уровнях  
сформированности**

<b>Компетенция</b>	<b>Уровни сформированности компетенций</b>	<b>Описание индикаторов (признаков проявления) через трудовые действия, необходимые умения и необходимые знания</b>
Формулировка компетенции	Высокий	Способен критически оценивать и свободно применять ...
	Базовый	Владеет основными способами применения ...
	Минимальный	Способен понимать основные требования ... Способен проводить ... по стандартным процедурам Владеет понятийно-терминологическим аппаратом и выделяет знания ...

Образец макета паспорта компетенции, разработанного коллективом преподавателей и методистов Ивановского государственного университета при участии автора диссертации, представлен в Приложении 2. Одной из серьезных сложностей при разработке измерителей для итоговой аттестации является большой объем совокупности знаний, умений, способов деятельности, содержащийся в базовых профессиональных дисциплинах. В таком многомерном случае, когда измерению подлежат сразу несколько компетенций, конструируются тестлеты (субтесты), содержание каждого из которых соответствует либо предметной области отдельной дисциплины, либо отдельному конструкту (компетенции) [277].

Процесс разработки измерителя, с одной стороны, стандартный, и многие коллективы и отдельные преподаватели овладели способами конструирования оценочных заданий и формирования фондов оценочных средств для промежуточной и итоговой аттестации. С другой стороны, этот процесс может иметь и не завершающийся характер, если поставить цель непрерывного совершенствования оценочного инструментария и постоянного учета изменений различных условий и содержания подготовки специалистов. Вместе с тем, при разработке оценочных средств для итоговой аттестации принята определенная последовательность этапов, которая представлена ниже [6, 263, 316].

*Первый этап.* Целеполагание, определение перечня подлежащих измерению компетенций и содержательных областей.

*Второй этап.* Соотнесение требований соответствующего профессионального стандарта и трудовых функций со знаниями, умениями, способами деятельности, лежащими в основе формирования компетенций. В случае большого числа компетенций необходима работа по их укрупнению (кластеризация). Такая работа проводится еще на этапе проектирования самой образовательной программы, а позднее может проводиться корректировка паспортов компетенций.

*Третий этап.* Определение структуры измерителя, числа квазипараллельных вариантов и стратегии предъявления заданий в нем.



*Четвертый этап.* Создание спецификации измерителя, выбор числа субтестов (тестлетов) для измерения выбранных компетенций, определение времени на их выполнение.

*Пятый этап.* Разработка заданий, размещение их по вариантам в соответствии со спецификацией и их упорядочение согласно выбранной стратегии предъявления и предварительным оценкам трудности.

*Шестой этап.* Экспертиза формы и содержания заданий, входящих в инструментарий.

*Седьмой этап.* Корректировка и переработка формы и содержания заданий по результатам экспертизы.

*Восьмой этап.* Разработка инструкции для аттестуемых и экзаменаторов.

*Девятый этап.* Проведение апробации.

*Десятый этап.* Сбор, статистическая обработка эмпирических результатов, интерпретация полученных данных для повышения качества измерителя.

*Одиннадцатый этап.* Корректировка формы и содержания заданий на основании данных предыдущего этапа. Улучшение системообразующих свойств инструментария, оптимизация количества заданий, диапазона значений параметра трудности ( $\beta$ ) и времени на их выполнение на основании полученных оценок характеристик заданий.

*Двенадцатый этап.* Установление стандартов оценивания (критериальных (пороговых) баллов).

Представленные двенадцать этапов образуют определенный цикл разработки оценочного инструментария для итоговой аттестации, который укрупненно представлен на рисунке 18. Соблюдение всех этапов, как и непрерывность цикла крайне важны, если стоит задача достижения целей многостадийного оценивания, соблюдения заявленных принципов и получение объективных и сопоставимых оценок результатов освоения выпускниками образовательных программ.

Особенность предлагаемого многостадийного оценивания состоит в том, что на каждой стадии используются различные методы и виды инструментария, использование которых обеспечивает получение всесторонней оценки

профессиональной готовности выпускника. При этом планирование содержания инструментария должно базироваться на следующих принципах: *репрезентативности, значимости и системности* [156].



Рис. 18. Обобщенный цикл разработки оценочного инструментария для итоговой аттестации

Первый принцип – *репрезентативности* – регламентирует процедуру отбора содержания таким образом, чтобы обеспечить оптимальную полноту и правильность пропорций содержания измерителя. Полнота достигается представительным отображением совокупности компетенций, заложенных во ФГОС и учета требований профессиональных стандартов. Пропорции должны быть адекватны той иерархии кластеров, которая представлена в компетентностной модели подготовки специалиста и распределении трудоемкости дисциплин и их разделов, заложенной в учебных планах и образовательных программах.

Второй принцип – *значимости* – предписывает включать в оценочный инструментарий наиболее существенные элементы содержания, имеющие определяющее значение для будущей профессиональной деятельности

выпускника. Выделение таких ядерных или базовых частей требует структурирования содержания перед его отбором. Этот же принцип, трактуемый по отношению к компетенциям, также предполагает обращение к компетентностной модели подготовки специалиста, в которой установлена иерархия компетенций.

Третий принцип – *системности* – предполагает подбор заданий, содержание которых связано между собой определенной иерархией и общей структурой знаний, умений и способов деятельности. В условиях компетентного подхода системность означает отражение в аттестационном измерителе иерархии кластеров компетентностной модели подготовки специалиста. Таким образом, становится ясным ответ на вопрос, который нередко возникает у преподавателей при разработке основной образовательной программы – зачем структурировать компетенции и связывать кластеры структурно-описательной моделью, отражающей приоритеты в подготовке специалиста? Ответ на него достаточно прозрачен – компетентностная модель должна лежать в центре внимания при обучении и оценивании результатов освоения образовательной программы.

Для адекватного представления выбранных для измерения конструктов указанная ранее паспортизация компетенций проводится с помощью процедуры операционализации. Чем выше будет достигнут уровень операционализации конструктов, тем адекватнее в содержании заданий будут отражены измеряемые переменные [156]. Для этого разработчики измерителей (оценочного инструментария) могут использовать одну из существующих таксономий. В Приложении 3 представлены четыре варианта различных таксономий, которые могут быть выбраны или адаптированы под конкретную цель оценивания при описании действий аттестуемого выпускника.

Процесс планирования содержания выполняется с помощью *спецификации*. В спецификации, которую предлагается разрабатывать в форме таблицы, содержится примерная раскладка процентного соотношения компетенций (кластеров компетенций) и содержательных разделов (дисциплин или модулей), включенных в измеритель. В этом документе определяется необходимое число

заданий, выбранное для оценивания в соответствии с важностью и трудоемкостью освоения компетенций или содержательных разделов. Раскладку начинают с подсчета априорно планируемого числа заданий, которое затем будет неоднократно меняться в процессе работы над оценочным инструментарием и совершенствования его характеристик в сторону увеличения или уменьшения. Спецификация может быть представлена в развернутой форме, включать указания форм заданий, которые будут использоваться для оценки результатов обучения, время выполнения заданий, правила подсчета правильных ответов, баллов и т. д. Пример структуры такой спецификации приведен в Приложении 4. Спецификация крайне важна при создании параллельных вариантов, необходимых в условиях массового оценивания, поскольку в ней закрепляется структура содержания, обязательная для всех вариантов измерителя.

Спецификация может иметь краткую или расширенную форму. В случае расширенной формы этот документ для многостадийного оценивания включает в себя следующие разделы:

1. Цель создания измерителя, описание подхода к его созданию и смежные цели.
2. Перечень нормативных документов (образовательных программ, образовательных и профессиональных стандартов и др.) и перечень базовой литературы, с опорой на которую предполагается создавать измеритель.
3. Описание общей структуры измерителя, перечень этапов с описанием видов используемого инструментария либо перечень тестлетов (субтестов).
4. Число заданий различной формы с указанием правил оценивания (дихотомическая или политомическая оценка).
5. Число параллельных вариантов измерителя.
6. Вес каждого задания, рекомендуемый его автором (если вводятся весовые коэффициенты при подсчете общего балла по измерителю), в общей оценке результатов выполнения оценочного средства.

7. Установленное время выполнения измерителя, в том числе время для каждого субтеста или стадии оценивания, среднее время выполнения одного задания с учетом специфики формы.

8. Соотношение заданий по различным разделам дисциплин, модулям, компетенциям или трудовым функциям профессиональных стандартов (в форме таблицы).

9. Рекомендации по контингенту испытуемых для апробации измерителя.

10. Перечень требований образовательных программ, образовательных или профессиональных стандартов, не охваченных в измерителе.

11. Рекомендуемая стратегия расположения заданий в измерителе, выстроенная на основе априорных оценок их трудности.

Таким образом, развернутая спецификация для аттестационного испытания высокой значимости является своеобразным паспортом измерителя, содержащим полную информацию о его свойствах и предназначении. Желательно, чтобы такая форма была принята в университете или любой другой организации, проводящей аттестацию на основе многостадийного оценивания. В какой-то мере разработанные многими университетами паспорта компетенций к каждой образовательной программе являются базисом для паспортов измерителей компетенций. Практический опыт работы автора диссертации в данном направлении показал, что целесообразно в рамках функционирования внутренней системы оценки качества образования (ВСОКО) организовать систему сертификации измерительных материалов. Такая система будет служить сразу нескольким целям, помогая развивать оценочную культуру и технологию, с одной стороны, и преодолевать присущую системе образования инертность, с другой стороны. Кроме того, проводимые в рамках процедур ВСОКО диагностические работы выступают той самой содержательной основой для апробации оценочного инструментария и формирования банка калиброванных заданий.

Краткая форма спецификации является частью развернутой версии и создается как промежуточный рабочий документ для авторов заданий и экспертов. Она имеет вид таблицы для сопряжения системы знаний, умений, компетенций или

трудовых функций с процентным соотношением заданий по различным разделам или содержательным модулям (п. 8 развернутой спецификации).

Спецификация содержания является концептуальной основой, которая очерчивает целевую область содержания измерителей для итоговой аттестации. Она считается обязательной частью работы авторов при создании аттестационных измерителей, поскольку включает их содержательный план с учетом различных ограничений временного или процедурного характера. При разработке спецификации важно, чтобы авторы заданий принимали во внимание четыре компонента, обеспечивающие формирование качественной структуры содержания измерителя. Эти компоненты включают: тематические единицы содержания, трудность заданий, диапазон охвата содержания и пропорции, которые определяются в соответствии со значимостью содержательных элементов [6, 156, 219].

Идентификация содержания измерителя, предназначенного для оценивания запланированных переменных (знаний, умений, способов деятельности, компетенций или готовности к выполнению трудовых функций), нетривиальная задача. С одной стороны, важен научно обоснованный подход к разработке спецификации, понимания психолого-педагогических закономерностей и особенностей формирования компетенций у обучающихся. С другой стороны, важно удерживать нормативную рамку при процедуре планирования, т. е. принимать во внимание требования ФГОС и профессиональных стандартов. И, наконец, способность преподавателей и экспертов беспристрастно и с высокой степенью объективности давать и принимать оценки качества разработанных заданий. Более того, качественная спецификация является необходимым условием обеспечения содержательной валидности результатов оценивания, что будет рассмотрено в следующих параграфах данной главы.

Для проведения аттестационного испытания на основе многостадийного оценивания в спецификации должна быть отражена примерная раскладка процентного соотношения субтестов по отдельным переменным измерения, структура содержания субтестов, проверяемые компетенции и необходимое число

заданий по каждому субтесту с указанием их формы. Пример такой спецификации для 4-х субтестов, предназначенных для оценивания уровня освоения четырех компетенций одного кластера, приведен в табл. 6 [87, 158]. В данной таблице показано, как могут распределяться задания по количеству и их форме в субтестах. Таким образом, можно убедиться, что для оценивания уровня сформированности первой компетенции в спецификации предусмотрено 42 задания, из чего можно сделать вывод, что вклад ее в общей компетентностной модели является существенным на фоне других компетенций. Такая работа синхронизируется с определением приоритетности компетенций и вклада тех или иных значимых способностей для подготовки соответствующего специалиста. Весовые коэффициенты (значимость) обычно выбираются экспертным путем в соответствии с компетентностной моделью подготовки специалиста, представляющей собой схематическое изображение связи компетенций, их иерархии и взаимного влияния.

Резюмируя эту часть – разработку спецификации, следует подчеркнуть, что это документ, который требует регулярного обновления и оптимизации. Такая работа должна проводиться после каждой процедуры итоговой аттестации, поскольку получаемые данные о результатах выполнения выпускниками тех или иных заданий, выводы об уровне сформированности отдельных компетенций при корректном их сравнении с учетом дополнительных факторов выступают основаниями для совершенствования как самого этого регламентирующего документа, так и разработанных на его основе измерителей.

Оценочные задания, входящие в состав инструментария, следует соотносить с этапом оценивания и уровнем сформированности компетентности, как было в обобщенном виде предложено во второй главе (рис. 16).

Пример распределения форм заданий по компетенциям и субтестам

Кластер компетенций (4 компетенции)	Измеритель				Итого
	Субтесты (содержательные разделы, модули, дисциплины)				
	Субтест 1	Субтест 2	Субтест 3	Субтест 4	
	20 заданий – 20 %	30 заданий – 30 %	40 заданий – 40 %	10 заданий – 10 %	
1-я компетенция	14 заданий	Не оценивается	18 заданий	10 заданий	42
	0 – с выбором ответа		10 – с выбором ответа	0 – с выбором ответа	
	10 – с краткими дополняемыми ответами		0 – с краткими дополняемыми ответами	3 – с краткими дополняемыми ответами	
	4 – с развернутыми ответами		3 – с развернутыми ответами	5 – с развернутыми ответами	
	0 – проблемные задания с наиболее правильным ответом		5 – проблемные задания с наиболее правильным ответом	2 – проблемные задания с наиболее правильным ответом	
2-я компетенция	9 заданий	10 заданий	Не оценивается	10 заданий	29
	5 – с выбором ответа	5 – с выбором ответа		0 – с выбором ответа	
	0 – с краткими дополняемыми ответами	0 – с краткими дополняемыми ответами		2 – с краткими дополняемыми ответами	
	2 – с развернутыми ответами	3 – с развернутыми ответами		3 – с развернутыми ответами	
	2 – проблемные задания с наиболее правильным ответом	2 – проблемные задания с наиболее правильным ответом		5 – проблемные задания с наиболее правильным ответом	
3-я компетенция	.....	.....	.....	.....	19
4-я компетенция	.....	.....	.....	.....	10
<i>Итого</i>					<i>100</i>



В данном параграфе опишем релевантный для каждой стадии тип оценочных заданий. Для краткости изложения здесь не приводятся примеры, классические ошибки и прочие методические аспекты разработки самих заданий, поскольку они достаточно подробно представлены в различных учебно-методических и учебных изданиях отечественных и зарубежных авторов, включая автора диссертации [6, 72, 116, 154, 156, 316, 365 и др.].

***Первая стадия оценивания. Количественный уровень образовательных измерений: задания закрытой формы для оценивания минимального уровня компетентности.***

В отечественной и зарубежной методической литературе выделяют четыре классических типа заданий закрытой формы:

- задания с выбором ответа, в которых следует выбрать правильный ответ из данного набора ответов;
- задания с конструируемым ответом (кратким регламентированным или развернутым в произвольной форме), предполагающие самостоятельное получение ответа;
- задания на установление соответствия, выполнение которых связано с выявлением соответствия между элементами двух множеств;
- задания на установление правильной последовательности, в которых требуется указать порядок элементов, действий или процессов, перечисленных в условии.

Вне зависимости от типа к оценочным заданиям предъявляются общие требования:

- каждое задание имеет свой порядковый номер, который может изменяться после статистической оценки трудности задания и выбора стратегии предъявления заданий;
- каждое задание имеет метку, которая присваивается заданию в момент его создания и остается неизменной при любой коррекции теста;
- каждое задание имеет эталон правильного ответа, в том числе задания с развернутым конструируемым ответом;

– в задании все элементы располагаются на четко определенных местах, фиксированных в рамках выбранной формы;

– для заданий разрабатывается стандартная инструкция по выполнению, которая не меняется в рамках каждой формы и предваряет формулировку заданий;

– для каждого задания разрабатывается правило выставления дихотомической или политомической оценки, общее для всех заданий одной формы и сопровождающееся инструкцией по проверке со стандартизированными процедурами подсчета сырых (первичных) баллов по тесту.

В целом выбор типа заданий и число вариантов зависят от целей и содержания оценивания, поэтому при итоговой аттестации обоснованно включать в измеритель больше заданий с выбором ответа, поскольку они благодаря высокой технологичности и автоматизированным процедурам проверки позволяют оценить большой объем содержания, что способствует росту надежности и содержательной валидности результатов образовательных измерений. Это вовсе не противоречит компетентностному подходу, когда в оценочной процедуре стремятся исключить такие задания, отдавая предпочтение открытым заданиям с развернутым ответом. Дело в том, что несмотря на относительно инновационную трактовку качества образования в компетентностной парадигме и формулировки требований к результатам освоения образовательных программ в виде совокупности компетенций во ФГОС, знания и умения остаются дидактически значимыми для целей обучения, поскольку они являются результатами текущего учебного процесса и закладывают фундамент для формирования компетенций. В связи с этим, при разработке измерителей есть смысл использовать все типы заданий, ориентируясь при их выборе на этап оценивания, специфику содержания, уровень ответственности принимаемых управленческих решений и т. д. Преимущества и недостатки различных типов заданий представлены ниже [87, 158].

### 1. Задания с двумя ответами.

*Преимущества:* благодаря краткости позволяют охватить большой объем материала, легко разрабатываются (только один дистрактор<sup>4</sup>), быстро с высокой объективностью обрабатываются результаты выполнения.

*Недостатки:* стимулируют механическое запоминание, поощряют угадывание, требуют увеличения количества заданий и, соответственно, времени тестирования для компенсации эффекта угадывания.

### 2. Задания с выбором из четырех – пяти ответов.

*Преимущества:* подходят для самых различных дисциплин, в силу краткости формулировок можно охватить большой объем содержания, обеспечивают возможность автоматизированной проверки и высокую объективность оценок аттестуемых, позволяют провести развернутый статистический анализ своих характеристик, скорректировать их и значительно повысить надежность образовательных измерений.

*Недостатки:* требуют значительной работы авторов при выборе дистракторов и коррекции баллов аттестуемых, не годятся для проверки продуктивного уровня деятельности и когнитивных умений.

### 3. Задания с конструируемыми регламентированными ответами.

*Преимущества:* просты в разработке, исключено угадывание, частично подходят для автоматизированной проверки.

*Недостатки:* проверяют, в основном, знание фактологического материала или понятийного аппарата, обычно легкие, иногда приводят к неоднозначно правильным и частично правильным ответам.

### 4. Задания со свободно конструируемыми ответами.

*Преимущества:* позволяют оценивать сложные конструкты, в том числе уровень сформированности компетенций, легко формулируются как традиционные задания, исключают угадывание.

---

<sup>4</sup> Дистрактор – вариант ответа на оценочное задание, близкий к правильному, но не являющийся таковым. Дистракторы должны иметь одинаковую форму (вербальную или невербальную), длину (быть в равной степени детальными или обобщенными), трудность, должны быть выражены одной частью речи и представлять вместе с правильным ответом одну область знаний.

*Недостатки:* требуют длительной ресурсозатратной процедуры проверки и значительного времени выполнения, не позволяют охватить значительный объем содержания, снижают надежность образовательных измерений.

#### 5. Задания на установление соответствия.

*Преимущества:* просты в разработке, идеально подходят для оценивания ассоциативных знаний, уменьшают эффект угадывания.

*Недостатки:* чаще всего используются при проверке репродуктивного уровня деятельности и алгоритмических умений, громоздки по форме представления.

***Вторая стадия оценивания. Количественный уровень образовательных измерений: профессионально-ориентированные (ситуационные) задания.***

Профессионально-ориентированные (ситуационные) задания часто отождествляют с кейсами, которые широко применяются в обучении и оценивании. Строго говоря, ситуационные задания являются составной частью образовательных кейсов. И те, и другие следует относить к оценочному инструментарию именно для измерения уровня профессиональной готовности обучающихся или выпускников [105, 116, 135]. Однако в условиях высокой значимости оценивания, когда стремятся минимизировать случайные ошибки измерения и повысить надежность и валидность результатов, способствовать этому будут именно ситуационные задания со строго детерминированным содержанием в отличие от кейс-измерителей, допускающих неоднозначные варианты решений, вариативность начальных и граничных условий контента и экспертные способы оценивания.

Для ситуационных заданий оценочные рубрики строго регламентируют выбор баллов экспертами, указывая число баллов в зависимости от степени правильности и полноты ответов, поскольку сама формулировка заданий имеет фиксированный характер. В силу неизбежного существования определенного количества неизвестной информации ситуационные задания

должны содержать ровно столько комментариев и пояснений, сколько нужно для понимания проблемы. Вместе с тем факты в заданиях должны носить правдоподобный, по возможности достоверный характер, чтобы не провоцировать сомнения или обсуждения. Как правило, основу при разработке ситуационного задания составляет отдельный эпизод из практического опыта профессиональной деятельности. К этому эпизоду или проблемной ситуации составляется от 3 до 5 вопросов, каждый из которых может иметь один или несколько оптимальных ответов в различной форме. Такой подход способствует повышению аутентичности и значимости профессиональной проблемы, сформулированной в ситуационном задании, ориентирован на оценивание именно профессиональных способностей и умений, лежащих в основе формирования тех компетенций, которые потребуются выпускникам при осуществлении трудовой деятельности по специальности.

К числу основных этапов разработки ситуационных заданий для многостадийного оценивания в итоговой аттестации относятся:

1. Выбор цели создания, предмета или предметов оценивания. Постановка проблемы. Определение границ контекста для проблемы. Идентификация основных вопросов, на которые должен ответить испытуемый (например, обозначение временных границ для контекста и вопросов, выбор ограничений по объему на доказательную базу и источники информации, определение организации или географической области, указание вида доказательств, которые должны быть собраны, определение приоритетов для анализа при ответах на вопросы и т.д.).

2. Разработка дизайна ситуационного задания.

3. Определение критериев для оценивания качества ситуационного задания и методов оценивания (внутренняя и внешняя согласованность, прогностическая валидность, надежность). Разработка методики содержательной экспертизы. Проведение экспертизы качества содержания и структуры ситуационных заданий.

4. Разработка оценочных рубрик и методик для экспертов.
5. Апробация ситуационных заданий и оценочных рубрик на выборке испытуемых.
6. Обработка данных, анализ результатов апробации, разработка рекомендаций по коррекции ситуационных заданий и оценочных рубрик.
7. Коррекция ситуационных заданий и оценочных рубрик.
8. Разработка методики шкалирования для использования ситуационных заданий в итоговой аттестации и формата протокола для сообщения результатов испытуемым (установление стандарта (критериального (порогового) балла).
9. Разработка методических рекомендаций для различных пользователей.

Конструирование ситуационных заданий строится на сочетании требований образовательных и профессиональных стандартов. Для этого используется определенный набор профессиональных компетенций из образовательной программы и конкретные трудовые функции из профессионального стандарта.

Разработка содержания эффективных ситуационных заданий имеет два направления: с одной стороны, она включает шаги по созданию типичных профессиональных ситуаций с описанием рабочего места и значимых вариантов ответов, характерных для определенной профессиональной деятельности, с другой стороны, предполагает выполнение всех шагов, необходимых для обеспечения качественного содержания заданий оценочного инструментария. Качество содержания таких заданий зависит в первую очередь от успешности формулирования ситуационной основы задания и вариантов ответов. Для применения в условиях оценивания высокой значимости, когда результаты выполнения задания во многом определяют решение об аттестации выпускника и выводах об уровне сформированности компетентности, необходимо выполнить дополнительные шаги, чтобы гарантировать, что задания и шкалы имеют адекватные свойства и

соответствуют всем требованиям к их качеству. Как отмечалось выше, при разработке профессионально-ориентированных заданий необходимо выполнить ряд этапов. Для выбора ситуации необходимо провести анализ содержания деятельности, представленной в формате трудовых функций профессиональных стандартов или совокупности компетенций. Как правило, к такому анализу привлекают экспертов, которые предлагают примеры ситуаций для содержания заданий на суждения. При разработке заданий для экзаменов высокой значимости экспертиза предшествует более глубокому анализу качества содержания, основанному на статистических методах и факторном анализе. Результаты проведенного анализа должны подтвердить, что содержание задания соответствует цели его создания, ответы коррелируются с условием, а ложные суждения, выполняющие роль дистракторов, правдоподобны.

Ситуации должны носить проблемный характер, поэтому для их формулирования используют анализ критических инцидентов, характерных для данного вида профессиональной деятельности. Ситуации, предложенные экспертами, подвергаются анализу, чтобы определить основные темы для разработки содержания заданий на суждения. В соответствии с ситуацией формируется набор качеств специалиста и характеристик результатов его работы. Как только этот набор качеств и характеристик получен, он вместе с ситуацией выбирается в качестве основы задания, к которой разрабатываются ответы.

Проблемность профессионально-ориентированному заданию придает включение в его описание дилеммы, когда аттестуемому предлагаются суждения на выбор в форме вариантов ответа. Ситуация должна выпукло подчеркивать проблемы профессиональной деятельности, гарантируя, что предлагаемые варианты ответа способны полностью или частично решить дилемму. Формулировки ответов могут представлять собой суждения, которые включают одно или несколько истинных среди остальных ложных, а также более или менее верные способы решения дилеммы.

Инструкции к ситуационным заданиям на суждения могут принимать две формы. Одна из форм носит категоричный характер и требует от испытуемого однозначного выбора, когда каждое суждение расценивается как истинное или ложное. Другая форма инструкции нацеливает на выбор предпочтительного, наиболее верного ответа или наименее верного ответа. Обе инструкции важны, но они служат различным целям. На экзамене высокой значимости первая инструкция предназначена для проверки понимания экзаменуемым того, как оценить каждое суждение и сделать такой выбор, который позволит решить проблемную ситуацию. Вторая инструкция обычно используется в тех случаях, когда есть более и менее правильные ответы, которые сформулированы применительно к проблемам в профессиональных ситуациях.

В первом случае используется стандартная форма инструкции «Классифицируйте каждое суждение как истинное или ложное». Во втором случае таких инструкций может быть две: «Выберите наиболее правильный ответ» или «Выберите наиболее правильный и наименее правильный ответы». Последний формат инструкции предлагает более высокую степень детализации в оценке понимания испытуемым требований ситуации и предполагает выбор двух ответов. На экзаменах высокой значимости в любом случае полностью правильный ответ оценивается одним баллом, в противном случае испытуемый получает ноль баллов.

Вместе с содержанием таких заданий необходимо также решить вопрос об их количестве и времени, отводимом на их выполнение. Учитывая многостадийность оценивания и то обстоятельство, что профессионально-ориентированные задания, – это задания, соответствующие базовому уровню сформированности компетентности выпускника, т. е. выше минимального удовлетворительного, их количество может варьироваться от 5 до 12. В каждом отдельно взятом случае итоговой аттестации по каждому профилю подготовки будет свое количество таких заданий. Это зависит от множества факторов. Но на первый план выходит количество переменных, которые



стремятся охватить при измерении конструкторов. Окончательное решение о количестве ситуационных заданий, включенных во второй этап оценивания, соответствующего оценке «хорошо», принимается после экспертного обсуждения и анализа данных апробации.

Апробацию ситуационных заданий, например, целесообразно проводить в рамках мероприятий по функционированию внутренней системы оценки качества (ВСОКО). При проведении диагностических работ, которые являются обязательной частью ВСОКО, включаются профессионально-ориентированные задания, затем проводится анализ их выполнения с целью включения в фонд оценочных средств для итоговой аттестации. В результате такой апробации определяется и время, которое следует отвести на выполнение каждого ситуационного задания, и в целом второго этапа оценивания. Сложившаяся практика в контрольно-оценочных процедурах предусматривает среднее время для ответа на одно задание от двух до пяти минут в зависимости от трудности этого задания. В целом значительная продолжительность итогового экзамена отражает природу оценочной процедуры высокой значимости, а также важность решений, принимаемых по его результатам об аттестации выпускника, и выдаче ему документа об образовании с присвоением квалификации.

Соблюдение перечисленных условий призвано породить от выпускника профессиональное решение профессиональной ситуации. Если, как отмечалось в начале параграфа, профессиональные компетенции изначально формулировались с учетом обобщенных трудовых функций, трудовых функций и трудовых действий специалиста из соответствующего профессионального стандарта, то для такой работы авторам заданий достаточно паспортов компетенций. И тем самым можно утверждать, что профессионально-ориентированные (ситуационные) задания обеспечивают аутентичные результаты оценивания уровня компетентности выпускника. Шаблон структуры для профессионально-ориентированного (ситуационного) задания может быть таким, как показано на рисунке 19.

**Структура профессионально-ориентированного (ситуационного) задания**

1. Инструкция – пояснительная записка.
2. Основная часть – формулировка ситуации.
3. Вопросы или задания к основной части  
(рекомендуемое количество: не менее 3 и не более 5).
4. Оценочные рубрики по проверке выполнения заданий

*Рис. 19.* Шаблон структуры профессионально-ориентированного (ситуационного) задания

По принятой в исследовании уровневой иерархии данный инструментарий по своей трудности должен соответствовать базовому уровню компетентности (оценке «хорошо»). Это условие на этапе разработки ситуационных заданий является одним из непростых для авторов заданий. Поэтому экспертиза качества структуры и содержания заданий в обязательном порядке должна проводиться специалистом или преподавателем, который не участвовал в его создании. Эксперты рассматривают и оценивают задания с позиций того, как и насколько оно будет отличать компетентное поведение выпускника от некомпетентного в различных профессионально-ориентированных ситуациях.

Как в случае заданий репродуктивного уровня, так и в ситуационных заданиях возникает вопрос, какого уровня трудность должна быть и каким образом ее определить. Ответ на этот вопрос дается в следующей четвертой главе при определении организационно-педагогических условий реализации многостадийного оценивания и решении задачи установления критериального (порогового) балла для аттестации выпускников на каждой стадии оценивания.

В целом при разработке оценочных заданий разного типа следует руководствоваться требованиями, которые сформулированы в Приложении 5,

а для формулировки заданий можно использовать специальный конструктор, предложенный в [229] (Приложение 6).

*Третья стадия оценивания. Качественный уровень образовательных измерений: структурированное собеседование или интервью.*

В рамках предлагаемой авторской методологии многостадийного оценивания третий этап проходит на качественном уровне образовательных измерений, что предусматривает экспертное участие в определении высокого уровня сформированности компетентности выпускника и выставлении оценки «отлично». Данная третья стадия оценивания, завершающая аттестационное испытание, – государственный экзамен, который реализуется в форме структурированного собеседования или интервью по определенной профессиональной теме. Такая организация оценивания позволяет актуализировать не только профессиональные и общепрофессиональные компетенции выпускника, но и универсальные компетенции. В частности – критическое мышление, коммуникацию, способность выпускником продемонстрировать свои образовательные и профессиональные предпочтения при выборе проблемы, отборе материала, рефлексии и т. д. Другими словами, на данном этапе происходит более глубокая индивидуализация оценивания, поскольку в ходе собеседования с экспертами, роль которых выполняют члены экзаменационной (аттестационной) комиссии, каждый аттестуемый, вышедший на данный этап после прохождения двух предыдущих, обосновывает и аргументирует свою позицию и точку зрения как компетентный специалист по обсуждаемой теме.

Структурированное собеседование проводится по заранее разработанным и апробированным кейсам. Эти этапы повторяют этапы разработки ситуационных заданий. Как показано выше, кейсы являются не столь детерминированными по условиям, допускают различные варианты решений, но требующие серьезной аргументации. Кейсы содержат дополнительную справочную информацию, позволяющую учитывать контекст самой задачи. Кроме того, эксперты могут уточнять позицию и

предлагаемое решение у аттестуемого выпускника, формируя тем самым свое оценочное суждение об уровне компетентности выпускника. Оценивание на данном этапе происходит по установленным критериям, которые заранее должны быть доведены как для аттестуемых, так и экзаменаторов. Структурированным такое собеседование называется, потому что аттестуемому выпускнику предлагается структура ответа по задаче. В рамках проводимой опытно-экспериментальной работы для итоговой аттестации на государственном экзамене выпускникам образовательной программы «Психология образования» по направлению подготовки 44.03.02 Психолого-педагогическое образование была выбрана следующая структура для ответа по решаемому кейсу (рис. 20).

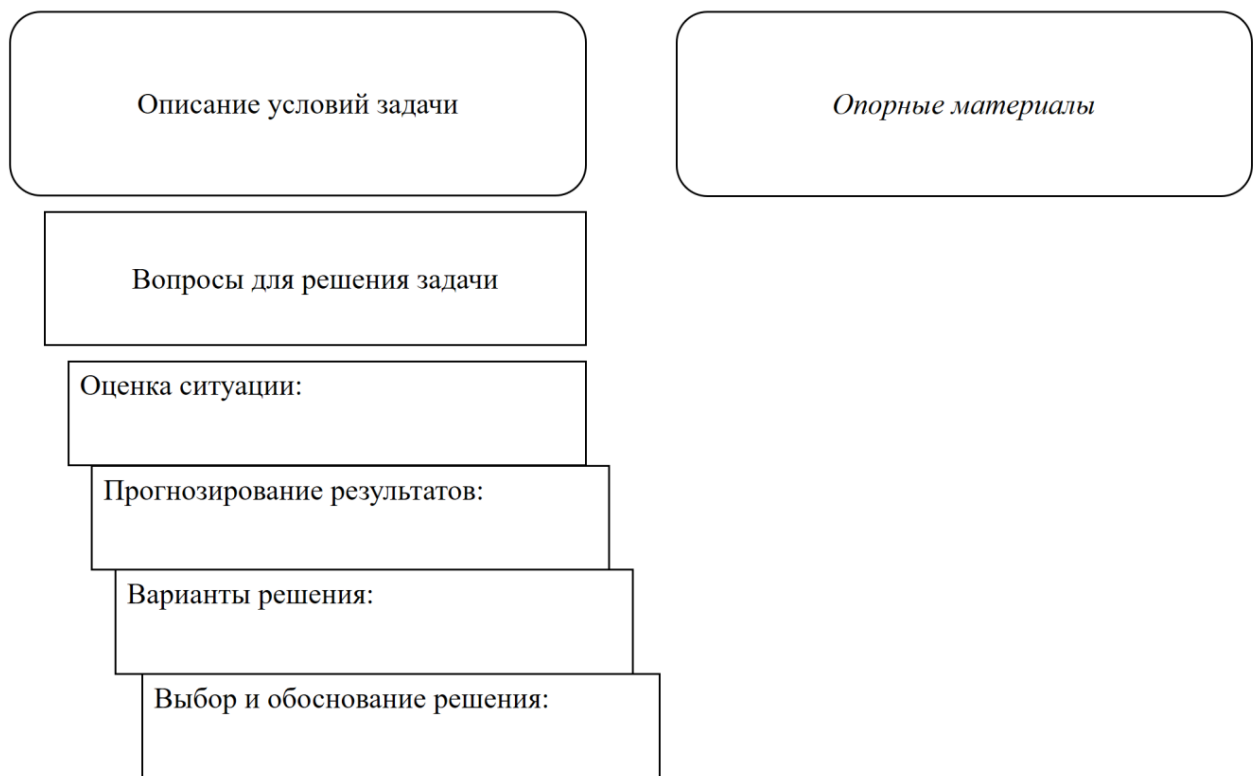


Рис. 20. Общая структура для собеседования по задаче на третьей стадии оценивания

Оценка решения проводится экспертами, в роли которых выступают председатель и члены экзаменационной комиссии, по критериям,

зафиксированном в оценочном листе (чек-лист) (табл. 7). Различия в оценках экспертов, произошедшие вследствие нечетких формулировок критериев, необходимо устранить до их применения в процессе аттестации. Перед началом работ по минимизации вариаций оценок необходимо выделить специальную группу экспертов, которая обобщит мнения остальных экспертов.

В данное время процедура образовательных измерений может быть предельно стандартизирована, когда измерители содержат только закрытые задания и их проверка проводится автоматически, а сам процесс характеризуется высокой технологичностью. Такие задания обеспечивают охват большого объема содержания, повышение точности измерений за счет длины измерителя (количества заданий), надежности и валидности результатов. Всё это является основаниями, по которым в аттестационное испытание на основе многостадийного оценивания следует включать задания закрытой формы, отдавая им приоритетную роль. Вместе с тем, неизменным компонентом в процедурах высокой значимости должны быть и открытые задания, к числу которых относятся и кейсы. Их включение в инструментарий является исключительно важным в силу возможности оценивать сложные конструкты в контексте смоделированной профессиональной ситуации, демонстрировать выпускникам способность к критическому мышлению, коммуникации, готовность к решению проблем. В таблице 8 приводится сравнение характеристик закрытых и открытых заданий.

*Таблица 7*

**Пример содержания рубрик оценочного листа**

<b>Критерии оценивания</b>	<b>Показатели оценивания</b>
Формулировка проблемы, обоснование выбора	Сформулирована тема в виде проблемного вопроса
Ориентация в понятийном поле	Определены основные понятия

<b>Критерии оценивания</b>	<b>Показатели оценивания</b>
Основные теоретические идеи и их объяснения	Охарактеризованы основные теоретические идеи Указаны и использованы не менее двух источников
Основные методические идеи и их разъяснение	Даны пояснения методических идей Указаны и использованы не менее одного источника
Практика ( примеры, ситуации, кейсы)	Приведены примеры, ситуации, кейсы
Выводы, включая <i>личное мнение, рефлексию</i>	Рассмотрены различные точки зрения, подходы, сформулирована своя точка зрения Сделаны выводы, проведена рефлексия

Для того, чтобы минимизировать неизбежные потери в точности измерений и выбрано *многостадийное оценивание* при последовательном предъявлении заданий с нарастанием их трудности и уменьшением степени детерминированности. Стратегии и алгоритмы такой технологии описаны в следующем параграфе данной главы.

*Таблица 8*

### **Сравнительные характеристики закрытых и открытых заданий**

	<b>Формы заданий</b>	
	Закрытые задания с выбором одного или нескольких правильных ответов	Открытые задания со свободно конструируемыми ответами
<b>Преимущества</b>	1) оптимально подходят для содержания многих дисциплин;	1) исключают угадывание;

	<b>Формы заданий</b>	
	Закрытые задания с выбором одного или нескольких правильных ответов	Открытые задания со свободно конструируемыми ответами
	2) обеспечивают представление большого объема содержательной области; 3) предполагают автоматизированную проверку; 4) позволяют повысить надежность и сопоставимость результатов измерений	2) позволяют оценивать уровень освоения разнообразных компетенций и трудовых функций; 3) формулируются легче; 4) способствуют росту валидности результатов измерений
<b>Недостатки</b>	1) трудозатратная работа по конструированию дистракторов; 2) возможность угадывания ответа; 3) не подходят для проверки продуктивного уровня деятельности	1) не позволяют включить большой объем содержания; 2) ресурсозатратная процедура проверки; 3) снижают надежность результатов измерений

### ***Портфолио.***

Портфолио как метод оценивания к настоящему времени достаточно изучен и методически проработан [38, 60, 87, 102, 112, 154, 196, 260 и др.]. В рамках проведения итоговой аттестации в предлагаемом авторском подходе многостадийного оценивания портфолио выступает своего рода подтверждающим дополнением к полученным оценкам уровня сформированности выпускника и его готовности к осуществлению профессиональной деятельности. Классическая ошибка при использовании портфолио в оценивании достижений обучающихся или результатов освоения выпускниками образовательной программы заключается в начислении баллов и суммировании их с другими качественными оценками, а затем – выведении

среднего. Эти операции недопустимы, что уже отмечалось выше в данной работе. Вместе с тем, опираясь на практическую деятельность автора диссертации и результаты опытно-экспериментальной работы, следует учитывать, что портфолио способствует получению аутентичных результатов в оценочных процедурах и снижает психолого-эмоциональное напряжение у аттестуемых, поскольку выступает интеллектуальным капиталом, накопленным за период освоения программы.

В целом под *портфолио* принято понимать содержательный комплект работ выпускника, подтверждающий его успехи и достижения за период освоения образовательной программы, способность к самооценке, умений самостоятельно решать проблемы различного содержания при выполнении законченных теоретических или практических исследований и работ, проектов, обладающих значимостью для поставленных целей обучения. Подборка обязательно осуществляется с участием самого обучающегося и включает его лучшие работы, которые оцениваются на основе четко определенных критериев независимыми экспертами (преподавателями, руководителями практик, представителями предприятий и организаций, сокурсниками и т. д.) и сопровождаются свидетельствами самостоятельной работы обучающегося. В эту подборку работ целесообразно включать отражение процесса и результатов прохождения практической подготовки выпускника, что усиливает выводы об уровне его компетентности в связи с взятым курсом на сближение образовательных и профессиональных стандартов в высшем образовании.

Одним из современных видов измерителей результатов обучения является *е-портфолио*. Целью внедрения *е-портфолио* является совершенствование системы подготовки специалиста за счет его вовлечения в профессиональное самоопределение, развитие рефлексивных навыков самооценки индивидуальных достижений, ориентации на выстраивание образовательной траектории, профессиональной карьеры. Такой вид



измерителя может использоваться в дистанционном и гибридном форматах обучения.

Портфолио вообще и е-портфолио в частности относятся к аутентичным методам оценивания, позволяющим студентам анализировать и критически оценивать процесс собственного развития, а также реализующим мотивационную функцию. Особенностью электронного портфолио является его динамичность, мобильность, модульность. Благодаря тому, что оно хранится в сети, его легко переносить, совместно редактировать, перемещать в новую систему или новую рабочую среду. Так, е-портфолио может существовать и развиваться после завершения курса или окончания университета, перехода на следующий уровень образования.

Для оценивания портфолио используются различные критерии. К их числу относятся:

- качество содержания работ, включая свидетельство того, что обучающийся наблюдает за изменением собственного понимания учебного материала, метакогнитивного мышления, уровня сформированности компетенции или продуктивной мыслительной деятельности;

- рост и развитие достижений обучающегося в сравнении с целевыми требованиями к качеству подготовки, например, с требованиями ФГОС или содержанием образовательной программы;

- понимание и правильность использования ключевых процессов когнитивной деятельности;

- полноту, правильность и уместность материалов портфолио;

- разнообразие форм предоставления данных портфолио.

Очевидно, что по трудозатратам оценивание портфолио выпускника может оказаться выше, чем при оценивании выполнения кейсов и проведении структурированного собеседования. Но преимущества его несомненны, поскольку портфолио обеспечивает:

- непрерывность в отслеживании и оценивании качества образовательных результатов;

- многомерность представления разнообразных компетенций;
- оценку мыслительной деятельности выпускника, его междисциплинарные способы деятельности при постановке и решении проблем и сформированности различных компетенций.

При правильном составлении и использовании материалы портфолио содержат достаточно много полезной информации и должны занять весомое место в структуре контрольно-оценочной деятельности. Особенностью портфолио является участие в оценке его результатов широкого круга независимых экспертов, в число которых могут входить научный руководитель, руководитель практики, куратор, другие преподаватели, сокурсники и т. д. Эти отзывы составляются в произвольной форме и служат дополнительным источником информации о результатах работы обучающегося. По своему назначению портфолио выступает аутентичным подтверждением качества подготовки выпускника как на внутреннем, так и внешнем уровнях оценивания. Особую роль оно должно занимать в аккредитационном мониторинге, независимой оценке качества подготовки специалистов и внутренней системе оценки качества образования.

Оценочный лист для портфолио, который может быть использован как самим студентом на этапе подготовки к итоговой аттестации в форме многостадийного оценивания, так и экспертами, представлен в следующей табличной форме (табл. 9).

## Оценочный лист для портфолио

№ п/п	Направления оценивания	Количество различных материалов				Качество и уровень представленного материала (высокий, базовый, минимальный)
		работ	отзывов	документов	доп. материалы	
1.	Участие в НИР: статьи, тезисы, выступление на конференциях, научных семинарах, круглых столах, участие в выполнении научных проектов и т.п.					
2.	Опыт практической деятельности в качестве педагога- психолога (участие в мероприятиях, методические разработки)					

№ п/п	Направления оценивания	Количество различных материалов				Качество и уровень представленного материала (высокий, базовый, минимальный)
		работ	отзывов	документов	доп. материалы	
3.	Участие в общественной деятельности: волонтерство, работа в общественных благотворительных фондах, органах студенческого самоуправления и др.					

### **3.2. Стратегии и алгоритмы реализации многостадийного оценивания**

Многостадийное оценивание с критериально-ориентированной интерпретацией его результатов представляет собой наиболее эффективное решение для классификации выпускников при итоговой аттестации. Как правило, аттестация предполагает простое дихотомическое решение – отнесение к группе аттестованных или не аттестованных, однако в случае итоговой аттестации выпускников, подтверждающих освоение образовательной программы высшего образования, требуется выставление

оценки по пятибалльной шкале. В этой связи дихотомическое правило должно быть перенесено на каждую стадию оценивания, проводимую последовательно от минимального уровня компетентности к высокому уровню и соответствующую выставлению определенной оценки аттестуемому («удовлетворительно», «хорошо», «отлично»). Эффективность такого решения подтверждается получаемыми объективными результатами уровня подготовленности выпускника к профессиональной деятельности за счет возможности всесторонней оценки его компетенций на разных уровнях и с помощью различного оценочного инструментария.

Строго говоря, речь следует вести не столько об эффективности, сколько об *эффективности* выбираемой стратегии. Данное понятие рассматривалось, прежде всего, в экономической и технологической сферах, включая научную, но с определенной долей условности может быть применено и к оценочным процедурам в образовании. Известно, что в англоязычной лексике существует два термина *effectiveness* и *efficiency*, оба из которых в русском языке будут переведены как эффективность. Однако этимологический анализ этих слов и контекст, в которых они используются, свидетельствуют, что *effectiveness* – это эффективность как впечатляющий результат, тогда как *efficiency* – это эффективность как рациональность и полезность. Поэтому реализацию многостадийного оценивания следует рассматривать с позиций эффективности за счет технологичности и прагматичности самого подхода к организации и содержанию аттестационного испытания в форме государственного экзамена, поскольку этапы количественного уровня измерений реализуются автоматически, минимизируя человеческие и временные ресурсы и повышая точность измерений, а качественный уровень обеспечивает экспертные оценки экзаменаторов, без которых немыслимо оценивание способностей выпускника к профессиональной деятельности.

Для реализации многостадийного оценивания можно предложить несколько стратегий. Общим для каждой стратегии является

предусмотренный на определенном или всех этапах компьютерный формат проведения. Преимущества такого формата очевидны – это экономия ресурсов (временных, финансовых, логистических), возможность мгновенного получения обратной связи, обеспечение высокой информационной безопасности, повышение надежности результатов измерений, эффективности и эффективности.

Одна из стратегий основывается на компьютерном адаптивном тестировании (КАТ), строящимся на аппарате современной теории тестов (IRT) и которая излагалась в ранних работах автора диссертации [138, 140, 160 и др.]. Суть ее применительно к оцениванию визуально представлена на рисунке 21.

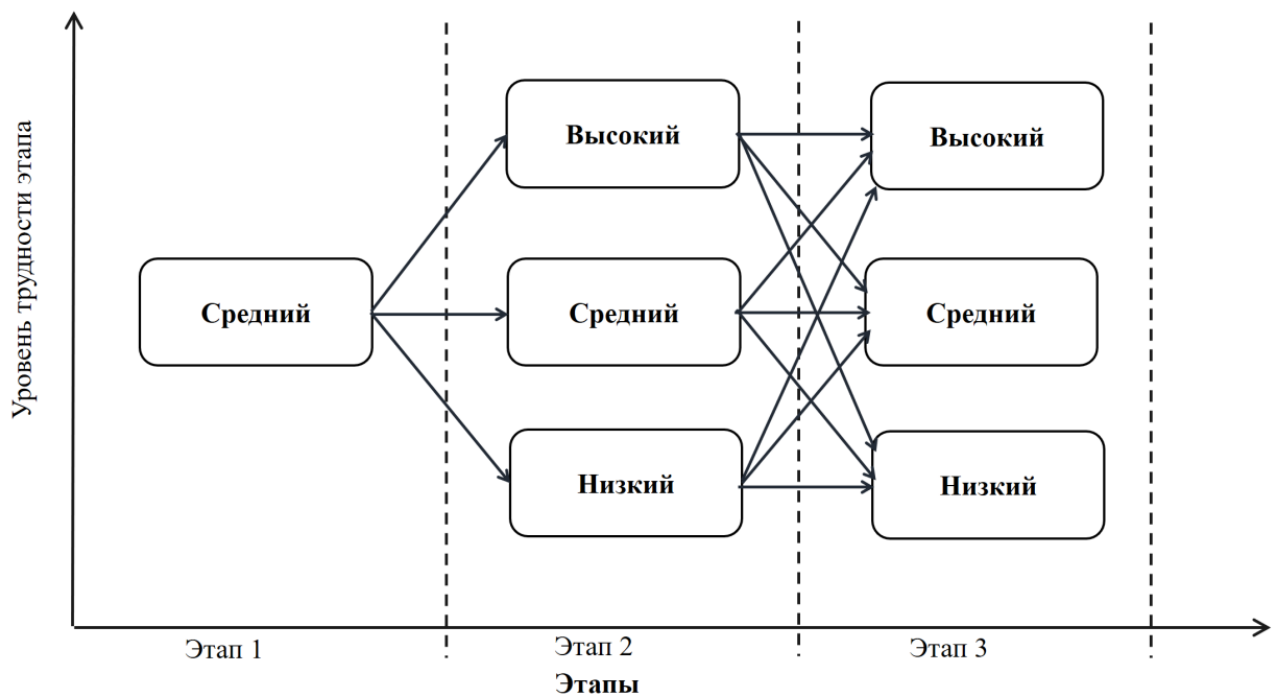


Рис. 21. Многостадийное оценивание в адаптивном режиме

Процедура состоит из нескольких этапов, отличающихся по целям и задачам. Первый этап предусматривает прохождение испытуемыми тестирования с заданиями средней трудности, так называемый входной тест, целью которого является осуществление максимальной дифференциации испытуемых вдоль оси переменной измерения. По результатам такой

дифференциации определяется интервал или множество точек на оси, применительно к которым на втором этапе организуется адаптивный режим и происходит адаптивное тестирование, но в одном из трех диапазонов трудности для каждого испытуемого – высоком, среднем, низком. По результатам выполнения заданий на втором этапе, учитывая полученную оценку уровня подготовленности аттестуемого студента, он переходит на следующий – третий этап тестирования на тот уровень трудности, который ему соответствует. Упрощенное представление самого адаптивного тестирования представлено на рисунке 22.

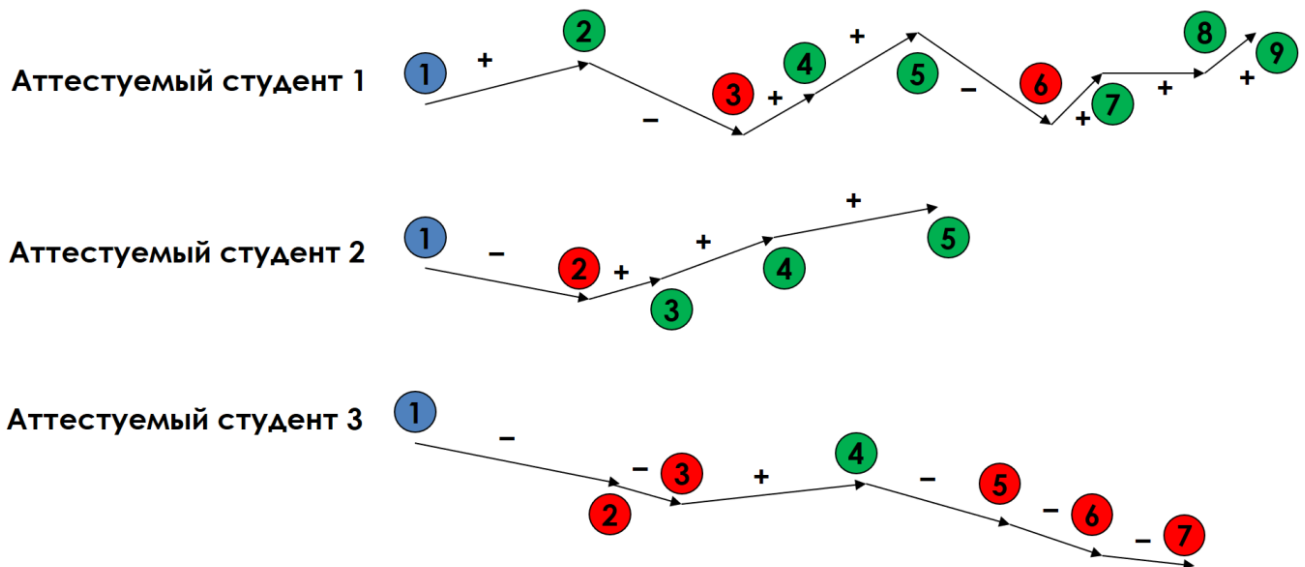


Рис 22. Визуализация адаптивного тестирования

Источник: Звонников В.И., Челышкова М.Б. Современные средства оценивания результатов обучения: учебник. 5-е изд., перераб. М.: Академия, 2013. 304 с.

Режим адаптивного тестирования представляет собой такую его организацию, когда каждый испытуемый движется по своей индивидуальной траектории в процессе выполнения отличающихся по числу и трудности наборов заданий. Алгоритм отбора и предъявления заданий строится по принципу обратной связи, когда при правильном ответе испытуемого очередное задание выбирается более трудным, а неверный ответ влечет за

собой предъявление последующего более легкого задания, чем то, на которое испытуемым был дан неверный ответ. Таким образом, такое тестирование строится на алгоритмах с полной контекстной зависимостью, в которых очередной шаг совершается только после оценки результатов выполнения предыдущего шага, поскольку зависит от этого результата.

В свою очередь многошаговые стратегии адаптивного тестирования подразделяются на фиксированно-ветвящиеся и варьирующе-ветвящиеся [139, 262]. Если один и тот же набор с фиксированным расположением заданий на оси трудности используется для всех испытуемых, но каждый движется по набору индивидуальным путем в зависимости от результатов выполнения каждого очередного задания, то стратегия адаптивного тестирования является фиксированно-ветвящейся. Задания по трудности в наборе обычно располагают на равном расстоянии друг от друга либо выбирают убывающий шаг сообразно нарастанию трудности, что позволяет подстроить темп тестирования под испытуемого. Приведем описание двух наиболее распространенных стратегий, относящихся к фиксированно-ветвящимся.

Суть *пирамидальной стратегии* заключается в том, что все испытуемые начинают процесс тестирования с задания средней трудности. Если ответ испытуемого на задание правильный, то ему предъявляется следующее по степени трудности задание; при неправильном ответе предлагается менее трудное задание. Процедура повторяется до тех пор, пока испытуемый не выполнит запланированное количество заданий. Для реализации пирамидальной стратегии необходимо при заданном количестве стадий тестирования (оно совпадает с количеством уровней трудности) определить количество заданий каждого уровня и в тесте в целом.

На рисунке 23 представлен пример 10-стадийного теста, в котором имеется 55 заданий. В начале тестирования студенту предъявляется задание средней трудности (5-го уровня). На втором шаге ему может быть предложено задание либо 5-го, либо 6-го уровня. Очевидно, что на каждом шаге могут быть



задействованы задания столько уровней трудности, каков номер шага. Если в тесте имеются задания 10 уровней трудности, то в целом каждому испытуемому предъявляется 10 из 55 входящих в тест заданий.

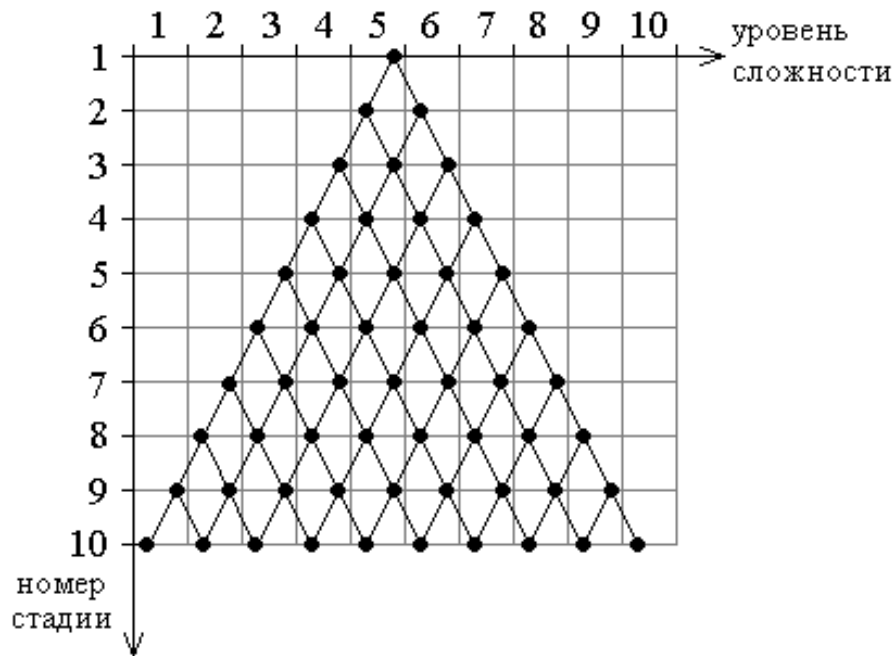


Рис. 23. Схема распределения заданий для 10-стадийного теста

Заметим, что в пирамидальной стратегии на каждом уровне трудности требуется определенное количество заданий. Наибольшее количество заданий (равное количеству уровней трудности) может быть задействовано на среднем уровне. Всего одно задание используется на максимальном уровне. На соседних уровнях количество заданий отличается на 2 (исключение составляют уровни, соседствующие со средним). Можно заметить, что на первом уровне требуется 2 задания, на втором – 4, на третьем – 6 и т. д. На последнем уровне используется 1 задание, на предпоследнем – 3 и т. д.

Можно использовать общую формулу для определения количества заданий на произвольном уровне трудности. Пусть всего имеется  $K$  уровней трудности. Номер среднего уровня трудности определяется как целая часть от деления количества уровней на 2: очевидно, при  $K$  четном это будет  $K/2$ , а для нечетного  $K$  получим  $(K+1)/2$ . На уровне трудности с номером, меньшим  $K/2$ ,

используется  $2i$  заданий, где  $i$  – номер уровня. Для уровней трудности с номерами, большими  $K/2$ , количество заданий равно  $2(K-i) + 1$ , где  $i$  – номер

уровня. Всего в тесте будет использовано  $\frac{(1+K)K}{2}$  заданий.

В таблице 10 представлены количества заданий каждого уровня и в тесте в целом для различных значений количества стадий тестирования. Очевидно, что пирамидальная стратегия может быть использована только при условии наличия достаточно большого банка заданий различного уровня трудности, однако она соответствует упрощенному пониманию адаптивного тестирования.

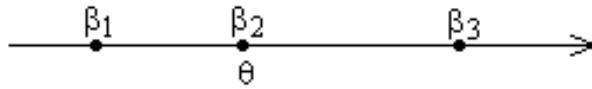
Таблица 10

### Количество заданий каждого уровня для пирамидальной стратегии

Количество стадий тестирования	Всего заданий в тесте	Заданий уровня																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	15	2	4	5	3	1															
6	21	2	4	6	5	3	1														
7	28	2	4	6	7	5	3	1													
8	36	2	4	6	8	7	5	3	1												
9	45	2	4	6	8	9	7	5	3	1											
10	55	2	4	6	8	10	9	7	5	3	1										
11	66	2	4	6	8	10	11	9	7	5	3	1									
12	78	2	4	6	8	10	12	11	9	7	5	3	1								
13	91	2	4	6	8	10	12	13	11	9	7	5	3	1							
14	105	2	4	6	8	10	12	14	13	11	9	7	5	3	1						
15	120	2	4	6	8	10	12	14	15	13	11	9	7	5	3	1					
16	136	2	4	6	8	10	12	14	16	15	13	11	9	7	5	3	1				
17	153	2	4	6	8	10	12	14	16	17	15	13	11	9	7	5	3	1			
18	171	2	4	6	8	10	12	14	16	18	17	15	13	11	9	7	5	3	1		
19	190	2	4	6	8	10	12	14	16	18	19	17	15	13	11	9	7	5	3	1	
20	210	2	4	6	8	10	12	14	16	18	20	19	17	15	13	11	9	7	5	3	1

*Стратегия половинного деления* строится на основе IRT, согласно которой значения уровня подготовленности аттестуемых и трудности заданий теста выражены в одних и тех же единицах измерения – логитах, а значит, могут быть помещены на одну стандартную шкалу, что позволяет соотнести

уровень подготовки любого студента с мерой трудности каждого задания теста. Абсолютная величина разности  $(\theta - \beta)$  – это расстояние, на котором находится испытуемый с уровнем подготовки  $\theta$  от задания с трудностью  $\beta$ .



Если отрицательная разность велика по модулю  $(\theta - \beta_3)$ , то это значит, что уровень подготовки студента намного ниже трудности задания, и с очень большой вероятностью он сделает его неверно. Большие положительные значения разности  $(\theta - \beta_1)$  также говорят о несоответствии уровня подготовленности и трудности задания, но в другую сторону, то есть с очень большой вероятностью испытуемый выполнит такое задание правильно.

Если задания расположены в порядке возрастания трудности, то можно утверждать следующее:

- если испытуемый справился с заданием, то с большой долей вероятности он правильно выполнит задания, уровень трудности которых меньше;

- если испытуемый не справился с заданием, то с большой долей вероятности он не сможет выполнить задания с более высоким уровнем трудности.

Эти утверждения позволяют выбирать задания таким образом, что студенту не предлагаются те задания, результат выполнения которых предсказуем с большой долей вероятности.

Пусть все задания расположены в порядке возрастания трудности, причём все задания по трудности различаются. Пусть испытуемому предложено задание среднего уровня трудности. Возможны две ситуации: студент с заданием не справился, студент выполнил задание верно. В первом случае с большой долей вероятности можно предположить, что испытуемый не справится с более сложными заданиями, во втором – справится с более легкими заданиями. Таким образом, отбирается та половина заданий, результат выполнения которых можно считать известным, и поэтому их

можно не предъявлять. Следовательно, можно выбрать оставшуюся половину заданий, относительно результата которых нельзя сделать никаких предположений: в первом случае это будут задания более легкие, а во втором – более трудные, чем предъявленное. То есть, для предъявления испытуемому имеем в два раза меньше заданий, чем было ранее. Повторяя на следующем шаге тестирования аналогичную процедуру, множество заданий будет делиться на две части (этим и объясняется название стратегии), то есть происходит сужение «области неопределенности» в два раза. Когда на очередном шаге мы получаем всего одно задание, тестирование завершается. Пример стратегии половинного деления для 16 заданий представлен на рисунке 24.

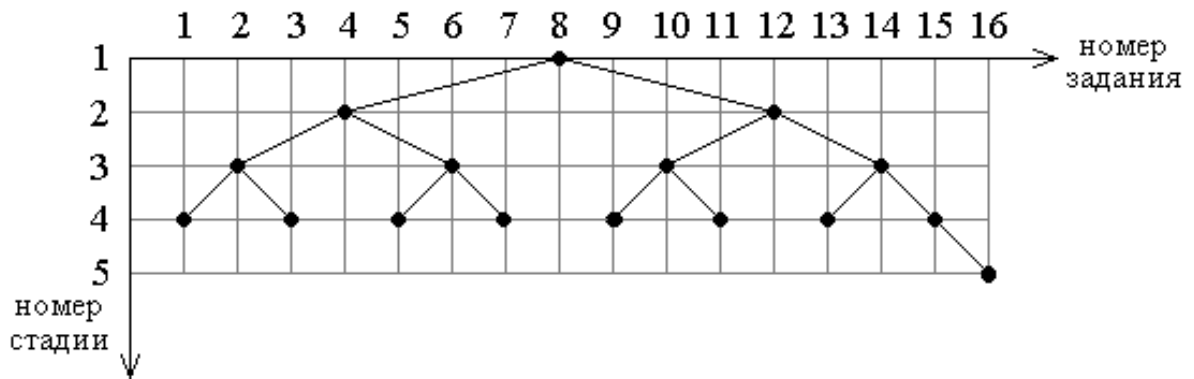


Рис. 24. Схема стратегии половинного деления для 16 заданий

Таким образом, если имеется  $K$  заданий, то после первого шага тестирования остается  $K/2$  заданий, после второй –  $K/4$  заданий, после третьей –  $K/8$  заданий и т. д. Тестирование продолжается, пока выполняется условие в

виде неравенства  $\frac{K}{2^i} > 1$ , где  $i$  – номер стадии. Это неравенство позволяет также по количеству заданий определить максимальное количество стадий, которое понадобится, – оно равно  $[\log_2(K+1)]$ , где  $[\cdot]$  обозначает целую часть числа.

**Количество стадий в стратегии половинного деления**

Количество заданий	8–15	16–31	32–63	64–127	128–255	256–511	512–1000
Количество стадий	4	5	6	7	8	9	10

Итак, стратегия половинного деления может быть описана следующим алгоритмом. Пусть всего имеется  $K$  заданий.

1. Предъявляется задание с номером  $j = [K/2]$ , номер стадии  $i = 1$ .

2. Если  $\frac{K}{2^i} > 1$ , то переходим к п.3, иначе тестирование завершается.

3. Увеличиваем номер стадии  $i$  на 1. Если задание с номером  $j$  выполнено

верно, то предъявляется задание с номером  $\left[ j + \frac{K}{2^i} \right]$ , иначе – с номером  $\left[ j - \frac{K}{2^i} \right]$ . Переход к п.2.

При имеющемся количестве заданий описанная стратегия позволяет получить результат за минимальное количество шагов, что влечет уменьшение длины теста, но происходит потеря точности измерения, а тем самым надежности адаптивного тестирования.

**Варьирующая ветвящаяся стратегия** заключается в предъявлении заданий по алгоритму, который прогнозирует оптимальную трудность последующего задания по результатам выполнения студентом предыдущего тестового задания. Главной особенностью и одновременно преимуществом варьирующей ветвящейся стратегии является пошаговая переоценка уровня подготовленности испытуемого после каждого выполнения очередного задания. В результате образуется своеобразная последовательность значений  $\theta$ , применительно к которым подбираются значения трудности задания  $\beta$ . Варьируются при такой стратегии адаптивного тестирования трудность заданий и шаг, определяемый разностью трудностей двух соседних заданий последовательности. Реализовать такую стратегию возможно только

с помощью вероятностных моделей теории IRT, как раз позволяющих прогнозировать успешность выполнения очередного задания при фиксированном значении  $\theta$ . Алгоритм варьирующей ветвящейся стратегии представлен на рисунке 25. Этот алгоритм был ранее реализован автором диссертации для проведения текущего контроля и промежуточной аттестации студентов в условиях дистанционного обучения [139]. Достоинствами этой стратегии являются:

- выбор в качестве критерия остановки алгоритма адаптивного тестирования величины ошибки измерения, получаемой при оценивании переменной измерения, что позволяет заранее прогнозировать надежность теста и результатов;

- индивидуальная траектория адаптивного тестирования для каждого студента;

- возможность вариации трудности и дифференцирующей способности при подборе оптимальных заданий для каждого контролируемого студента.

Опросы студентов, их реакция на формат проведения оценивания на основе адаптивного тестирования говорят о положительном отношении к таким формам и методам, особенно в условиях дистанционного обучения, когда принципы справедливости оценки и ее объективности, а также создание равных условий для каждого звучат с особой остротой.

Вместе с перечисленными достоинствами остается нерешенным вопрос о том, как входить в режим адаптивного тестирования, т. е. с задания какой трудности следует начинать оценочную процедуру. Можно предложить несколько подходов для начала тестирования. Первый подход заключается в том, что если отсутствует информация о начальном уровне подготовленности контролируемого студента, то следует ориентироваться на среднюю оценку в группе, проходившей тестирование. Вторым подходом относится к двухшаговому тестированию и строится на использовании результатов выполнения входного теста (претеста). При третьем подходе тестирование

предлагается начинать с относительно легких заданий, дав время испытуемому для привыкания и адаптации [262].



Рис. 25. Алгоритм варьирующе-ветвящегося адаптивного тестирования

Вместе с тем реализация адаптивного тестирования в исключительно компьютерном формате без общения с экспертами в условиях оценочной процедуры «с высокими ставками» в современных условиях представляется не вполне оправданной. Это связано как с формальной стороной вопроса, регламентацией итоговой аттестации, так и с психолого-педагогическими аспектами. В частности, это связано со временем, отведенным на выполнение каждого задания в адаптивном режиме, невозможностью для испытуемого видеть весь набор заданий, их количество, отсутствие возможности вернуться к предыдущему заданию и пересмотреть выбранный вариант решения и др. В условиях психологически напряженной ситуации экзамена высокой значимости указанные обстоятельства будут только усиливать тревожность аттестуемых.

Производным от точечного адаптивного тестирования с пошаговой переоценкой трудности предъявляемого задания и в соответствии с пересчитанным уровнем подготовленности является структурированное многостадийное оценивание. Действительно, переход от более легкой стадии оценивания, содержащей задания, посильные для выполнения большинством испытуемых, к стадиям с нарастанием трудности заданий, релевантным уровню компетентности, в полной мере отвечает идеям адаптивного тестирования, а также гуманистическому характеру контрольно-оценочной деятельности. Стратегия такого структурированного последовательного многостадийного оценивания представлена на рисунке 26.

В роли измерителей на каждом этапе выступает определенный инструментарий, разработка которого представлена в предыдущем параграфе. Для описания методики проведения многостадийного оценивания приведем его декомпозицию в соответствии с определенными уровнями образовательных измерений и инструментарием (рис. 27).



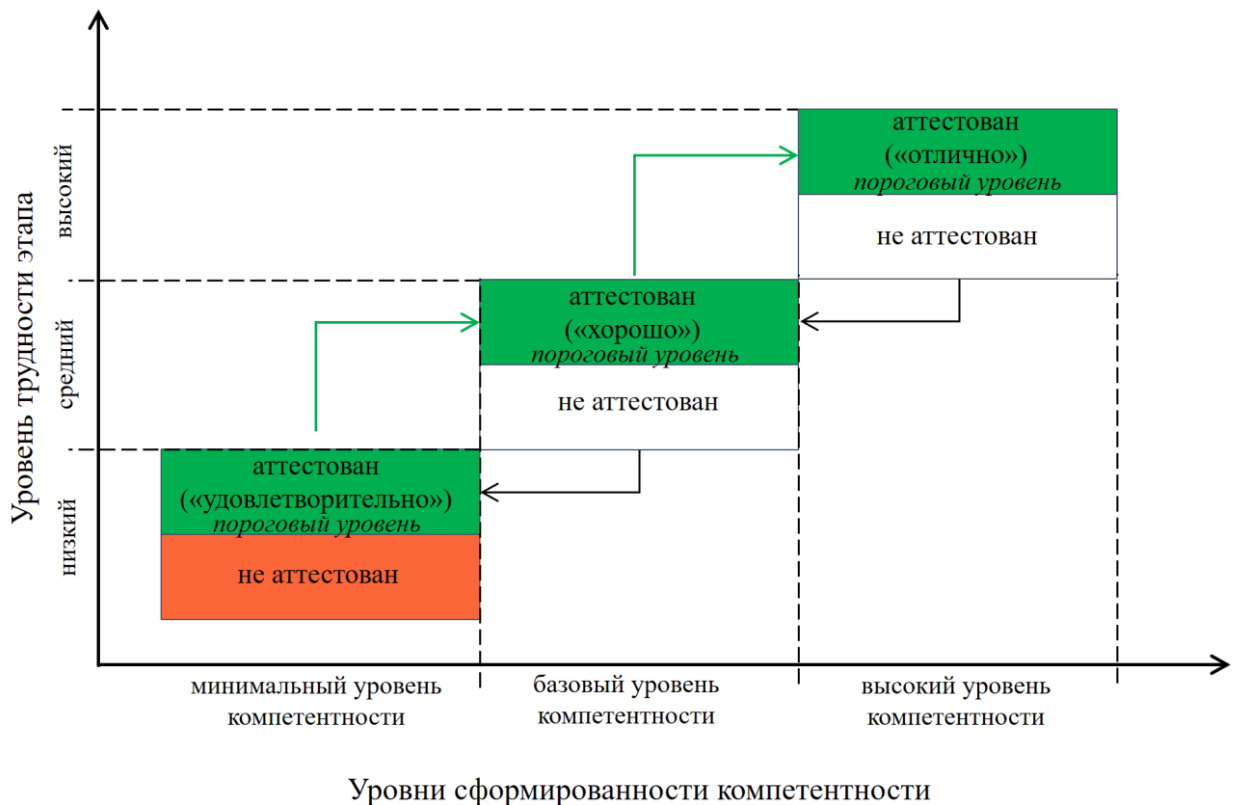


Рис. 26. Стратегия многостадийного оценивания при итоговой аттестации

Первый этап оценивания минимального уровня компетентности проводится методом компьютерного тестирования с помощью субтестов, содержание каждого из которых измеряет необходимые знания и умения по базовой дисциплине либо конкретной компетенции, имеющих определяющее значение для будущей профессиональной деятельности выпускника и увязанные с трудовыми функциями профессионального стандарта. Как обосновывалось при описании спецификации в параграфе 3.1, таких субтестов на данном этапе будет несколько, уровень трудности заданий и самих субтестов должны быть такими, чтобы реализовывался критериально-ориентированный подход в образовательных измерениях, а длина субтестов – достаточной для обеспечения надежных результатов. При этом у испытуемого есть возможность переключаться между заданиями, возвращаясь обратно. Другими словами, каждый аттестуемый выпускник видит весь набор заданий по каждому субтесту, которые ему предстоит выполнить, и может сам

регулировать свою траекторию тестирования. Время выполнения на каждый субтест не регламентировано, оно определяется в целом временем на прохождение первого этапа оценивания, исходя из спецификации, итогов экспертной оценки и предварительной апробации. При преодолении аттестуемым критериального (порогового) балла на первой стадии оценивания его уровень компетентности признается компетентным с оценкой «удовлетворительно».

Переход на второй этап оценивания, который также реализуется в компьютерном формате, осуществляется по желанию аттестуемого, если он готов подтвердить базовый уровень компетентности, соответствующий оценке «хорошо».

Второй этап оценивания предусматривает выполнение аттестуемым профессионально-ориентированных (ситуационных) заданий, уровень трудности которых выше минимального, а количество может варьироваться от 5 до 12. Выполнение ситуационных или контекстуализированных заданий проблемного характера с множественным выбором, позволяющих провести измерения по совокупности компетенций и решить проблемы привлечения для проверки эксперта. За счет множественности переменных измерения и проблемного характера содержания такие задания позволят оценить способность студентов к критическому и абстрактному мышлению, анализу и синтезу информации. Испытуемый может выполнять эти задания в произвольном порядке, пропуская и возвращаясь снова. Время для прохождения второй стадии оценивания определяется, исходя из времени, предусмотренного для выполнения ситуационных заданий в соответствии со спецификацией. Выполнение необходимого количества заданий для аттестации на данном этапе, т. е. выполнение стандарта оценивания и достижение критериального (порогового) балла, является основанием для оценки «хорошо» и перехода на этап подтверждения высокого уровня компетентности и оценки «отлично» на государственном экзамене. В случае, если стандарт оценивания не выполнен, аттестуемый будет отнесен к

категории аттестованного на минимальном уровне компетентности с оценкой «удовлетворительно».

Третий этап на основе структурированного собеседования по полученному кейсу относится к продуктивному уровню оценивания, поскольку предполагает не только воспроизведение необходимых для будущей профессиональной деятельности знаний, умений, демонстрацию способов деятельности через решение контекстуализированных заданий проблемного характера, но и предусматривает устную коммуникацию с экспертами, роль которых выполняют члены экзаменационной комиссии.

Как правило, на данном этапе выпускнику дается время для ответа с обоснованием решения по полученному кейсу, после чего аттестуемый обосновывает предлагаемое решение кейса, отвечает на уточняющие и дополнительные вопросы членов экзаменационной комиссии, которые проводят оценивание по оценочным листам (чек-листам). Итоговая оценка носит коллегиальный характер, учитывающий мнение каждого члена комиссии. При ответе аттестуемый также имеет возможность представить портфолио, оценивание которого может быть проведено заблаговременно до начала аттестационного испытания. Важно, что портфолио вносит вклад в итоговую оценку, в идеале подтверждая выводы о высоком уровне компетентности выпускника. В том случае, если аттестуемый не справился с выполнением кейса, он классифицируется как успешно прошедший второй этап оценивания с оценкой «хорошо» и базовым уровнем компетентности.

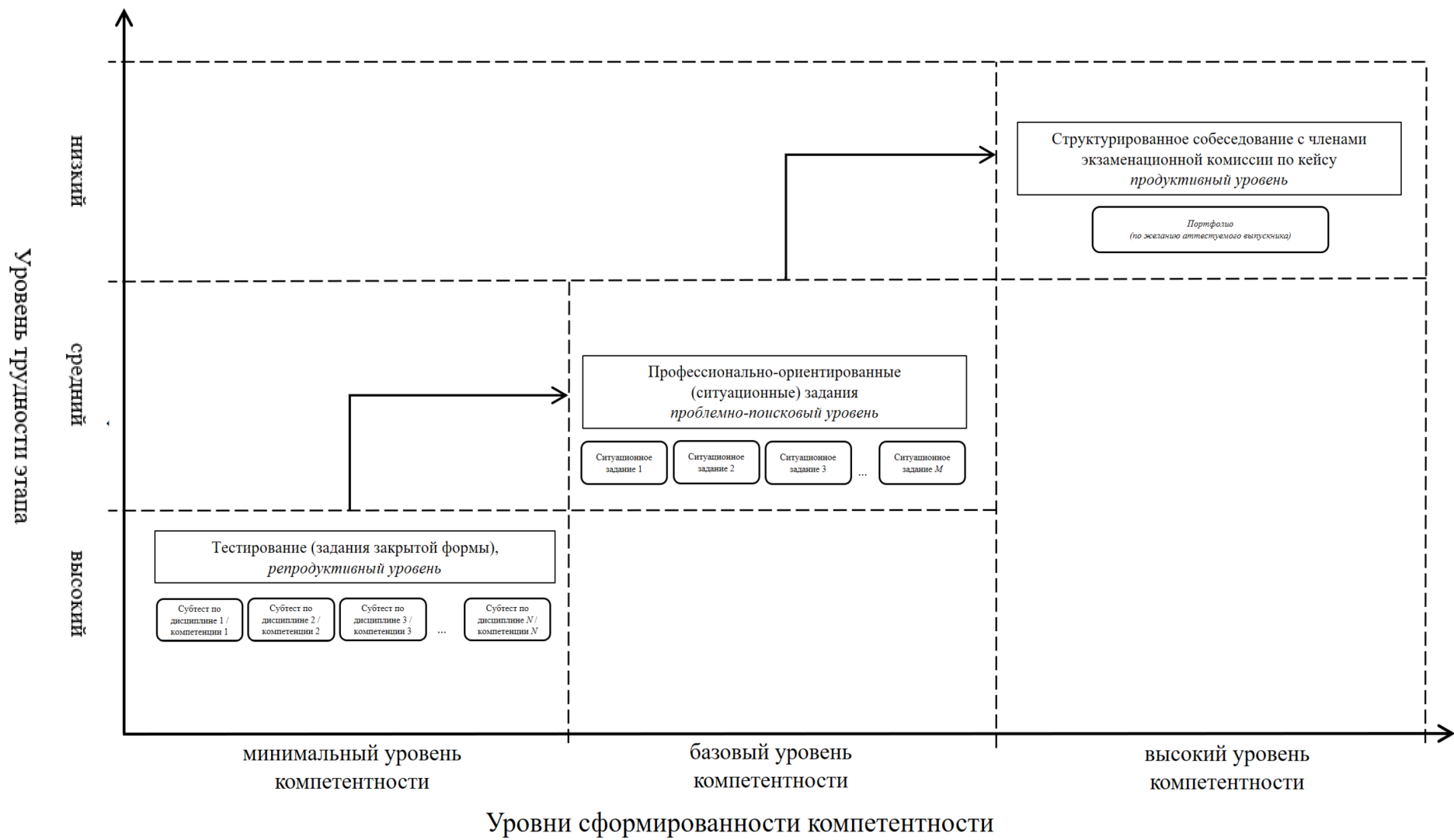


Рис. 27. Декомпозиция многостадийного оценивания

### **3.3. Методика определения надежности результатов многостадийного оценивания**

Концептуальные представления о надежности были даны во второй главе при обосновании методологии многостадийного оценивания. Следует подчеркнуть, что основывающееся на образовательных измерениях оценивание отличается от традиционных методов, как правило, строящихся на субъективных суждениях экзаменаторов, две ключевые характеристики – надежность и валидность. В различной литературе можно встретить словосочетания «надежный тест» или «надежный инструментарий», но корректнее говорить о надежности или ненадежности результатов измерений, которые получаются при использовании конкретного инструментария. И, как подчеркивалось выше, разработку инструментария следует вести сначала с опорой на классическую теорию тестов, а затем переходить к современной. Этой логики следовал и автор диссертации, когда принимал участие в разработке и апробации оценочного инструментария для аттестационных процедур высокой значимости (аккредитация специалистов в области здравоохранения, итоговая аттестация выпускников по программам бакалавриата и магистратуры). Только после того, как оценочный инструментарий пройдет апробацию на репрезентативных выборках студентов, можно говорить о начале формирования фондов оценочных средств или, говоря иными словами, – банка калиброванных заданий, обеспечивающих получение надежных результатов при их использовании. Чем надежнее результаты измерений, тем меньше ошибка измерения и выше объективность. Именно стремление к объективности оценок при итоговой аттестации заставляет обращаться к рассмотрению различных методов надежности и выборе тех, которые релевантны условиям многостадийного оценивания.

В силу сочетания количественного и качественного уровней образовательных измерений, отвечающих компетентностной трактовке при оценке качества подготовки специалистов, необходимо понимать природу явления и

причины снижения надежности результатов образовательных измерений при многостадийном оценивании. Это понимание крайне важно в условиях итоговой аттестации, чтобы стремиться к способам минимизации этих причин. Ниже в табличной форме представлены причины снижения надежности результатов измерений, которые сформированы не только на основе изучения соответствующей литературы, но и практического опыта автора диссертации, и способы минимизации этих причин (табл. 11).

Традиционные способы определения надежности, которые используют при нормативно-ориентированном подходе, анализируя дисперсию результатов измерений и стремясь к ее максимизации, при итоговой аттестации в условиях компетентностного подхода в высшем образовании не годятся. Как отмечается в исследованиях R. Berk, R. Brennan, W. Popham, критериально-ориентированный подход требует иных, отличных от нормативно-ориентированных методов определения надежности, валидности, анализа самих заданий [283, 302, 343]. Сделанный в данном исследовании выбор в пользу нескольких этапов оценивания с критериально-ориентированной интерпретацией результатов в рамках аттестационного испытания (государственного экзамена) во многом позволяет преодолевать указанные в таблице 11 причины снижения надежности измерений в целом.

*Таблица 11*

**Причины снижения надежности результатов образовательных измерений и способы их минимизации**

<b>№ п/п</b>	<b>Причина</b>	<b>Способ ее минимизации</b>
1.	Субъективность оценки результатов выполнения заданий измерителя	Связан с большей долей заданий со свободно конструируемым ответом, целесообразно включение достаточного числа заданий закрытой формы

№ п/п	Причина	Способ ее минимизации
2.	Угадывание ответов на задания в закрытой форме	Тщательная работа по подбору и анализу дистракторов
3.	Несоблюдение технологии измерения (от этапа конструирования инструментария до его применения и некорректной интерпретации)	Строго следовать определенным правилам и алгоритмам, повышая квалификацию разработчиков заданий и организаторов оценочных процедур. Использование инструкций для испытуемых
4.	Недостаточная длина измерителей	Соблюдать баланс количества оценочных заданий, применяя различные виды в несколько этапов оценивания
5.	Высокая гетерогенность содержания измерителей	Разрабатывать субтесты (тестлеты), нацеленные на измерение результатов освоения отдельных дисциплин (знаний, умений, способов действий), компетенций или трудовых действий
6.	Отсутствие согласованных критериев оценивания для открытых заданий (таких как кейсы)	Проводить предварительную апробацию критериев методом экспертных оценок (например, метод Дельфи, экспертное интервью и пр.)

Надежность результатов многостадийного оценивания будет складываться из надежности результатов по каждому этапу и используемым измерителям.

Следуя предложенной логике, на каждом этапе необходим стандарт его выполнения, т. е. критериальный (пороговый) балл, делящий выпускников на две взаимоисключающих группы (категории) – аттестованных и не аттестованных. В работе В. Ю. Переверзева надежность для критериально-ориентированного подхода определяется как «согласованность решений, принятых на основании проведения двух испытаний» [203, с. 82]. Исследователь предлагает рассматривать надежность как постоянство результатов аттестуемых, для измерения уровня подготовленности которых применяются задания из одного фонда оценочных средств, соответствующих общему содержанию. Другими словами, согласованность решений касается степени принятия одинаковых решений по двум различным результатам измерения. Измерение может проводиться либо одним и тем же набором заданий одного измерителя, либо по двум взаимозаменяемым измерителям. Под взаимозаменяемостью понимается форма эквивалентности. В результате такого измерения получится два индивидуальных балла для каждого испытуемого, а на основании совокупности всех результатов по группе испытуемых можно составить таблицу сопряженности (табл. 12).

Таблица 12

**Таблица сопряженности результатов измерений при критериально-ориентированном подходе**

		Измерение 1	
		Не аттестован	Аттестован
Измерение 2	Не аттестован	--	+ -
	Аттестован	- +	+ +

Таблица 12 отражает результаты двух измерений по одному и тому же набору заданий (один тест, или взаимозаменяемый тест, или данные из ретестового метода). Знаки «--» означает, что в эту категорию попадут те испытуемые, кто не преодолел критериальный (пороговый) балл в обоих измерениях. В данном поле на пересечении двух категорий «не аттестован» и «не аттестован» вписывается доля таких испытуемых, рассчитываемая как отношение количества не аттестованных



испытуемых по обоим измерениям к общему числу испытуемых, проходящих измерение. Аналогично, знаки «+ +» на пересечении «аттестован» и «аттестован» означают, что здесь указывается доля испытуемых, которые аттестованы по результатам обоих измерений. В той же логике заполняются и оставшиеся два поля «- +» и «+ -».

В работах, посвященных критериально-ориентированным измерениям, указывается, что две группы (в наших обозначениях «- -» и «+ +») рассматривается как область устойчивых решений, а группы «- +» и «+ -» как область противоречивых неустойчивых решений в части отнесения испытуемых к аттестованным и не аттестованным [203, 283]. В первую очередь, доли таких групп будут зависеть от качества подготовки выпускников, ее особенностей, а не от качества инструментария. В этой связи для оценки надежности при критериально-ориентированном измерении необходимо стремиться к репрезентативным выборкам. В условиях итоговой аттестации в высшем образовании данное условие может выполняться только в случае объединения большого количества университетов в апробацию оценочного инструментария для процедур высокой значимости. Автор диссертации с коллегами пошли по пути включения таких оценочных заданий в диагностические работы, которые проводятся в рамках процедуры ВСОКО, что дает основания для расчетов надежности. Опыт других коллег, который отражен в ряде публикаций, также указывает на то, что для получения надежных результатов в условиях компетентностного подхода необходима предварительная апробация оценочного инструментария [8].

Для определения коэффициента надежности, обозначаемого символом  $\varphi$ , рассматривают таблицу сопряженности (табл. 12) и на основании исходных данных вычисляют корреляцию, используя следующую формулу:

$$\varphi = \frac{P_{(- -)} \cdot P_{(+ +)} - P_{(+ -)} \cdot P_{(- +)}}{\sqrt{(P_{(- -)} + P_{(+ -)}) \cdot (P_{(+ -)} + P_{(+ +)}) \cdot (P_{(- -)} + P_{(+ -)}) \cdot (P_{(+ -)} + P_{(+ +)})}}, \quad (7)$$

где  $P_{(- -)}$  – доля испытуемых, не преодолевших критериальный (пороговый) балл в обоих измерениях, т. е. не аттестованных;  $P_{(+ +)}$  – доля испытуемых, преодолевших

критериальный (пороговый) балл в обоих измерениях, т. е. аттестованных;  $P_{(-+)}$  – доля испытуемых, которые не преодолели критериальный балл при первом измерении и преодолели его при втором измерении;  $P_{(+-)}$  – доля испытуемых, которые были аттестованы при первом измерении и не аттестованы при втором измерении. Последние два случая, т. е.  $P_{(-+)}$  и  $P_{(+-)}$ , относят к ошибкам классификации, поскольку испытуемые при двух измерениях не подтвердили свой уровень, изменив его на противоположный.

Формула (7) достаточно легко объясняется. Знаменатель представляет собой нормирующий множитель, а числитель разность произведений долей, обозначающих меру согласованности и несогласованности классификационных решений. Из этого следует вывод, чем больше по значению произведение  $P_{(-+)} \cdot P_{(+-)}$  и меньше  $P_{(++)} \cdot P_{(--)}$ , тем надежнее результаты измерений при критериально-ориентированном подходе.

Также для критериально-ориентированного подхода R. Hambleton и M. Novick предложили рассматривать сумму  $P_{(-+)} + P_{(+-)}$  как меру согласованности принятия решений, а остальные варианты как противоречивые [294]. Из чего следует, что мера несогласованной классификации представляет собой разность  $1 - (P_{(-+)} + P_{(+-)})$ . При этом показывая только первичные оценки надежности, исследователи предложили использовать каппу Коэна, которая может принимать значения в интервале от  $-1$  до  $+1$  и также служить оценкой надежности результатов измерений при критериально-ориентированном подходе. Коэффициент Коэна или  $\kappa$ -коэффициент согласованности рассчитывается по следующей формуле:

$$\kappa = \frac{(P_{(--)} + P_{(++)}) - (P_{(++)} + P_{(-+)})(P_{(++)} + P_{(+-)}) - (P_{(--)} + P_{(+-)})(P_{(--)} + P_{(-+)})}{1 - (P_{(++)} + P_{(-+)})(P_{(++)} + P_{(+-)}) - (P_{(--)} + P_{(+-)})(P_{(--)} + P_{(-+)})}, \quad (8)$$

где  $(P_{(++)} + P_{(-+)}) \cdot (P_{(++)} + P_{(+-)}) + (P_{(--)} + P_{(+-)}) \cdot (P_{(--)} + P_{(-+)}) = P_c$  обозначается как вероятность случайного принятия согласованного решения.

Рекомендуется вести подсчет обоих показателей надежности –  $\phi$ -коэффициент корреляции по формуле (7) и  $\kappa$ -коэффициент согласованности по формуле (8). Близкие значения этих коэффициентов к величине 0,8 будут

свидетельствовать о надежности результатов измерений. Расчеты представленных здесь коэффициентов для целей нашего исследования при реализации первых двух этапов многостадийного оценивания представлены в Приложении 8.

### **3.4. Обеспечение содержательной валидности результатов многостадийного оценивания**

В широком смысле валидность (от англ. *validity* – пригодность, весомость, обоснованность,) следует рассматривать как методологическую характеристику образовательных измерений, поскольку она не только характеризует обоснованность и пригодность применения определенных методик и результатов в конкретных условиях, но также связана с вопросами размерности измерений. В своих исследованиях по измерениям В. И. Звонников отмечает, что «валидность – это характеристика того, в какой мере и насколько удачно удастся измерить именно тот конструкт, который планировался к измерению» [83, с. 208]. В более узкой постановке рассматриваемого вопроса говорят не о валидности результатов оценивания, а о валидности измерителя (оценочного средства). Но такое употребление термина «валидность» не корректно, поскольку оценивание валидности предназначено для осознания того, какие выводы можно сделать по полученным результатам измерения. В целом выводы о качестве результатов многостадийного оценивания осложняются тем, что нет и не может быть раз и навсегда определенной надежности и валидности. Ситуация вполне объяснимая, поскольку оценки надежности и валидности зависят от ряда факторов, меняющихся при различных ситуациях образовательных измерений. По этой причине при каждом измерении приходится оценивать надежность и валидность, а уж потом говорить о достоверной интерпретации результатов измерения [88].

В идеализированной ситуации разрабатываемый инструментарий (измеритель) должен быть валиден по содержанию (*содержательная валидность*), по измеряемым переменным (*конструктная валидность*) и по способности

прогнозировать успехи в профессиональной деятельности выпускников как молодых специалистов на рынке труда (*прогностическая валидность*).

Для оценочных процедур высокой значимости в условиях итоговой аттестации выпускников на первый план выходит содержательная валидность. Не преуменьшая значимость конструктивной и прогностической валидности, задача репрезентативного представления содержания по отношению к запланированным результатам освоения образовательной программы в форме компетенций в оценочном инструментарии и его экспертиза является первоочередной для многостадийного оценивания. Именно поэтому разработке спецификации для последующего конструирования оценочного инструментария уделяется большое внимание. Как было отмечено выше, качественное описание заявленных образовательных результатов в паспортах компетенций является залогом качественно разработанной спецификации. Если измеритель позволяет проверить все то, что задумано авторами в его спецификации, он считается валидным. Представление о содержательной валидности не следует связывать только с полнотой отображения в измерителе оцениваемых переменных, необходимы также правильные пропорции содержательных элементов. Содержательная валидность будет тем выше, чем глубже и полнее отображено всё то, что запланировано к оцениванию.

Обоснование содержательной валидности и ее повышение связано, в первую очередь, с обязательным проведением экспертизы. Эта обязательность обусловлена характером аттестационной процедуры, в которой будет использоваться оценочный инструментарий и по результатам которой будут приниматься решения о присвоении квалификации выпускнику. В этой связи следует особенно подчеркнуть важность экспертного оценивания качества содержания заданий, входящих в оценочный инструментарий.

Стандарты образовательного тестирования определяют содержательную валидность как степень, в которой эмпирические доказательства и теоретические выводы свидетельствуют об адекватности содержания измерителя предполагаемой интерпретации результатов измерений [305]. Для определения содержательной

валидности необходимо привлекать экспертов, являющихся признанными специалистами в той сфере профессиональной деятельности, к которой готовился обучающийся при освоении образовательной программы и у него формировались компетенции, а также эти эксперты не участвовали в разработке самих заданий. В противном случае говорить об экспертизе просто не приходится.

Количество экспертов, как правило, составляет не менее трех человек. Иногда в условиях кадрового дефицита и других ограничений, можно обратиться к коллективам кафедр других образовательных организаций, реализующих такие же или смежные образовательные программы. Подобное приглашение к экспертной деятельности будет служить сближению и выработке общего понимания профессионального сообщества в вопросах оценки качества подготовки специалистов. Кроме того, автором диссертации в период с 2010 по 2019 гг. была развернута работа по внутренней сертификации педагогических измерительных материалов, которая позволила начать формирование банка оценочных средств для текущего контроля, промежуточной и итоговой аттестации. Обязательным условием для положительного результата этой деятельности является личная заинтересованность в совершенствовании как своего педагогического мастерства, так и стремление к повышению качества образования. Другие условия организационно-педагогического характера описаны в следующей главе, представляющей опытно-экспериментальную работу по реализации многостадийного оценивания при итоговой аттестации.

Методика экспертизы качества содержания оценочных заданий включает следующие шаги.

*Первый шаг.* Знакомство эксперта со спецификацией измерителя.

*Второй шаг.* Выполнение оценочных заданий, содержащихся в измерителе.

*Третий шаг.* Анализ заданий с позиций:

- ясности инструкции к заданиям;
- лексической точности формулировок, грамматических ошибок;
- однозначности и недвусмысленности информации, заложенной в тексты заданий;

- корректности и привлекательности дистракторов;
- уровня трудности заданий и их корректности отнесения к минимальному, базовому или высокому уровню компетентности выпускника;
- соответствия содержания задания тем знаниям, умениям, способам деятельности и компетенциям, которые заявлены для измерения;
- уровня значимости содержания каждого задания с учетом критериально-ориентированного подхода к образовательным измерениям и исключения заданий второстепенной важности;
- отводимого времени на выполнение заданий.

Кроме того, при анализе заданий целесообразно напомнить экспертам о необходимости избегать лексической громоздкости формулировок в заданиях, стремясь к их предельной четкости, точности и корректности с учетом принятого для той или иной сферы профессиональной деятельности понятийно-терминологического аппарата.

*Четвертый шаг.* Анализ качества всего измерителя по критериям:

- полнота охвата требований соответствующего федерального государственного образовательного стандарта и образовательной программы в содержаниях оценочных заданий измерителя;
- отражение требований соответствующих профессиональных стандартов в содержании оценочных заданий через трудовые функции, трудовые действия, необходимые умения и знания;
- соблюдение пропорций содержания и их соответствие спецификации.

*Пятый шаг.* Подготовка общих выводов и рекомендаций по совершенствованию содержания оценочных заданий и измерителя в целом.

Несмотря на долю субъективности, присутствующей в такой экспертиза содержания измерителя, она способствует повышению качества разрабатываемых заданий и содержательной валидности. Как можно сделать вывод, подобные шаги достаточно трудоемкая работа и перед организаторами аттестационных испытаний

стоит нетривиальная задача подбора и мотивации экспертов на данный важный вид деятельности.

После того, как будут получены заключения от экспертов, их необходимо проанализировать и организаторам, и авторам заданий. В случае согласованности замечаний и исправлений администраторы, отвечающие за ведение банков заданий (фондов оценочных средств), вносят все изменения. После этого может происходить апробация оценочного инструментария.

В целом для повышения валидности результатов многостадийного оценивания можно выделить следующие условия [87, 156, 219]:

1. Подбор оптимальной трудности заданий для обеспечения планируемого характера распределения баллов испытуемых – смещенное нормальное распределение вправо.

2. Проведение нескольких экспертиз качества содержания измерителя.

3. Подбор валидных заданий с высокой дискриминативностью (способностью различать испытуемых).

4. Соответствие содержания измерителя его спецификации (содержательному плану).

5. Четкое формулирование существенных признаков измеряемой переменной и её отличий от других переменных, не планируемых к включению в данный измеритель.

6. Создание банка калиброванных заданий, обладающих устойчивыми статистиками, и его использование при композиции измерителя.

Таким образом, для обеспечения содержательной валидности результатов многостадийного оценивания необходимо проведение экспертизы качества содержания измерителей, предназначенных для каждого этапа оценивания. Такая экспертиза должна проводиться не формально, а с максимальной отдачей, что, бесспорно, будет давать возможность создания качественных оценочных средств и получать надежные и валидные результаты оценивания. Высокая надежность и валидность инструментария является необходимым, но не достаточным условием качества результатов образовательных измерений. Большое значение имеет также

стандартизация процедуры оценивания, которая должна обеспечивать равенство условий для аттестуемых выпускников, и это направление опытно-экспериментальной части исследования, рассматриваемое в следующей главе.

### **Выводы по третьей главе**

В случае итоговой аттестации, когда выставляемые оценки и принимаемые решения имеют ответственный и даже судьбоносный характер, оценивание целесообразно проводить в несколько этапов образовательных измерений. Это обусловлено тем, что многоэтапность измерений хотя и приводит к дополнительным трудозатратам со стороны организаторов оценивания (разработка разнообразных оценочных средств), но обеспечивает рост надежности по сравнению с использованием одного этапа измерений. Преимущества нескольких этапов измерений состоят в том, что на каждом из них используются разнообразные формы заданий, отличающиеся по своим возможностям (задания с выбором ответов, с установлением правильной последовательности или соответствия, ситуационные задания с конструируемыми ответами и так далее). И, наконец, за счет сочетания различных оценочных средств достигается репрезентативный охват требований образовательных и профессиональных стандартов, что приводит к повышению надежности и валидности результатов аттестации.

При этом технология реализации такого многостадийного оценивания рассматривается как системная совокупность взаимосвязанных методик – от планирования всего процесса оценивания, проектирования и разработки оценочного инструментария для каждого этапа оценивания, формата реализации (компьютерного или бланкового, письменного или устного) до обеспечения надежности и валидности результатов, которые получаются при его использовании в аттестационном испытании. Такое понимание технологии многостадийного оценивания при итоговой аттестации делает его органической частью всей



дидактической системы реализации образовательной программы – от замысла, разработки компетентностной модели специалиста, проектирования результатов и соответствующего содержания через учебный план и рабочие программы дисциплин (модулей) и практик до оценки качества подготовки специалистов.

Важным этапом в реализации многостадийного оценивания является разработка инструментария, предполагающего более тщательную работу авторов оценочных заданий и организаторов итоговой аттестации, чем это принято для текущего контроля и промежуточной аттестации. Этот процесс начинается с разработки спецификации, которая должна отражать всё то значимое содержание, которое подлежит операционализации и на основе которого будут создаваться оценочные задания. В целом же разработка оценочного инструментария должна проходить определенный цикл, имеющий характер незавершенности в силу необходимости совершенствования и улучшения измерителей.

Само многостадийное оценивание можно рассматривать как производное от адаптивного тестирования, в котором пошагово происходит переоценка трудности и уровня подготовленности. Алгоритм многостадийного оценивания предполагает, что на первом этапе проводится стандартизированная оценочная процедура с помощью заданий закрытой формы (с выбором ответом, установление правильной последовательности, соответствия или с кратким регламентированным ответом), на втором этапе – решение профессионально-ориентированных (ситуационных) заданий, а на третьем этапе – оценка профессиональной готовности через структурированное собеседование по полученному кейсу с членами экзаменационной комиссии. Переход аттестуемых к каждому следующему этапу возможен только, если их результат превышает установленный критериальный балл (стандарт выполнения этапа) предыдущего этапа. Аутентичность результатов оценивания обеспечивает портфолио, которое выступает подтверждающим дополнением к полученным оценкам уровня сформированности компетентности выпускника и его готовности к осуществлению профессиональной деятельности.

Надежность как одна из ключевых характеристик образовательных измерений при критериально-ориентированном подходе имеет свои особенности,

зависит от ряда факторов (характеристик измерителя, условий проведения оценивания, уровня стандартизации этой процедуры и др.). Для повышения достоверности вычисление показателя надежности следует проводить несколькими способами – с помощью  $\phi$ -коэффициента корреляции и  $\kappa$ -коэффициента согласованности. Близость этих значений и их величина около 0,8 будет говорить о надежности результатов.

Валидность показывает степень адекватности результатов измерения поставленной цели. При критериально-ориентированном подходе в итоговой аттестации на первый план выходит задача обеспечения содержательной валидности. Для этого должна проводиться тщательная качественная экспертиза разработанных оценочных заданий и измерителя в целом с последующей их корректировкой и улучшением.

## **ГЛАВА 4. ОПЫТНО-ЭКСПЕРИМЕНТАЛЬНАЯ РАБОТА ПО РЕАЛИЗАЦИИ ТЕХНОЛОГИИ МНОГОСТАДИЙНОГО ОЦЕНИВАНИЯ ВЫПУСКНИКОВ**

### **4.1. Организационно-педагогические условия реализации многостадийного оценивания выпускников**

Определение организационно-педагогических условий реализации многостадийного оценивания является необходимым, чтобы обеспечить достижение заранее поставленных целей при его использовании в итоговой аттестации. Под *организационно-педагогическими условиями* будем понимать такие *особые обстоятельства* организации многостадийного оценивания при итоговой аттестации выпускников, которые приводят к достижению, прежде всего, специальных целей, определенных в методологическом обосновании нашей работы. Именно соблюдение совокупности организационных и педагогических обстоятельств при реализации многостадийного оценивания должно обеспечить объективность и независимость аттестационной процедуры, ее эффективность, обоснованность выводов об уровне компетентности выпускников, создание благоприятной психологической атмосферы на экзамене высокой значимости. Представленные ниже организационно-педагогические условия, прошедшие апробацию в рамках опытно-экспериментальной работы и подтвердившие свою состоятельность, должны выполняться на всех этапах проведения аттестационного испытания на основе многостадийного оценивания – от планирования до принятия классификационных решений.

*1. Повышение уровня квалификации педагогических работников в области объективного оценивания и образовательных измерений.*

Можно утверждать, что это первоочередное условие, которое должно выполняться не только для организации многостадийного оценивания и не только для итоговой аттестации, а в целом для развития контрольно-оценочной

деятельности в системе образования. Безусловно, что переход на компетентностно-ориентированные стандарты выступил своего рода триггером, поскольку заставил многих администраторов и педагогических работников не только переосмыслить содержание подготовки специалистов и используемые образовательные технологии, но также искать новые подходы, доказывающие, что заявленные в стандартах компетенции сформированы у выпускников, а полученные ими оценки при итоговой аттестации релевантны оценкам работодателей. В этой связи наряду с программами повышения квалификации по теории и практике реализации компетентностного подхода в образовании были разработаны и реализовывались дополнительные программы повышения квалификации для педагогических работников в области объективных методов оценивания и образовательных измерений. Реализация таких программ позволила решить сразу несколько задач. Первая задача состояла в расширении представлений и обновления знаний профессорско-преподавательского состава о современных подходах и технологиях, применяемых в контрольно-оценочной деятельности при реализации ФГОС третьего поколения. Вторая задача, носившая прагматичный характер, предусматривала разработку оценочного инструментария на основе полученных теоретических знаний о способах формирования и оценивания компетенций. Тем самым теоретические положения компетентностного подхода переносились в практическую плоскость его реализации через проектирование образовательных результатов и способов их измерений.

Кроме того, кардинальное обновление с 2012 года образовательного законодательства и многих подзаконных актов привели к дефицитам компетенций в сфере управления образованием на основе независимых исследований качества образования, что также послужило основанием подготовки и переподготовки кадров. Под руководством и при участии автора диссертации реализовывались дополнительные образовательные программы «Управление образованием» (576 час.), «Независимая оценка качества образования: порядок проведения, анализ результатов и принятие управленческих решений» (72 час.), «Моделирование системы оценки качества образования с учетом результатов независимых

мониторинговых исследований» (72 час.), каждая из которых включала модуль, связанный с оценкой качества образования, разработкой инструментария для проведения образовательных измерений, обеспечение качества самих оценочных процедур и интерпретации результатов.

Появление в числе первых профессионального стандарта «Педагог (педагогическая деятельность в сфере дошкольного, начального общего, основного общего, среднего общего образования) (воспитатель, учитель)», утвержденного приказом Минтруда России от 18 октября 2013 г. № 544н, также выступило основой для учета требований к профессии учителя в части контрольно-оценочной деятельности. Эти требования были положены автором диссертации в основу разработанных рабочих программ дисциплин «Оценивание в образовании», «Образовательные стандарты и программы», «Современные средства оценки учебных достижений» для образовательных программ бакалавриата и специалитета по направлениям подготовки для классического университета (01.03.01 Математика, 03.03.02 Физика, 04.03.01 Химия, 04.05.01 Фундаментальная и прикладная химия, 06.03.01 Биология, 44.03.01 Педагогическое образование, 44.03.02 Психолого-педагогическое образование, 45.03.01 Филология) и дополнительную профессиональную программу профессиональной переподготовки с присвоением квалификации «Преподаватель старшей школы в условиях ФГОС общего образования».

Также в силу установленных приказом Минтруда России от 12 апреля 2013 г. № 148н «Об утверждении уровней квалификации в целях разработки проектов профессиональных стандартов» соответствий уровней квалификаций и путей их достижения через освоение образовательных программ, включая программы уровня магистратуры, и требований ряда профессиональных стандартов к 7 квалификационному уровню по определенным профессиям, для образовательных программ магистратуры по психологическим и психолого-педагогическим направлениям подготовки были разработаны программы по дисциплинам «Измерения в образовании», «Нормативно-правовое обеспечение психолого-педагогической деятельности».

Указанные организационные и содержательные изменения, которые затронули пересмотр структуры, формы и содержания подготовки и переподготовки педагогических кадров, привели к большому пониманию и лучшему освоению современных подходов и требований к контролю и оценке, а также развитию культуры оценочной деятельности. Это подтверждается используемыми в образовательном процессе оценочными заданиями различных форм для разных направлений (44.03.02 Психолого-педагогическое образование, 39.03.01 Социология, 31.05.01 Лечебное дело), а также в системе аккредитации специалистов в области здравоохранения. Часть результатов такой работы легла в основу формирования паспортов компетенций и создания фондов оценочных средств. Всего за период с 2012 по 2020 гг. по различным видам программ прошло обучение и переподготовку 370 человек из числа педагогических работников, включая профессорско-преподавательский состав, и 428 студентов.

*2. Наличие банка калиброванных оценочных заданий, разработанных с учетом требований профессиональных стандартов.*

Данное условие также является обязательным для реализации многостадийного оценивания при итоговой аттестации, и оно является следствием выполнения первого условия. Кроме того, как кратко приводилось в третьей главе, обеспечение этого условия возможно за счет различных форм и мер поддержки коллективов кафедр, ведущих подготовку специалистов по тем или иным образовательным программам. Наконец, объединение усилий нескольких коллективов различных университетов также приводят к положительному синергетическому эффекту и формированию таких фондов оценочных средств по родственным или смежным образовательным программам.

Вместе с этим, необходимо понимать, что задание становится калиброванным только после апробации и математико-статистической обработки результатов его выполнения представительной группой студентов. Очень часто в условиях малокомплектности академических групп, специфики содержания самих образовательных программ и отдельных профессиональных компетенций, а также неготовности к межвузовскому взаимодействию приходится идти по пути

экспертной оценки качества разработанных авторами заданий, чтобы отнести их к категории калиброванных. Однако, эти трудности вполне преодолимы, если, как в случае развития системы аккредитации специалистов в области здравоохранения, существует общее согласованное понимание необходимости формирования таких банков на репрезентативных выборках студентов. В этой связи в данном исследовании опытно-экспериментальная работа велась по двум направлениям. Первое связано с апробацией заданий для крупномасштабных оценочных процедур (таких, как аккредитация специалистов в области здравоохранения) с их последующей математико-статистической обработкой и калибровкой. Так, автор диссертационного исследования участвовал в 2019 г. в масштабной апробации оценочного инструментария, в которой приняло участие свыше 10 000 студентов медицинских вузов по специальности 31.05.01 Лечебное дело, и обработке результатов [231].

В целях обеспечения возможности математико-статистической обработки и интерпретации результатов тестирования автором диссертации предлагались практические рекомендации и комплекс программ, предназначенных для разработчиков, которые стремятся к овладению методами создания контрольно-оценочных средств и совершенствованию используемых в учебном процессе педагогических тестов. Такое руководство, разработанное при ведущей роли дипломированного специалиста-тестолога Щанициной С. В., послужило качественным основанием для понимания технологии математической обработки результатов тестирования в рамках классической и современной теорий тестов, а также способом получения количественной и качественной интерпретации результатов апробации заданий. Сами практические рекомендации представлены в Приложении 9.

Второе направление опытно-экспериментальной работы было связано с апробацией и формированием банка стандартизированных заданий закрытой формы для многостадийного оценивания при итоговой аттестации по конкретным направлениям подготовки (44.03.02 Психолого-педагогическое образование, 39.03.01 Социология). Первая стадия оценивания, которая по своему характеру

является репродуктивной, требовалось отобрать такие задания по профессионально значимым дисциплинам в будущей деятельности выпускника, которые бы соответствовали критериально-ориентированному подходу, т. е. характеризовались невысокой трудностью и соответствовали минимальному уровню компетентности. Формирование банка таких заданий происходило в процессе выполнения студентами диагностических работ при проведении мероприятий внутренней системы оценки качества образования. Эта процедура достаточно стандартизирована, если стандартизированы и сами оценочные задания (например, с выбором ответов), результаты выполнения которых получались при прохождении студентами бланкового или компьютерного тестирования. При этом автором диссертации и коллегами отстаивалась следующая позиция. Первый этап многостадийного оценивания на итоговом экзамене, объем которого является достаточно обширным по содержанию проверяемых знаний и умений, должен содержать посильные для выполнения большинством студентов задания.

Исследовательский и методический интерес представляли задания для второго и третьего этапов оценивания. Их структура, выбор формы заданий и содержательные основания формировались следующим образом. Поскольку открытые задания не предусматривают автоматическую проверку, предполагают только экспертную оценку их качества, то это условие означает тщательную экспертизу каждого такого задания еще на этапе разработки спецификации и обеспечения содержательной валидности результатов. Только после согласованного положительного заключения о надлежащем качестве задания (например, кейсов) от трех экспертов, его можно включать в банк калиброванных заданий для соответствующего этапа оценивания.

Вместе с тем, для определения наиболее оптимальной формы заданий для второго и третьего этапов оценивания в рамках промежуточной аттестации студентов по дисциплинам психолого-педагогического цикла образовательной программы «Психология образования» по направлению 44.03.02 Психолого-педагогическое образование проводилась апробация следующих форм:



1) беседа-обоснование выбора студентом ответов на основе анализа практической ситуации (ситуационные задания с выбором ответов);

2) свободное собеседование по проблеме, сформулированной самим студентом;

3) структурированная беседа по результатам анализа студентом предложенного кейса;

4) оценка портфолио студента.

*Первая форма*, связанная с обоснованием выбора студентом ответов, предусматривала несколько вариантов (6–7) выходов из предложенной практической ситуации на основе ее анализа. Студенту необходимо было проанализировать содержание, выбрать два варианта, которые, с его точки зрения, являются наиболее продуктивными и привести аргументы к обоснованию выбора (устно или письменно). Такая форма была полезна на этапе письменного варианта государственного экзамена до многостадийного оценивания, чтобы можно было по ответам выпускников накапливать информацию и проводить корректировку формулировок и содержания заданий. Пример такого задания приводится ниже.

Ситуация. *«В ответ на соответствующее замечание учителя ученик говорит, что для того, чтобы усвоить учебный предмет, ему не нужно много работать: «Меня считают достаточно способным по жизни человеком».*

Что должен ответить ученику на это учитель?

1. *«Это мнение, которому ты вряд ли соответствуешь».*

2. *«Те трудности, которые ты до сих пор испытывал, и твои знания отнюдь не говорят об этом».*

3. *«Многие люди считают себя достаточно способными, но далеко не все на деле таковыми являются».*

4. *«Я рад(а), что ты такого высокого мнения о себе».*

5. *«Это тем более должно заставить тебя прилагать больше усилий в учении».*

6. *«Это звучит так, как будто ты сам не очень веришь в свои способности».*

В данном ситуационном задании важную роль играет возможность двух выборов. Тогда аргументация будет строиться и на их сравнении.

*Вторая форма* – свободное собеседование по проблеме, которая ставится самим студентом на основе содержания, выносимого на экзамен. Для этого необходимо было сформулировать требования, которые затем и становились основой для оценки самого собеседования. В нашем случае были выделены следующие критерии такого свободного собеседования:

- 1) Формулировка проблемы, обоснование выбора;
- 2) Ориентация в понятийном поле;
- 3) Основные теоретические идеи и *их объяснения*;
- 4) Основные методические идеи и *их разъяснение*;
- 5) Практика (примеры, ситуации);
- 6) Выводы;
- 7) Личное мнение;
- 8) Рефлексия.

Преимущества данной формы состоят в индивидуализации экзамена для каждого студента, в раскрытии творческого потенциала и коммуникативных возможностей студента вступать в диалог с преподавателем. Это выражается как в формулировке каждым студентом своего проблемного вопроса, отражающего его интересы и предпочтения, так и в отборе и форме представления материала по решению данного вопроса. Студенты при этом демонстрируют умения:

- формулировать проблемные вопросы;
- использовать практический материал (собственные истории, ситуации, примеры из практической деятельности);
- осуществлять рефлексивную деятельность;
- вступать в устный диалог с преподавателем (экзаменатором).

*Третья форма*, предусматривающая структурированную беседу по результатам анализа студентом предложенного кейса, проводится по заданной

структуре, подробно описанной в разделе 3.1 третьей главе и представленной на рис. 20. Приведем пример такого кейса.

*Кейс: Аня (3 года) стала называть себя Дарьей. Мама, подыгрывая дочери, стала тоже называть ее Дарьей. Однако мама ни разу не спросила, почему девочке вдруг захотелось сменить имя, и тем самым не желать быть самим собой.*

*Родители приняв игру дочери, укрепили ее в мысли стать кем-то другим и позволить вести себя так, будто у нее другие родители.*

Вопросы:

1. В чем могут быть причины такой ситуации?
2. Как Вы оцениваете поведение родителей?
3. Какие последствия в развитии личности девочки возможно могут проявиться в дальнейшем и как скорректировать их негативные проявления?

Критерии, которые использовались для оценки выполнения кейса представлены в таблице 13.

Таблица 13

#### Критерии оценки выполнения кейса

Критерии	Оценка
Даны полные и правильные ответы на все <i>три</i> вопроса, аргументировано дано разъяснение теоретическим позициям, высказано свое отношение по спорным вопросам. Студент ясно, четко и доказательно излагает вариант решения.	<i>отлично</i>
Даны полные и правильные ответы на <i>два</i> вопроса, аргументировано дано разъяснение теоретическим позициям, высказано свое отношение по спорным	<i>хорошо</i>

Критерии	Оценка
вопросам, но могут быть допущены ошибки, некорректность объяснения. Или даны ответы на все три вопроса. Но допускаются неточности и существенные ошибки.	
Даны полные и правильные ответы на <i>один</i> вопрос, или даны ответы на два, или даже на все три, но не аргументировано, не дано разъяснение теоретическим позициям, не высказано свое отношение по спорным вопросам, допущены существенные ошибки, некорректность объяснения. Решение представлено не доказательно.	<i>удовлетворительно</i>
Ни один из вопросов не раскрыт полно и правильно, или содержание ответа не соответствует вопросу, не может пояснить и аргументировать свою точку зрения, подвести теоретическую базу, допускает грубые ошибки и неточности.	<i>неудовлетворительно</i>

Анализ выполнения студентами заданий по предложенным формам показал, что самой трудной и объемной является вторая форма – свободное собеседование. В тоже время студенты отмечали, что это самая творческая форма, вызывающая интерес у наиболее подготовленных и заинтересованных обучающихся.

Выбор определенной формы для второго и третьего этапа многостадийного оценивания был обусловлен различными факторами, к числу которых относились уровень подготовленности студентов и выпускников; содержание заданий (однородные, междисциплинарные, полидисциплинарные по своему характеру) и время, отводимое на их выполнение; возможность индивидуализации и дифференциации с помощью оценочных заданий.

Отметим, что все апробированные формы в процессе аттестации студентов и оценивании результатов обучения отвечают положениям доказательного подхода,

хотя и требуют для его полнокровной реализации перевода в компьютерную автоматизированную систему. Вместе с тем, учитывая роль оценок при итоговой аттестации и значимость принимаемых решений исследовательская позиция автора диссертации, подкрепляемая экспертными суждениями педагогов и практиков, предполагает, что нельзя исключать из завершающего этапа оценивания человека.

Исходя из результатов трехгодичной апробации (2014, 2015, 2016 гг.) таких форм в рамках промежуточной аттестации студентов, обучающихся по образовательной программе бакалавриата «Психология образования» (32 студента) и с учетом экспертной оценки, на второй стадии оценивание был сделан выбор в пользу профессионально-ориентированных (ситуационных) заданий, которые можно перевести в стандартизированную форму компьютерного тестирования. Для третьего этапа оценивания было определено структурированное интервью с членами экзаменационной комиссии.

*3. Наличие компьютерных программ либо программно-инструментальной среды (сервиса, платформы), обеспечивающих реализацию алгоритмов многостадийного оценивания, а также получение обратной связи от выпускников и работодателей для развития и совершенствования аттестационных процедур и качества подготовки специалистов.*

Данное условие можно рассматривать применительно как в целом к многостадийному оцениванию, так и к отдельным компонентам его реализации. Современный уровень развития компьютерных технологий и открывшихся технологических возможностей допускают интеграцию всех составляющих многостадийного оценивания – от непрерывного оценивания качества используемых заданий до моментальной обработки результатов выполнения их испытуемыми и предъявление окончательной оценки.

Модели современной теории тестов IRT необходимы для того, чтобы на количественном уровне измерений, работая на интервальной шкале, получать оценки трудности заданий и уровня подготовленности в единых единицах измерения – логитах. Это позволяет отбирать для измерения уровня компетентности аттестуемого выпускника в каждом субтесте такие задания,

которые оптимальны по трудности в интервале значений около установленного критериального (порогового) балла для соответствующего этапа.

Первые два этапа представленной технологии многостадийного оценивания реализуются в компьютерном формате без участия человека, что обеспечивает повышение точности образовательных измерений по описанным выше основаниям. В силу особенностей итоговой аттестации и ее высокой значимости для получения документов об образовании и присвоения выпускнику квалификации, подтверждающей соответствие требованиям образовательных стандартов, в процедуре компьютерного оценивания должны быть выполнены некоторые допущения. А именно, речь идет о запрограммированной возможности каждому аттестуемому выпускнику переходить от задания к заданию в каждом субтесте при прохождении первого этапа, соответствующего минимальному уровню компетентности, и в случае затруднения пропустить такое задание, а затем к нему снова вернуться. Также испытуемые должны видеть количество всех заданий по каждому субтесту, необходимых для выполнения. Наконец, после завершения первого этапа выпускник должен видеть на экране свой результат выполнения каждого субтеста, общую оценку за первый этап оценивания и должна быть предусмотрена возможность апелляции в случае его несогласия с результатами. При преодолении критериального (порогового) балла аттестуемый должен иметь возможность перейти к следующему этапу оценивания, где аналогично будут соблюдаться такие же возможности, как и на первом этапе.

Приведенные требования обоснованы и согласованы при обсуждении их как со студентами, которым предстояло проходить аттестацию в таком формате, так и с ведущими преподавателями, имеющими большой стаж работы и высокий экспертный статус. Более того, пропуск задания и возможность к нему впоследствии вернуться создают более комфортную ситуацию, с точки зрения некоторого снижения психологического напряжения во время оценочной процедуры.

В силу того, что нормативные требования предполагают обязательное наличие и функционирование в каждой организации высшего образования

электронной информационной образовательной среды (ЭИОС), то описанное требование по реализации компьютерного формата проведения первых двух этапов многостадийного оценивания может быть легко реализуемо, но при соблюдении первых двух условий и готовности самой образовательной организации.

Кроме того, целесообразно заложить в качестве одной из возможностей ЭИОС при ее проектировании – опцию для формирования студентом своего портфолио. Оно будет доступно для просмотра и оценки членами экзаменационной комиссии, участвующими в итоговой аттестации.

Также значимым компонентом в системе оценки качества образования (как внешней, так и внутренней) является обратная связь от представителей работодателей, участвующих в реализации образовательной программы. Обработка результатов и их учет крайне важны при принятии различного рода административных и содержательных решений в части реализации и совершенствования образовательной программы. Пример такого опроса для выпускников представлен в Приложении 10, для работодателей – в Приложении 11.

В целом описанные возможности изначально закладывались коллективом Ивановского государственного университета при проектировании ЭИОС «Мой университет», расположенной по адресу: <https://uni.ivanovo.ac.ru/>.

*4. Наличие спецификации к оценочному инструментарию для аттестационных процедур, которая обеспечивает содержательную валидность результатов образовательных измерений.*

Это также следует отнести к организационно-педагогическим условиям реализации многостадийного оценивания при итоговой аттестации.

Данное условие достаточно подробно раскрыто в третьей главе при описании того, как обеспечивать содержательную валидность результатов, но важно подчеркнуть следующее. Говоря о содержательной валидности, следует отдавать отчет в том, что речь идет, в первую очередь, о качественной экспертизе оценочных заданий, которые должны попасть в банк для проведения многостадийного оценивания. Поэтому этап разработки спецификации, включающий обязательность

соотнесения измеряемых компетенций с индикаторами их проявления, обоснованность и вклад каждого задания в общую оценку уровня компетентности выпускника, учет при необходимости положений профессионального стандарта в части описания трудовых функций специалиста является архиважным для итоговой аттестации. В этой связи для выполнения данного условия необходимо привлечение экспертов к такой работе. В их роли могут выступать либо работодатели, принимающие студентов в свои организации и на предприятия для прохождения практической подготовки, либо опытные и авторитетные преподаватели-практики, четко представляющие профессиональную сферу деятельности выпускников.

*5. Наличие согласованного регламента проведения аттестационного испытания на основе многостадийного оценивания.*

Это условие предусматривает четкую регламентацию всей процедуры многостадийного оценивания, которая оформляется в виде документа, доступного для ознакомления всеми выпускниками. Такой регламент должен включать порядок прохождения многостадийного оценивания, время на выполнение заданий каждого этапа оценивания, критерии оценивания, предусматривать возможность прохождения пробного оценивания, а также случай апелляции, если выпускник будет не согласен с полученной оценкой. К традиционным консультациям, которые проходят перед аттестационными испытаниями (государственным экзаменом), следует разработать инструкции для выпускников по прохождению многостадийного оценивания и предоставить возможность пройти репетиционный экзамен на основе этой технологии.



## 4.2. Шкалирование результатов и установление критериальных баллов для многостадийного оценивания

Шкалирование при критериально-ориентированном подходе в образовательных измерениях при итоговой аттестации выпускников имеет своей целью установление уровня компетентности выпускников. Оно сводится к необходимости установления критериального (порогового) балла [154, 157]. В многостадийном оценивании, когда процедура осуществляется в несколько этапов, переход к следующему этапу оценивания происходит после того, как аттестуемый выпускник преодолел пороговый балл на текущем этапе.

Критериальный (пороговый) балл может быть задан в описательной форме или выражен в виде числа или процента на некоторой шкале. В случае выбора некоторого числа или процента критерий называют *проходным баллом*, позволяющим разделять выпускников на две группы – аттестованных и не аттестованных. В данном исследовании выбор сделан в пользу совмещения качественных и количественных критериев с целью построения уровневой вертикальной шкалы, на которой качественные оценочные суждения об уровне сформированности компетентности выпускника (минимальный, базовый, высокий) соотносятся с пороговым баллом, представляющим собой процент правильно выполненных заданий, и оценкой по пятибалльной шкале. Оценка по пятибалльной шкале, а точнее по четырехбалльной – «отлично», «хорошо», «удовлетворительно» и «неудовлетворительно», необходима для выполнения нормативно установленных правил при итоговой аттестации [189].

Задача установления порогового балла на итоговой аттестации не нова, но она приобретает иное звучание в данном исследовании в силу нескольких обстоятельств. Во-первых, проведение аттестационного испытания – государственного экзамена – не в традиционном формате по ответам аттестуемых выпускников на 2-3 вопроса случайно полученных билетов в устной или письменном формах, а прохождение нескольких этапов оценивания. Во-вторых, каждый этап отличается по уровню трудности, и это необходимо учитывать при

установлении порогового балла на каждой итерации оценивания. В-третьих, компетентностная трактовка образовательных результатов с учетом требований профессиональных стандартов должна быть отражена и в правилах установления критериальных баллов, и получения тех или иных оценок, понятных каждому выпускнику и принятых теми, кто проводит итоговую аттестацию. Конечно, существует позиция, что нововведения, предусматривающие усложнение ряда привычных процедур, включая выставление традиционных оценок, не принесут значимого результата в деле подготовки специалистов и повышение качества образования. Но исследовательская и практическая позиция автора данного исследования, подкрепленная экспертными позициями и эмпирическими результатами, свидетельствует об обратном. Чем больше строгих методов оценивания с четко детерминированными критериями выставления тех или иных оценок будет применяться в аттестационных процедурах высокой значимости, тем выше будет ответственность за качество подготовки специалистов и тем меньше будет рекламаций со стороны работодателей и сферы труда в целом к системе высшего образования.

Разработка описательных критериев начинается с анализа наблюдаемых признаков проявления оцениваемых переменных, в роли которых при итоговой аттестации выступают компетенции, а также способность выполнять трудовые функции, если учитываются требования профессионального стандарта. Примеры такого соотнесения профессиональных компетенций образовательной программы «Психология образования» по направлению подготовки 44.03.02 Психолого-педагогическое образование и трудовых функций из профессионального стандарта «Педагог-психолог (психолог в сфере образования)» представлены в Приложении 12.

Пороговый балл, являющийся границей между аттестованными и не аттестованными испытуемыми, выступает целевым численным эквивалентом, свидетельствующим о достижении необходимого уровня. Переход за него означает, что балл аттестуемого превысил необходимый уровень, а сам выпускник обладает необходимым уровнем компетентности. Процесс установления

порогового балла состоит в обосновании выбора точки на шкале, позволяющей с большой вероятностью гарантировать, что все испытуемые с баллами, больше чем пороговый балл, обоснованно попали в группу аттестованных.

Научный интерес к определению методов установления пороговых баллов возник с развитием критериально-ориентированного подхода к интерпретации результатов образовательных измерений в 60-е годы XX века.

Все методы установления порогового балла на шкале наблюдаемых оценок делятся на две группы, которые отличаются по отношению к ошибкам классификации [125, 203, 219, 290]. Первая группа методов решает задачу минимизации вероятностей ошибочной классификации при предположении о том, что все ложные положительные и ложные отрицательные ошибки одинаково серьезны при принятии решений о распределении испытуемых по уровням на шкале. Вторая группа методов связана с минимизацией ожидаемых потерь из-за ошибочной классификации. Эти методы используют пороговую функцию потерь, которая предполагает, что внутри классов ложных положительных и ложных отрицательных ошибок не устанавливается никаких различий, но эти ошибки отличаются по значимости друг от друга.

Также методы установления пороговых баллов делят по другим основаниям. Первое основание центрировано на результатах выполнения измерителя испытуемыми, а второе – сам измеритель и входящие в него оценочные задания. В первом случае проводится анализ оценок параметров испытуемых, т.е. уровня подготовленности, а во втором случае в основе экспертных суждений относительно порогового балла лежит анализ содержания заданий. Как и в теории тестов, первое основание связано с нормами, а второе – с критериями.

Процедура установления порогового балла всегда связана с экспертной оценкой и, как правило, это происходит еще на этапе разработки заданий для оценивания. На первый взгляд, может показаться, что это сугубо субъективный подход. Однако, для минимизации этой субъективности следует (1) в обязательном порядке проводить валидизацию выбранных пороговых баллов после эмпирической апробации и (2) использовать несколько методов, соотнося в

последующем согласованность экспертных позиций. Критериально-ориентированный подход к установлению пороговых баллов предусматривает работу экспертов по трем направлениям.

Первое направление работы эксперта, а лучше группы экспертов, основывается на целостном впечатлении содержания оценочного инструментария, которые провели исследование и, основываясь на собственном впечатлении, предлагают пороговый балл, т.е. процент заданий, с которыми должен справиться аттестуемый выпускник с минимальным или базовым уровнем компетентности в зависимости от этапа многостадийного оценивания. Как правило, каждый эксперт устанавливает свой критерий, а для принятия окончательного решения берется среднее значение. Второе направление связано с экспертизой каждого оценочного задания. Третье направление, которое является логическим продолжением первых двух, предусматривает анализ эмпирических результатов выполнения заданий, по результатам которого делается вывод о содержании заданий инструментария.

Этап определения порогового балла можно считать определяющим, поскольку от него зависит валидность окончательного аттестационного решения, а присвоение квалификации выпускнику будет означать, что он на минимальном, базовом или высоком уровне компетентен для выполнения трудовых функций по определенному виду профессиональной деятельности.

Трудно переоценить роль порогового балла при итоговой аттестации для любой сферы деятельности, но особенно остро этот вопрос стоит, когда речь идет о национальной безопасности, здоровье людей, их образовании. Другими словами, при определении порогового балла необходимо учитывать политические, социальные, экономические и другие контексты. Конечно, существует искушение административным решением снизить пороговый балл, за который пройдут все аттестуемые выпускники, создавая иллюзию высокого качества подготовки специалистов. Другая крайность, это установление высоких пороговых баллов и усложнение всей процедуры аттестационного испытания, тем самым получение большого числа неаттестованных выпускников. Всё это лишь свидетельствует о необходимости содержательного и подробного обоснования выбора порогового

балла в совокупности с высоким качеством разработанного оценочного инструментария, учетом множества контекстных факторов и анализа последствий.

В целом для установления порогового балла при многостадийном оценивании можно рассматривать несколько вариантов. В исследовании отдано предпочтение первому варианту, когда на каждом этапе многостадийного оценивания определяется свой пороговый балл, подлежащий обязательной эмпирической валидации. Но можно выделить еще два варианта, которые могут быть уместны, если аттестационное испытание проводится в один или два этапа, или организаторы аттестационных процедур и создатели оценочного инструментария находятся в начале пути по совершенствованию контрольно-оценочной деятельности и только начинают формировать банки оценочных заданий. Итак, второй вариант может предусматривать установление единственного порогового балла для высокого уровня компетентности, достижение которого и будет свидетельствовать об аттестации выпускника. Такой подход может быть реализован для сферы деятельности с очень высокими ставками (например, в области безопасности в той или иной отрасли экономики). Третий вариант установления порогового балла связан исключительно с необходимостью получения определенного результата за выполнение заданий различной трудности, которые не ранжированы ни по уровням компетентности, ни по сложности. Такой вариант ближе к оценочным процедурам при текущем контроле и промежуточной аттестации, если не используются специальные алгоритмы IRT, учитывающие вклад каждого задания в общую оценку уровня подготовленности.

Как уже отмечалось при обосновании образовательных измерений, ошибки неизбежны в любых оценочных процедурах. И чем больше детерминированы методы и соблюдаются все необходимые сопутствующие измерительной процедуре условия, тем точнее будет результат и меньше возможность ошибки. В общем случае эта проблема решается либо за счет многократного измерения, что невозможно в случае итоговой аттестации, либо в удлинении оценочного средства, что также приведет к увеличению времени на прохождение аттестационного испытания и быстрой утомляемости испытуемых. В литературе по критериально-

ориентированному подходу отмечается, что вокруг порогового балла всегда существует «нейтральная зона» – 10% слева и 10% справа, в которую могут попасть аттестуемые, классифицированные как не прошедшие за пороговый балл [203, 283]. Такая ошибка классификации называется ложно отрицательной (неправильная неаттестация). А в случае, если испытуемый не должен быть отнесен в группу аттестованных, но оказался прошедшим за пороговый балл, то такая ошибка классификации называется ложно положительной (неправильная аттестация).

Главный вывод, который необходимо сделать при рассмотрении ошибок классификации, заключается в следующем. Классификационные решения по тем аттестуемым выпускникам, кто попал в «нейтральную зону», необходимо аргументировать на основе дополнительной информации о каждом из них. Эта информация может представлять из себя определенные характеристики на выпускника, его успеваемость за время освоения образовательной программы, наконец, портфолио. Кроме того, у аттестуемого, который не согласен с результатами аттестационного испытания и, следовательно, своей классификацией, остается право на апелляцию.

Результаты опытно-экспериментальной работы показали, что такие ошибки присущи, в первую очередь, в крупномасштабных оценочных процедурах, к числу которых относится система аккредитации специалистов в области здравоохранения. В условиях итоговой аттестации выпускников по образовательным программам высшего образования, как правило, это исключительные ситуации и при анализе классификационной ошибки, из которых чаще встречаются ложно отрицательные, проведения апелляции вопрос решается положительным образом и в пользу выпускника.

В целях минимизации ошибок классификации, как правило, пользуются несколькими методами установления пороговых баллов. Не существует строгой иерархии этих методов применительно к используемому в аттестационной процедуре оценочного инструментария, но можно выделить наиболее

распространенные методы, которые приняты профессиональным сообществом в области образовательных измерений [203, 283, 290, 294, 310].

*Метод Недельски*, используемый только для закрытых заданий с выбором одного ответа. Суть этого метода заключается в определении экспертами вероятности правильного выполнения задания той частью испытуемых, уровень подготовленности которых находится на границе между компетентными и не компетентными. После анализа заданий и вариантов ответов к ним, эксперты исключают такие варианты, которые, по их мнению, гипотетический минимально компетентный аттестуемый никогда не выберет. Величина, получаемая из соотношения одного правильного ответа к оставшимся неправильным, рассматривается как вероятность того, что минимально компетентный испытуемый верно выполнит задание. Затем находится среднее значение по каждому заданию по полученным от экспертов величинам, т. е. используется усредненная величина по всем их суждениям [87, 159]. Далее находится сумма величин Недельски по всем заданиям и это будет пороговый балл, который должны преодолеть компетентные аттестуемые. Этот метод не лишен ряда недостатков и ограничений.

*Метод Ангоффа и его модификации* достаточно распространены в процедурах итоговой аттестации. Это обусловлено его простотой и тем обстоятельством, что он зарекомендовал себя как высоко согласованный и надежный метод получения оценок пороговых баллов. Метод Ангоффа также начинается с работы экспертов, которые анализируют содержание заданий. По результатам анализа они предлагают значения долей минимально компетентных аттестуемых, которые могут выполнить верно каждое задание. В основном методе Ангоффа каждый эксперт рассматривает задание за заданием, представляя себе гипотетического «минимально компетентного испытуемого» и принимая решение о том, мог ли такой испытуемый ответить правильно на очередное задание теста. В случае положительного решения эксперт присваивает значение, равное 1, каждому заданию, а при отрицательном решении — нулевое значение (0). В этом случае сумма оценок по заданиям будет равняться гипотетической не преобразованной

оценке, полученной «минимально компетентным испытуемым». Далее, анализируя каждое задание, эксперт должен представить уже группу минимально компетентных испытуемых и предложить вероятность того, что «минимально компетентный испытуемый» ответит правильно для каждого задания оценочного средства. Оценка вероятности для каждого задания дается в виде доли минимально компетентных испытуемых, которые ответят на это задание верно. Сумма этих вероятностей по всем заданиям выбирается в качестве порогового балла по оценочному средству. Идея о возможностях минимально компетентного испытуемого является ключевой для основного метода Ангоффа, поэтому успешность его применения при аттестации зависит от того, насколько хорошо подготовлены эксперты к такой практической работе [87, 159].

Таким образом, данный метод предусматривает участие квалифицированных преподавателей и практиков, подробно разбирающихся в будущей профессиональной деятельности подлежащего аттестации выпускника и способных решить предлагаемые оценочные задания. В работах по установлению стандартов [87, 290] отмечается, что в настоящее время преимущественно используются модификации основного метода Ангоффа. Эта модификация метода Ангоффа предполагает несколько итераций, в ходе которых эксперт каждый раз определяет вероятность правильного выполнения каждого задания минимально компетентным выпускником. В практике, как правило, принято проводить три таких итерации, при этом допускается обмен мнениями между экспертами перед началом следующей итерации, который приводит к эффекту «конвергенции» экспертных пороговых оценок. Считается, что тем самым уменьшается рассогласованность экспертных оценок и повышается ответственность самих экспертов. Несомненным достоинством метода является его возможность использовать для заданий различных форм. Для заданий с конструируемым ответом используют так называемый *расширенный метод Ангоффа*. Вместо выставления априорных оценок вероятности правильного выполнения заданий «минимально компетентными испытуемыми» эксперты предлагают оценки к заданиям, которые, как они полагают, получают пограничные испытуемые по



каждому заданию с конструируемым ответом. Эти процедуры могут быть объединены для случаев, когда измеритель включает сочетание различных форм заданий, нуждающихся как в дихотомических оценках, так и в выборе баллов из оценочных рубрик для политомически оцениваемых заданий. Пороговые баллы для расширенного метода Ангоффа вычисляются таким же образом, как в оригинальном и модифицированном методах. В облегченном варианте метода Ангоффа, получившем название метод «Да / Нет», каждому заданию эксперты присваивают одно из двух суждений. Если испытуемый в состоянии верно выполнить задание, то ему присваивается «Да», в противном случае «Нет». К несомненным преимуществам модифицированного метода Ангоффа и метода «Да / Нет» относятся их простота и возможность применения для оценочных средств с различными формами заданий, предполагающими как дихотомические, так и политомические оценки. Также модифицированный метод Ангоффа используется для установления нескольких пороговых баллов тогда, когда это необходимо в рамках одного этапа оценивания [87, 290].

*Метод Эбея*, используемый при критериально-ориентированном подходе в образовательных измерениях, допускает для установления порогового балла использование различных заданий – закрытых с выбором ответа и открытых со свободно конструируемым ответом. Так же, как и в методе Недельски, должны быть определены эксперты, которые выносят суждения по каждому заданию отдельно, устанавливая уровень его трудности (как правило в процентах). И кроме этого, они должны оценить содержательную значимость задания относительно поставленной цели для данного оценочного инструментария. Для использования этого метода составляется специальная двумерная таблица с оценочными категориями, в которую эксперты заносят свои оценки по трудности (легкое, среднее, трудное) и значимости содержания (наиболее существенное, важное, допустимое, сомнительное) каждого задания, пример которой показан ниже (табл. 14).

**Общий вид таблицы для установления порогового балла  
по методу Эбеля**

<b>Уровень трудности задания</b>	<b>Уровень значимости содержания</b>				<i>Итого</i>
	Наиболее существенный	Важный	Допустимый	Сомнительный	
Легкий					
Средний					
Трудный					
<i>Общее число</i>					

Заполнение экспертами такой таблицы предусматривает, что они имеют четкое представление как о выпускниках, которые будут проходить каждый этап многостадийного оценивания, так и целях аттестационной процедуры для принятия решений об их классификации. К заполненным данным по двум категориям добавляется процент заданий, на которые должен правильно ответить аттестуемый с минимальным или базовым уровнем компетентности в зависимости от того, для какого этапа оценивания проводится процедура определения порогового балла по методу Эбеля. При определении процента выполнения заданий стремятся получить согласованное мнение всех экспертов, участвующих в определении порогового балла. Для получения порогового балла по методу Эбеля суммируют произведение числа заданий по каждой категории и соответствующий процент их верного выполнения, а затем делят на общее число полученных экспертных суждений. Число экспертных суждений – это количество экспертов, умноженное на количество тестируемых заданий. Пример расчета порогового балла по методу Эбеля для отдельного субтеста по дисциплине «Педагогика и педагогическая психология», содержащего 50 заданий, для первого этапа многостадийного оценивания при итоговой аттестации представлен в таблице 15.

Примечательно, что данный метод можно распространить на ситуацию, когда вместо экспертных оценок процента правильных ответов на задания используются эмпирические оценки трудности заданий, т. е. доля верно выполнивших задание при его апробации. В данном исследовании в рамках опытно-экспериментальной работы использовались полученные при апробации величины трудности заданий во время выполнения студентами диагностических работ, которые включали задания субтестов по дисциплинам, входящим в процедуру многостадийного оценивания при итоговой аттестации. Преимущество метода Эбеля заключается в том, что его можно применить для заданий как с дихотомической, так и полиномической оценкой результатов их выполнения.

Обязательным требованием при определении порогового балла одним из представленных методов является высокие требования к привлекаемым для этого экспертам. Во-первых, эксперты должны представлять, какой контингент выпускников будет проходить оценивания по предлагаемому инструментарию и понимать цели итоговой аттестации. Во-вторых, эксперты должны быть специалистами в той сфере профессиональной деятельности, к которой готовились обучающиеся, осваивая образовательную программу. В-третьих, они должны быть готовы аргументировано, беспристрастно и с высокой долей объективности оценивать как отдельные оценочные задания, так и весь измеритель, предназначенный для итоговой аттестации выпускников.

Таблица 15

**Расчет порогового балла по методу Эбеля для отдельного субтеста по дисциплине  
«Педагога и педагогическая психология» для первого этапа многостадийного оценивания**

Уровень значимости содержания	Уровень трудности задания	Эксперт 1	Эксперт 2	Эксперт 3	Число заданий, оцененных как принадлежащие различным категориям (А)	Оценка процента правильных ответов на задание (В), %	Произведение (А × В)
<b>Наиболее существенный</b>	Легкий	7	8	7	22	100	2200
	Средний	0	0	1	1	70	70
	Трудный	0	0	0	0	25	0
	<i>общее число</i>				23	-	
<b>Важный</b>	Легкий	10	11	12	33	85	2805
	Средний	9	2	2	13	40	520
	Трудный	2	1	1	4	25	100
	<i>общее число</i>				50		
<b>Допустимый</b>	Легкий	15	14	15	44	80	3520
	Средний	3	10	6	19	40	760
	Трудный	1	1	2	4	25	100
	<i>общее число</i>				67		
<b>Сомнительный</b>	Легкий	2	1	2	5	50	250
	Средний	1	2	2	5	40	200
	Трудный	0	0	0	0	30	0
	<i>общее число</i>				10		
<i>Общее число суждений</i>		50	50	50	150		10525

**Пороговый балл      70,17**

Очень часто можно встретить ситуацию, когда критериальный или пороговый балл определяется экспертным путем просто «на глазок», исходя из собственной позиции экспертов. Достаточно долго в аттестационном тестировании, которое проходило в рамках процедур государственной аккредитации образовательных организаций, таким критериальным баллом выступало условие, что в тестируемой группе не менее 60 % студентов освоили все дидактические единицы, входящие в содержание дисциплины. Выбор таких дисциплин, как и сам критериальный балл, чаще всего не обосновывался, что вызывало недоумение у профессионального сообщества.

Во избежание подобного рода ошибок, а также для повышения обоснованности и прозрачности следует предложить объединение нескольких подходов к определению выбора порогового балла. Такой метод, предложенный R. Jaeger, получил название «итеративный структурированный метод оценки заданий», и он предусматривает комбинацию трех составляющих [125, 203]. Первая – это целостная оценка экспертов об измерителе (оценочном инструментарии), вторая – экспертиза качества каждого задания, а третья – анализ результатов выполнения этих заданий.

Состав экспертов, которые участвуют в такой процедуре, желательно расширить до представителей работодателей. В начале каждый эксперт оценивает задание по описанному выше методу «Да / Нет», определяя необходимость правильного ответа на данное задание выпускником. Каждому заданию в категории «Да» присваивается цифра «1», а «Нет» – «0». Общее количество ответов «Да» выступает рекомендуемым критериальным баллом, который определил каждый эксперт. Затем экспертам предъявляются результаты выполнения испытуемыми заданий, а также оценки, которые дали по каждому заданию другие эксперты. Ознакомившись с полученной информацией, каждый эксперт может пересмотреть свои первоначальные оценки, изменив их. Эту итерацию повторяют еще раз. Если группа экспертов состоит из представителей разных сфер деятельности (например, работодатели, преподаватели-предметники, методисты), то рекомендуется эти

итерации проводить для каждой подгруппы экспертов. Тогда значение медианы распределения критериальных баллов, полученное от каждой подгруппы экспертов, можно рассматривать как общий критериальный балл.

Конечно, описанный подход будет требовать большого количества ресурсов (человеческих, временных и, возможно, материальных), поэтому следует исходить из имеющихся возможностей, степени значимости и роли выставляемых оценок на аттестационном испытании.

Для проведения многостадийного оценивания пороговый балл определялся по методу Эбеля для каждого этапа оценивания применительно к каждому субтесту, входящему в измеритель на первом этапе, и ситуационным заданиям на втором этапе. Пороговый балл для первого этапа был установлен на уровне 70 % верно выполненных заданий по каждому субтесту. Пороговый балл для второго этапа многостадийного оценивания был установлен – 60 %, т. е. три из пяти верно выполненных ситуационных задания.

В нашем случае, отображенном в таблице 15, в рамках опытно-экспериментальной работы (2018 г.) и определения пороговых баллов для аттестуемых выпускников по образовательной программе «Психология образования» участвовало три эксперта. Их характеристики следующие.

Эксперт 1. Женщина, 63 года, кандидат педагогических наук, доцент. Стаж работы – более 40 лет, в том числе в системе общего образования классным руководителем, учителем. Высокая степень экспертности подтверждается общественным признанием, подготовкой кадров высшей квалификации (подготовлено и защищено 18 кандидатов наук), председатель и член конкурсов педагогического мастерства областного и муниципальных уровней. Учитель высшей квалификации.

Эксперт 2. Женщина, 60 лет, учитель высшей квалификации, заслуженный учитель. Стаж работы в общем образовании – 38 лет, в том числе учителем, классным руководителем, завучем, директором. Высокая степень экспертности подтверждается общественным признанием, подготовкой призеров олимпиад, победителей различных конкурсов.

Эксперт 3. Мужчина, 58 лет, кандидат педагогических наук, доцент. Стаж работы в системе образования более 35 лет, в т. ч. учителем в школе, классным руководителем, завучем, руководителем районного отдела образования, преподавателем высшей школы. Высокая степень экспертности подтверждается общественным признанием, профессиональными достижениями.

Третий этап многостадийного оценивания происходил на основе экспертной оценки ответа аттестуемого выпускника в соответствии с чек-листом, представленном в следующем параграфе данной главы. Для прохождения порогового балла на этом этапе аттестуемому выпускнику необходимо было соответствовать не менее 6 критериям из 7 установленных, при этом не должно быть ни одного не проявленного критерия. Итоговое решение принималось комиссией коллегиально.

Резюмируя данный раздел, общая последовательность действий организаторов и разработчиков оценочного инструментария по установлению порогового балла для многостадийного оценивания будет следующая.

1. Определить целевую группу аттестуемых, принадлежащих одному направлению подготовки (допустимо включить студентов, осваивавших различные по направленности (профилю) образовательные программы в рамках одного направления подготовки).

2. Подготовить паспорта компетенций, имеющиеся фонды оценочных средств для промежуточной и итоговой аттестации по дисциплинам (модулям), которые являются определяющими для будущей профессиональной деятельности выпускника, спецификацию к оценочному инструментарию для итоговой аттестации.

3. Выбрать один или несколько методов для установления пороговых баллов, проконсультировать группу экспертов, как выполняется определение пороговых баллов (при необходимости подготовить методические материалы и инструкции).

4. Получить от экспертов оценки содержания заданий.

5. После проведения апробации оценочных заданий предоставить возможность взаимодействия и обсуждения экспертов с разработчиками заданий предполагаемого порогового балла.

6. В случае необходимости провести коррекционную работу в отношении оценочных заданий.

7. Организовать экспертное общение для установления пороговых баллов.

8. Согласовать с кругом ответственных и заинтересованных лиц (администрация, профессорско-преподавательский состав, представители работодателей) пороговые баллы выполнения оценочных заданий в составе инструментария для каждого уровня компетентности на каждом этапе оценивания.

9. Документально обобщить процесс установления пороговых баллов, обосновывающий их валидность.

10. После завершения аттестационных испытаний на основе многостадийного оценивания проанализировать полученные результаты (оценки) уровня сформированности выпускников с учетом дополнительной вторичной информацией. При необходимости повторить цикл по установлению порогового балла.

#### **4.3. Классификационные решения об аттестации выпускников на основе многостадийного оценивания**

Строго говоря, этап принятия решений об уровне сформированности компетентности относится в большей степени к управленческой сфере, чем к дидактике. Вместе с тем, последствия принятия таких решений оказывают влияние на все составляющие образовательного процесса, а также на государственную политику в области образования и развития экономики



страны, технологии и содержание подготовки субъектов образовательных отношений, к числу которых в принципе можно отнести выпускников, получающих документ о высшем образовании и квалификацию. В этот момент меняется и статус выпускников, выходящих на рынок труда. В большинстве случаев они становятся молодыми специалистами. Федеральный закон от 30.12.2020 г. № 489-ФЗ «О молодежной политике в Российской Федерации» определяет эту категорию как молодой специалист, являющийся гражданином Российской Федерации в возрасте до 35 лет включительно, завершившим обучение по основным профессиональным образовательным программам и (или) по программам профессионального обучения, впервые устраивающимся на работу в соответствии с полученной квалификацией, в том числе имеющий трудовой стаж, полученный в период обучения по основным профессиональным образовательным программам и (или) по программам профессионального обучения. Тем самым, присваивая квалификацию выпускнику, образовательные организации высшего образования и конкретные структурные подразделения (кафедры, институты, центры, департаменты и пр.) несут ответственность за качество его подготовки как специалиста и за оценки, которые отражают уровень его компетентности. Поэтому оценка становится краеугольным камнем, когда обсуждаются самые различные аспекты, касающиеся образования в целом, и качества подготовки специалистов в частности. И именно поэтому подходы к получению оценок должны быть понятны и прозрачны всем участникам образовательных отношений, им должны доверять и отождествлять со справедливостью. В данном исследовании в реализацию многостадийного оценивания как раз и закладывается эта ключевая идея, которая приводит к обоснованным решениям об уровне компетентности выпускников, завершивших подготовку к выбранным видам профессиональной деятельности и переходящих из сферы образования в сферу труда.

В связи с тем, что принимаемые экзаменационной комиссией решения по результатам прохождения выпускником аттестационного испытания

предусматривают отнесение его к одной из категорий и выставление определенной оценки, то корректнее использовать словосочетание «классификационные решения». Под *классификационным решением* об аттестации выпускника понимается формально оформленное решение о выставлении одной из нормативно установленной оценки выпускнику на основании прохождения им аттестационного испытания и отнесение его к категории с определенным уровнем компетентности. Алгоритм принятия классификационного решения представлен на рисунке 28.

В этом случае речь идет о ситуации, когда итоговая аттестация предусматривает два аттестационных испытания, одним из которых является государственный экзамен и именно он проводится на основе многостадийного оценивания, а второе аттестационное испытание – это защита выпускной квалификационной работы. Поэтому присуждение квалификации выпускнику происходит по итогам успешного прохождения двух аттестационных испытаний. Успешность означает получение одной из трех возможных положительных оценок по традиционной пятибалльной шкале.

В данном исследовании было определено, что прохождение аттестационного испытания в несколько этапов позволяет соотносить уровень сформированности компетентности выпускника с нормативно установленными оценками. Несмотря на то, что информация об уровне компетентности выпускника в настоящий момент ни в какой форме, кроме традиционных оценок, в установленных документах не указывается, эти данные крайне полезны именно с точки зрения принятия дидактических решений. Такие решения являются следствием анализа результатов прохождения аттестуемыми выпускниками аттестационного испытания, позволяющих проводить корректировку и изменения в содержании подготовки специалистов. Это может быть как изменение конкретных дисциплин учебного плана, последовательности их изучения, трудоемкости, так и уточнение формулировок в результатах освоения образовательной программы в целом, совершенствование оценочного инструментария и др.

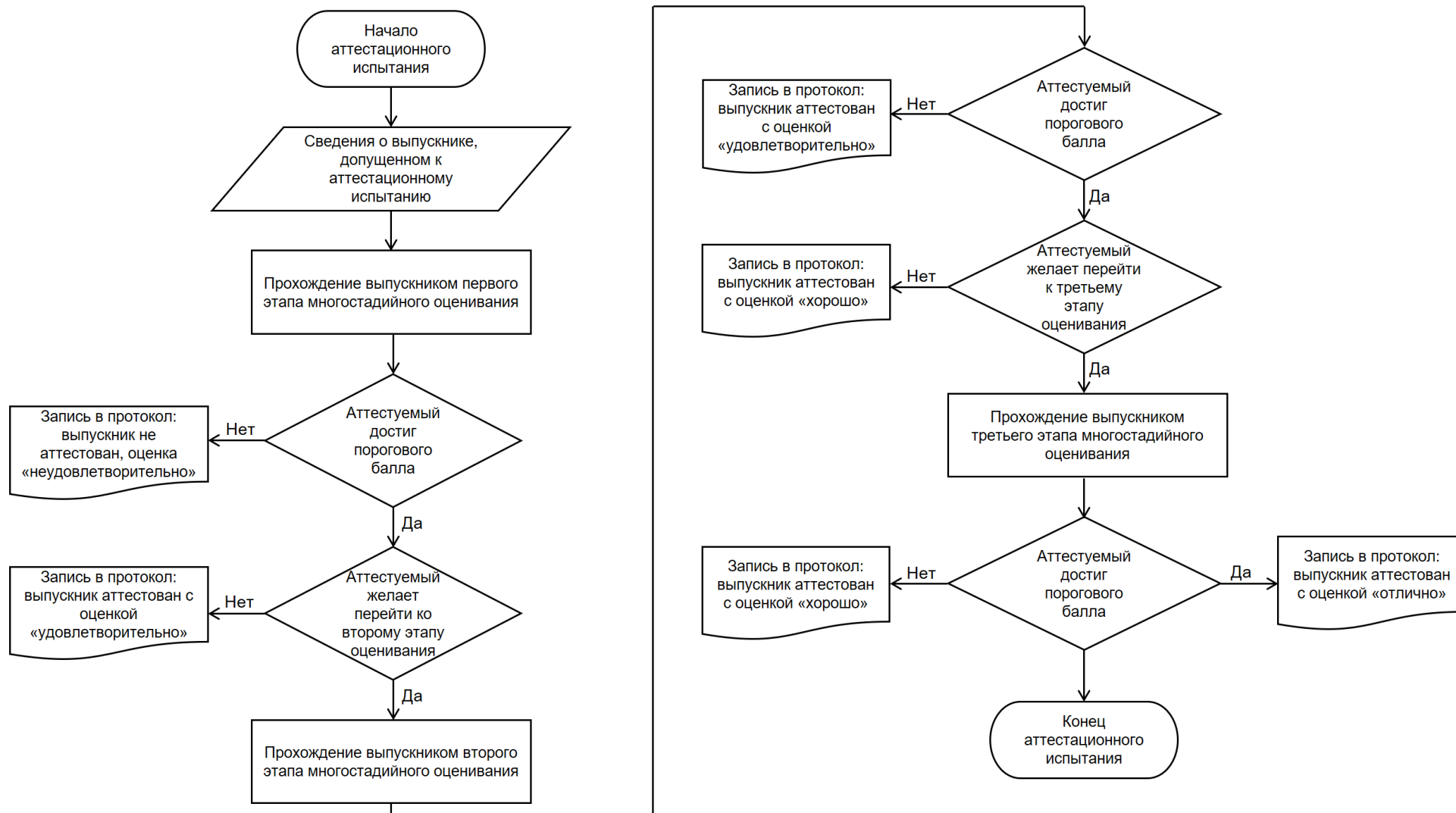


Рис. 28. Алгоритм классификационного решения об аттестации выпускника

Кроме того, указанные выше ссылки на получение обратной связи от работодателей по определенным анкетам (Приложение 11), участвующих в итоговой аттестации, и, следовательно, в принятии классификационных решений, также являются содержательной базой для руководителей образовательных программ и заведующих кафедрами, ответственных за их реализацию. Сводные результаты такого социологического опроса «Оценка удовлетворенности работодателей условиями, организацией и качеством образовательного процесса» по всем образовательным программам, которые реализовывались в 2022-2023 учебном году и по которым проводилась итоговая аттестация выпускников, представлены в Приложении 13.

Далее приведем описание принятия классификационного решения об аттестации выпускников на основе многостадийного оценивания в аттестационной процедуре по программе «Психология образования» направления подготовки 44.03.02 Психолого-педагогическое образование. Для наглядности декомпозиция многостадийного оценивания представлена на рисунке 29.

Как было описано в третьей главе, первые два этапа многостадийного оценивания проводятся в форме компьютерного тестирования. Первый этап, соответствующий минимальному уровню компетентности, включает в себя три субтеста по базовым дисциплинам для указанной образовательной программы:

Субтест 1. Педагогика и педагогическая психология, трудоемкость дисциплины 12 зачетных единиц, 50 заданий.

Субтест 2. Общая и экспериментальная психология, трудоемкость дисциплины 15 зачетных единиц, 100 заданий.

Субтест 3. Психология развития и возрастная психология, трудоемкость дисциплины 10 зачетных единиц, 35 заданий.

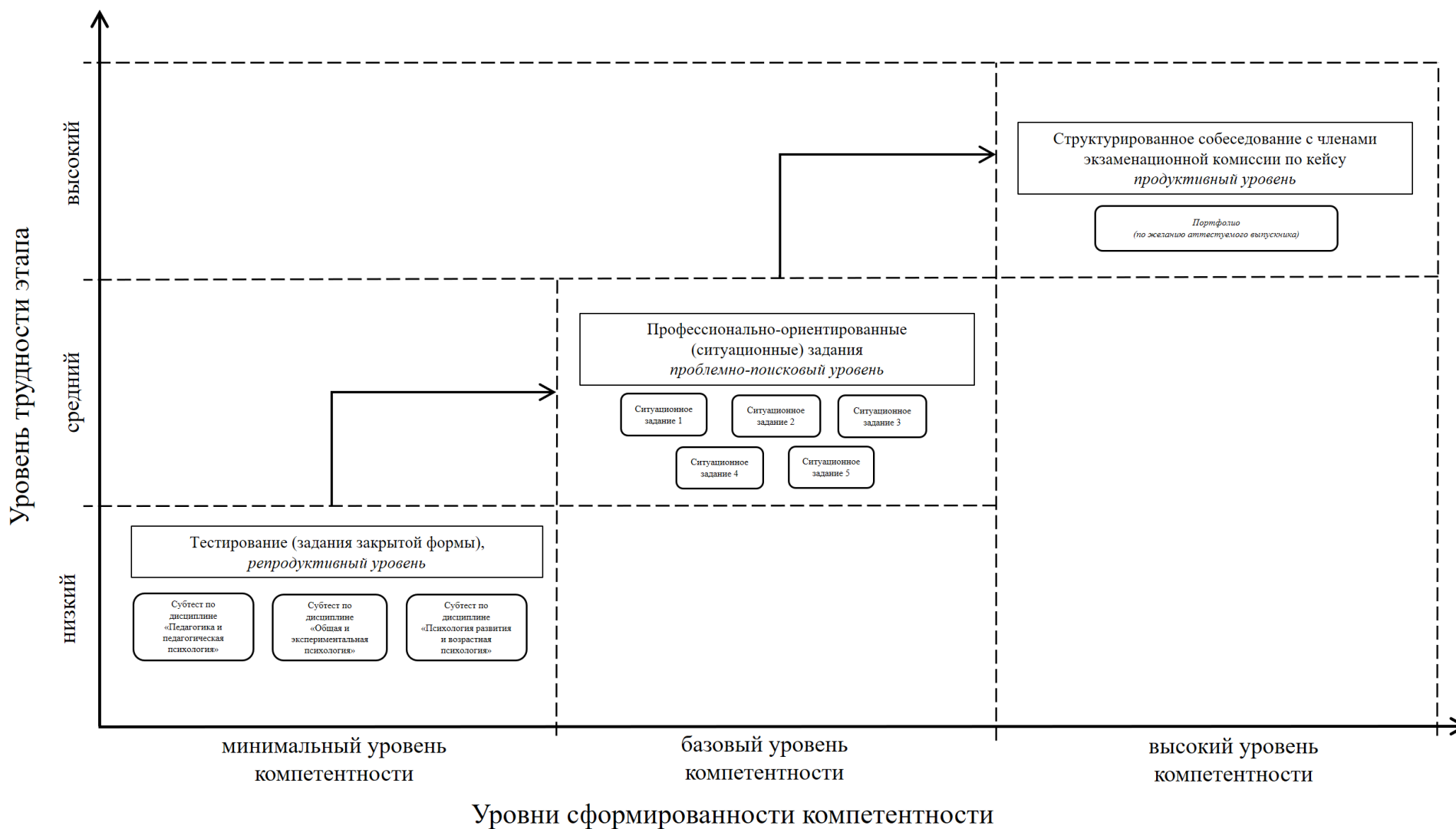


Рис. 29. Многостадийное оценивание по образовательной программе «Психология образования»

Для критериально-ориентированного подхода принято отбирать задания определенной трудности. Трудность каждого задания может быть рассчитана по следующей формуле:

$$\beta_j = \ln \frac{q_j}{p_j}, \quad j=1, 2, \dots, n \quad (9)$$

где  $p_j$  и  $q_j$  – доли правильных и неправильных ответов на  $j$ -е задание.

Теоретически значения параметра трудности могут лежать в интервале  $(-\infty; +\infty)$ , но для проведения измерений на основе критериально-ориентированного подхода оптимальные считаются задания, трудность которых находится в интервале  $[-2; +2]$  логитов. Расчеты вероятностей правильного ответа по однопараметрической модели IRT (модель Раша) для 17 заданий различной трудности при различных значениях уровня подготовленности  $\theta$  вместе с характеристическими кривыми представлены в Приложении 14. На прохождение первого этапа отводится 4 часа (240 минут).

Пороговый балл, установленный итеративным структурированным методом оценки заданий и методом Эбея, составляет 70 % верно выполненных заданий по каждому субтесту. В случае, когда испытуемый не преодолевает пороговый балл, он считается не аттестованным и получает оценку «неудовлетворительно». Это первое классификационное решение, которое может быть принято по отрицательным результатам прохождения тестирования. Вторым классификационным решением на первом этапе оценивания в том случае, если аттестуемый выпускник преодолевает пороговый балл, является аттестация выпускника с оценкой «удовлетворительно», соответствующей минимальному уровню компетентности. Аттестуемый может перейти к следующему второму этапу оценивания для подтверждения своего базового уровня компетентности. В случае преодоления порогового балла на втором этапе, он получает оценку «хорошо», в противном случае за ним сохраняются результаты первого этапа и оценка «удовлетворительно». В рассматриваемом случае на второй стадии каждому испытуемому предлагается выполнить пять ситуационных заданий, имеющих междисциплинарный характер. Это также вариант компьютерного

представления заданий с выбором вариантов ответов. Для прохождения этого этапа аттестуемый должен выполнить верно не менее трех из пяти заданий. Таким образом, третье классификационное решение связано с установлением соответствия выпускника базовому уровню компетентности и выставлении оценки «хорошо» за аттестационное испытание.

Наконец, на третьей стадии оценивания, предусматривающей устное собеседование с членами государственной экзаменационной комиссии по полученному для решения кейсу, будет принято четвертое классификационное решение. Данное классификационное решение уже принимается коллегиально после того, как аттестуемый представит комиссии свое решение проблемы, которая поставлена в задании. На ее выполнение аттестуемому отводится до 20 минут. Для оценивания ответа аттестуемого определен экспертный критериальный лист, представленный в таблице 14, в котором сформулированы критерии и индикаторы их проявления в ответе.

Таблица 14

**Форма экспертного критериального листа для оценивания аттестуемого выпускника на третьем этап (решение кейса)**

№ п/п	Критерий	Индикаторы проявления критерия		
		полностью присутствует	частично присутствует	отсутствует
1.	Формулировка темы в виде проблемного вопроса и обоснование выбора			
2.	Выделены основные понятия, даны им определения, в ответе используется понятийно-терминологический аппарат			
3.	Предлагаемое решение объяснено теоретическими положениями и			

№ п/п	Критерий	Индикаторы проявления критерия		
		полностью присутствует	частично присутствует	отсутствует
	концепциями, указано не менее двух источников			
4.	Даны пояснения методическим идеям, указаны и использованы не менее одного источника			
5.	Приведены обоснованные примеры из практики			
6.	Рассмотрены различные точки зрения, подходы, сформулирована своя точка зрения и позиция. Сделаны выводы, проведена рефлексия			
7.	Отсутствие фактических ошибок			

В результате заполнения формы экспертного критериального листа каждым экспертом делается основной вывод о компетентности или не компетентности выпускника на высоком уровне. Также членам комиссии заранее предоставляется возможность ознакомиться с портфолио выпускников, дать экспертную качественную оценку представленным материалам, которые также учитываются при общих выводах о его компетентности. Пример того, как выглядит электронное портфолио, доступное для просмотра экспертом-экзаменатором, представлен на рисунке 30.

Общее решение принимается простым большинством голосов. В случае разногласий членов комиссии право решающего голоса закреплено за председателем комиссии.



Пороговым баллом на данном этапе выступает не менее 6 из 7 полностью присутствующих в ответе выпускника критериев и не должно быть ни одного индикатора «отсутствует».

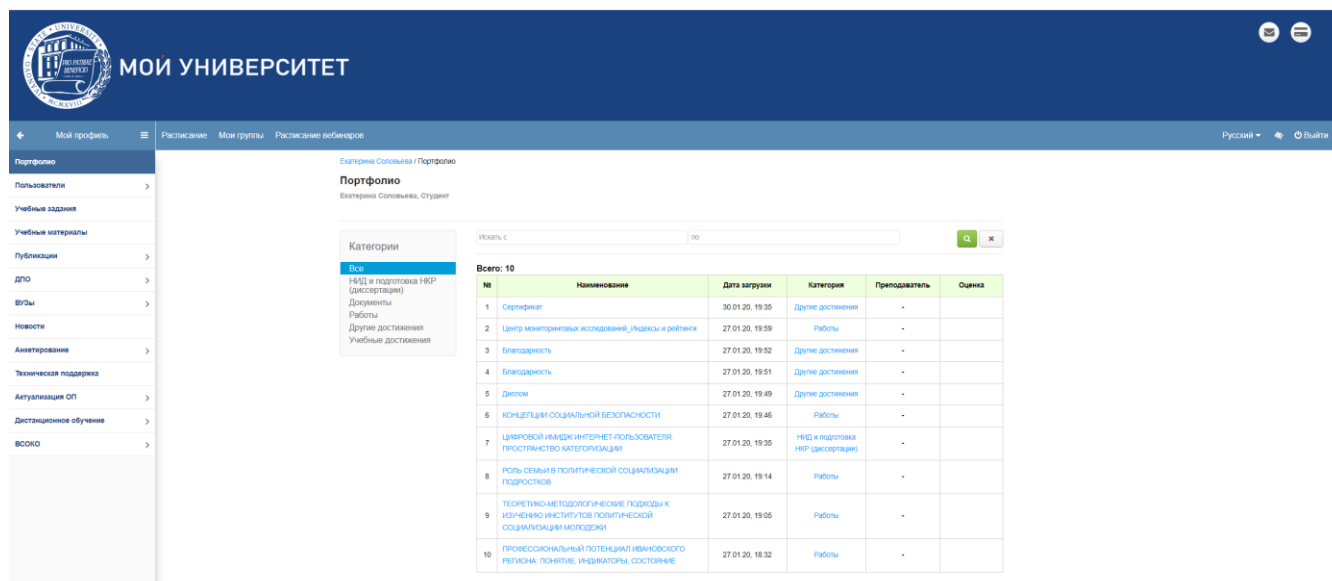


Рис. 30. Фрагмент структуры электронного портфолио студента в ЭИОС

Все классификационные решения оформляются протоколами, которые ведет секретарь комиссии. Они же контролируют порядок и время прохождения этапов оценивания. Примеры протоколов представлены в Приложении 15.

В случае несогласия аттестуемого выпускника с полученной оценкой на аттестационном испытании он имеет право подать апелляцию и тогда начинает работать апелляционная комиссия. Это направление работ в данном исследовании не рассматривается.

Таким образом, приведенная в данном параграфе последовательность прохождения аттестационного испытания (государственного экзамена) на основе многостадийного оценивания и принятия одного из классификационных решений об аттестации выпускника служат поставленным целям итоговой аттестации, обеспечивая всестороннюю оценку уровня профессиональной подготовленности, ее объективность и независимость. Для аттестуемых выпускников такая процедура выступает альтернативой традиционному экзамену, повышающей мотивацию за счет уровневости и нескольких этапов, снижающей вербальное давление

экзаменаторов и способствующей созданию комфортной психологической атмосферы. Это подтверждается и результатами изучения мнения выпускников о самой процедуре итоговой аттестации и определения ответа на вопрос: способны ли существующие процедуры оценивания измерить уровень готовности выпускников вуза к профессиональной деятельности?

В опросе, который проводился для выпускников ИвГУ по программам бакалавриата, принял участие 251 выпускник по образовательным программам различных направлений подготовки – гуманитарных, естественных и технических. Большинство выпускников (64,3 %) удовлетворены существующей процедурой государственной итоговой аттестации. Стоит отметить, что приведенный выше показатель является совокупной характеристикой. Так, респонденты могли вкладывать в него не столько качественные стороны процедуры оценивания (например, объективность, независимость), сколько соблюдение требований к организации ГИА (например, учебно-методические, технические условия) и личное отношение к университету и педагогическому коллективу. Подчеркнем, что в большинстве случаев студенты также отмечают положительный, доброжелательный психологический климат во время прохождения итоговой аттестации. Вместе с тем, больше половины опрошенных (54%) отмечают, что оценивание, осуществляемое в ходе сдачи государственного экзамена, является формальным и не отражает их действительного уровня знаний, умений и уровня сформированности компетенций. При этом более высокий уровень доверия выпускники выражают процедуре защиты ВКР и отмечают, что в 70 % случаев возможно оценить их готовность к профессиональной деятельности. Такая же тенденция наблюдается в оценках респондентами этого аттестационного испытания с позиций объективности – по пятибалльной шкале уровень объективности государственного экзамена оценивается в 3,58, а защиты ВКР – 4,6. Соблюдение принципа независимости при итоговой аттестации выпускники оценивают высоко, отвечая положительно по следующим критериям: все члены комиссии присутствовали на аттестационном испытании – 77,5 %, у каждого члена

комиссии были оценочные листы – 86,7%, возможность предварительно ознакомиться с требованиями, критериями и шкалой оценки – 81,6 %.

В ходе опроса выпускникам предлагалось оценить адекватность существующих аттестационных испытаний. 62 % респондентов соглашались с тем, что в ходе государственного экзамена нельзя в полной мере оценить готовность выпускника к будущей профессиональной деятельности. При этом 70,9 % опрошенных считают, что успешность подготовки и защиты выпускной квалификационной работы свидетельствует о готовности выпускника к будущей профессиональной деятельности. Выпускники также высказались о необходимости совершенствования существующей системы оценивания в рамках итоговой аттестации (безусловно необходимо – 30,6%, скорее необходимо – 31%).

Наиболее оптимальными способами для оценки профессиональных навыков и компетенций выпускники считают написание и защита ВКР – 77,7%, решение профессиональных ситуационных задач – 43,4%, а также решение кейсов – 35,9%. Данный результат подтверждает популярность использования инновационного оценочного инструментария и индивидуальных форматов образовательного оценивания. Эти результаты еще раз подтверждают необходимость совершенствования сложившихся практик итоговой аттестации в университете с обязательным всесторонним обсуждением и разъяснением аспектов и аргументов оценивания со всеми участниками образовательного процесса.

### **Выводы по четвертой главе**

Для практической реализации многостадийного оценивания при итоговой аттестации выпускников необходимо соблюдение ряда организационно-педагогических условий. Эти условия, сформулированные в данной главе, были выявлены и апробированы в ходе практической деятельности автора данной диссертации, доказав свою необходимость.

Выбор критерия оценивания и назначение порогового балла имеет принципиальное значение в условиях итоговой аттестации. Во многом валидность аттестационного решения при многостадийном оценивании будет зависеть не только от содержательной валидности, но и от оптимального выбора такого стандарта оценивания – порогового балла. Выбор такого критерия должен быть определен перед применением оценочного инструментария, до начала интерпретации результатов оценивания и обязательно доведен до аттестуемых. Кроме того, в рассматриваемом в данной работе многостадийном оценивании, состоящем на первых двух этапах из нескольких субтестов, которые выполняет аттестуемый выпускник, необходимо установить и правила выставления оценки за каждую стадию и условие перехода на следующую стадию. Выпускник будет считаться аттестованным на определенной стадии оценивания, если достиг установленного стандарта выполнения (порогового балла). Методы его получения представлены в соответствующем разделе данной главы.

Применение критериально-ориентированного подхода связано с принятием классификационных решений относительно уровня компетентности аттестуемых выпускников и выставления им оценки. На основании подтверждения своих способностей, соответствующих одному из установленных уровней – минимальному, базовому или высокому, выпускники получают одну из нормативно установленных оценок. Следует отметить, что критериально-ориентированный подход в оценивании может использоваться и в целях текущего педагогического контроля, когда становится возможным классифицировать обучающихся по группам для освоения образовательных программ на различных уровнях.

## ЗАКЛЮЧЕНИЕ

Существуют такие сферы деятельности человека, в которых нет права на ошибку. К числу таких сфер наряду с национальной безопасностью, здравоохранением, энергетикой, строительством, транспортной и другими технологическими отраслями относится педагогическая наука и образование. Эта ошибка неразрывно связана с оценкой, выступающей краеугольным камнем во многих вопросах. Если говорить о системе профессионального образования, то сегодня с новой силой обсуждаются вопросы о качестве подготовки специалистов. Это касается и форм обучения, и содержания подготовки, и ее продолжительности, и, конечно, соответствия запросам рынков труда на определенный уровень компетентных специалистов. Ответ на такой запрос система профессионального образования реализует не только через ежегодный прием абитуриентов на приоритетные направления подготовки и выпуск соответствующих специалистов, но также через обеспечение гарантий качества их подготовки.

Конец первой четверти XXI века вновь актуализировал проблему качества, объективности, справедливости и доверия к оценкам и оценочным процедурам в формальном и неформальном образовании. Открывшиеся возможности получения высшего образования, подготовки и переподготовки специалистов для различных отраслей экономики с помощью дистанционных образовательных технологий и электронного обучения с бесконтрольным применением когнитивных помощников, технологий искусственного интеллекта на сегодняшний день вызывают больше настороженности, нежели удовлетворения в вопросах качества образования. В профессиональном образовании усугубляет ситуацию разрыв между ожиданиями работодателей и выпускников, получивших документ о высшем образовании и квалификацию. Для минимизации этого разрыва на уровне государственного регулирования созданы соответствующие условия – введены профессиональные стандарты, учет которых является обязательным при проектировании образовательных результатов в форме профессиональных

компетенций, приняты нормативные акты о независимой оценке квалификаций, в образовательные стандарты высшего образования заложены условия функционирования внешней и внутренней систем оценки качества, в отдельных отраслях развивается система аккредитации специалистов (как например, в здравоохранении).

Вместе с тем, на теоретико-методологическом уровне проводимые исследования проблем качества подготовки специалистов не привели еще к согласованным подходам в оценочных процедурах и организации итоговой аттестации в высшем образовании. Практическая деятельность в области образовательного оценивания, как правило, проводится преподавателями в отрыве от последних достижений теории и основывается, прежде всего, на интуиции и следующей логике: чем лучше студент справляется с трудными заданиями, тем выше качество его подготовки. Введение компетентностного подхода в высшем образовании было призвано изменить предметный взгляд на подготовку специалистов, что, конечно, частично удалось, но слабо сказалось на контрольно-оценочной деятельности. Такой разрыв приводит к негативным последствиям, когда результаты оценивания, полученные студентами и выпускниками по одним и тем же направлениям подготовки, но осваивавших образовательные программы в разных университетах, являются несопоставимыми и необоснованными при одних и тех же нормативно установленных оценках по пятибалльной шкале. Не изменяет эту ситуацию на сегодняшний день и внедрение демонстрационного экзамена в практику итоговой аттестации в высшем образовании в силу различных содержательных и организационных трудностей.

Проведенное исследование призвано решить проблему получения надежных, валидных и аутентичных оценок для обоснованных классификационных решений об уровне сформированности компетентности выпускников при итоговой аттестации за счет обоснования методологии многостадийного оценивания и технологии его реализации в аттестационном испытании высокой значимости (государственный экзамен). Новое понимание результатов освоения образовательных программ не отрицает привычной базовой триады – знаний,

умений, владений, а напротив на их основе и с учетом психофизиологических, личностных, ценностных установок и особенностей, они интегрируются в понятие компетентность, обуславливающее эффективную профессиональную деятельность выпускника. Это обстоятельство необходимо учитывать при ответе на основной вопрос в итоговой аттестации выпускников: каким образом должен быть построен оценочный процесс, чтобы одновременно можно было измерять большое количество заявленных в образовательных программах компетенций, при этом обеспечивая независимость самой процедуры, а также сопоставимость получаемых оценок? Для ответа на этот вопрос в данной работе предложен переход от оценивания устных или письменных ответов аттестуемых выпускников на 2–3 случайно полученных вопроса экзаменационных билетов к многостадийному оцениванию, предусматривающему применение релевантного оценочного инструментария на нескольких этапах аттестационной процедуры.

Применительно к проблеме исследования выделен состав понятийно-терминологического аппарата многостадийного оценивания, определен сам термин «многостадийное оценивание» и уточнены производные от него понятия, связанные с компетентностным подходом, оценочным инструментарием, ошибкой классификации аттестуемого, пороговым баллом, а также дополнены такие категории, как диагностика, образовательные измерения, квалиметрия, уровни сформированности компетентности в контексте целей итоговой аттестации в высшем образовании.

Предложенная авторская методология многостадийного оценивания включает целеполагание, функции, принципы, теоретические основы и подходы, методы и соответствующий им инструментарий для измерения на количественном и качественном уровнях результатов освоения выпускниками образовательной программы. В число целей многостадийного оценивания входят как общие, носящие социально значимый характер для системы высшего образования и совершенствования системы подготовки кадров для экономики страны, так и специальные. Последние связаны с достижением объективности, обоснованности, нацелены на обеспечение повышения эффективности аттестационных процедур,

создание ситуации успеха и благоприятной психологической атмосферы, снижающей невербальное давление экзаменаторов. На многостадийное оценивание возлагаются оценочная, информационная, сравнительная и прогностическая функции. С одной стороны, эти функции традиционно присущи педагогическому контролю, но введение многостадийного оценивания в итоговую аттестацию с новыми возможностями вычислительных мощностей и обработки данных позволяет их реализовывать с большей эффективностью. Принципы как центральная категория в методологическом аппарате многостадийного оценивания классифицируются по трем группам: нормативные (принципы объективности и независимости), общедидактические (принципы научности, релевантности, достоверности, справедливости) и специальные. К числу специальных принципов, повышающих состоятельность многостадийного оценивания, относятся: дифференциация, индивидуализация, последовательность, критериальность, открытость, сочетание количественного и качественного уровней образовательных измерений, кумулятивность.

Основу методологии многостадийного оценивания составляют компетентностный подход к содержанию высшего образования и трактовке результатов освоения образовательных программ, бипарадигмальный подход к образовательным измерениям и теория тестов. Разработка инструментария и его использование в многостадийном оценивании требует опоры на специальные положения, обеспечивающие научное обоснование качества всех компонентов процесса, средств и результатов образовательных измерений. В обоснование качества сложность вносит латентный характер подлежащих измерению конструктов (компетенций). Поэтому интерпретация результатов измерений при итоговой аттестации без общения с выпускником, без предыстории его обучения может оказаться крайне опасной и привести к ложным выводам относительно уровня его компетентности. В этой связи предлагается сочетать количественный и качественный уровни образовательных измерений. Обработка, анализ и интерпретация результатов измерений различного уровня должна проводиться в строгом соответствии с правилами и группами допустимых операций, выделенных



для каждого уровня измерения, с учетом специфики оцениваемых переменных. Невыполнение правил обработки данных может привести к нарушению принципов справедливости и некорректным аттестационным выводам о выпускниках. Положения обеих теорий – и классической, и современной – необходимы для реализации многостадийного оценивания. Классическая теория тестов обеспечивает проведение первичной апробации оценочного инструментария и начального этапа анализа результатов, а современная теория IRT существенным образом продвигает подходы в измерении латентных переменных за счет математического аппарата и предлагаемых моделей (однопараметрической, двух- и трехпараметрических, политомических). Современная теория тестов развивает положения классической теории, позволяя переходить на количественный уровень образовательных измерений. В ее основе лежат вероятностные модели, которые предсказывают успех или неуспех выполнения аттестуемыми с разным уровнем подготовки тех или иных заданий различной трудности.

Предложенная в работе структурная модель многостадийного оценивания обладает свойствами полноты и системности, охватывая практически весь спектр задач, решаемых с помощью образовательных измерений для определения уровня компетентности выпускников при итоговой аттестации. Модель восполняет существующий пробел в дидактике профессионального образования, касающийся теоретико-методологического обоснования контроля и оценки в русле современных тенденций сочетания количественных и качественных методов образовательных измерений, и отвечает требованиям ингерентности (согласованности), простоты (ясности) и адекватности. Данная модель включает все основные структурные компоненты многостадийного оценивания при итоговой аттестации: методологические основания, содержание, технологию и получение выводов для принятия решений.

Важным научным результатом является обоснование оценивания в несколько стадий (этапов) образовательных измерений, последовательно сочетая количественный и качественный уровни. Преимущества нескольких этапов измерений состоят в том, что на каждом из них используются разнообразные

формы заданий, отличающиеся по своим возможностям (задания с выбором ответов, с установлением правильной последовательности или соответствия, ситуационные задания с конструируемыми ответами и так далее). И, наконец, за счет сочетания различных оценочных средств достигается репрезентативный охват требований образовательных и профессиональных стандартов, что приводит к повышению надежности и валидности результатов аттестации.

Реализация полученных теоретических результатов осуществляется с помощью специальной технологии многостадийного оценивания, включающей в себя системную совокупность взаимосвязанных методик – от планирования всего процесса оценивания, проектирования и разработки оценочного инструментария для каждого этапа оценивания, формата реализации (компьютерного или бланкового, письменного или устного) до обеспечения надежности и валидности результатов, которые получаются при его использовании в аттестационном испытании. Такое понимание технологии многостадийного оценивания при итоговой аттестации делает его органической частью всей дидактической системы реализации образовательной программы – от замысла, разработки компетентностной модели специалиста, проектирования результатов и соответствующего содержания через учебный план и рабочие программы дисциплин (модулей) и практик до оценки качества подготовки специалистов.

Разработка инструментария для многостадийного оценивания ведется с другой степенью ответственности, чем это принято для текущего контроля и промежуточной аттестации. Этот процесс начинается с разработки спецификации, которая должна отражать всё то значимое содержание, которое подлежит операционализации и на основе которого будут создаваться оценочные задания. В целом же разработка оценочного инструментария должна проходить определенный цикл, имеющий характер незавершенности в силу необходимости совершенствования и улучшения измерителей.

Само многостадийное оценивание можно рассматривать как производное от адаптивного тестирования, в котором пошагово происходит переоценка трудности и уровня подготовленности. Алгоритм многостадийного оценивания предполагает,

что на первом этапе проводится стандартизированная оценочная процедура с помощью заданий закрытой формы (с выбором ответом, установление правильной последовательности, соответствия или с кратким регламентированным ответом), на втором этапе – решение профессионально-ориентированных (ситуационных) заданий, а на третьем этапе – оценка профессиональной готовности через структурированное собеседование по полученному кейсу с членами экзаменационной комиссии. Переход аттестуемых к каждому следующему этапу возможен только, если их результат превышает установленный пороговый балл (стандарт выполнения этапа) предыдущего этапа. Аутентичность результатов оценивания обеспечивает портфолио, которое выступает подтверждающим дополнением к полученным оценкам уровня сформированности компетентности выпускника и его готовности к осуществлению профессиональной деятельности.

Особое внимание в исследовании уделено вопросам обеспечения надежности и валидности результатов многостадийного оценивания. Надежность как одна из ключевых характеристик образовательных измерений при критериально-ориентированном подходе имеет свои особенности, зависит от ряда факторов (характеристик измерителя, условий проведения оценивания, уровня стандартизации этой процедуры и др.). Для повышения достоверности вычисление показателя надежности следует проводить несколькими способами – с помощью  $r$ -коэффициента корреляции и  $k$ -коэффициента согласованности. Значения этих параметров, близкие к 0,8 и выше, говорят о надежности результатов. Валидность показывает степень адекватности результатов измерения поставленной цели. При критериально-ориентированном подходе в итоговой аттестации на первый план выходит задача обеспечения содержательной валидности. Для этого проводится тщательная качественная экспертиза разработанных оценочных заданий и измерителя в целом с последующей их корректировкой и улучшением.

Для практической реализации многостадийного оценивания при итоговой аттестации выпускников необходимо соблюдение ряда организационно-педагогических условий. Эти условия были выявлены и апробированы автором в ходе многолетней практической деятельности, доказав свою необходимость.

Выбор критерия оценивания и назначение порогового балла имеет принципиальное значение в условиях итоговой аттестации. Во многом валидность аттестационного решения при многостадийном оценивании будет зависеть не только от содержательной валидности, но и от оптимального выбора такого стандарта оценивания – порогового балла. Выбор такого критерия должен быть определен перед применением оценочного инструментария, до начала интерпретации результатов оценивания и обязательно доведен до аттестуемых. Выпускник считается аттестованным на определенной стадии оценивания, если достиг установленного стандарта выполнения или порогового балла. Установление порогового балла основывается на итеративно структурированном методе оценки заданий.

Применение критериально-ориентированного подхода связано с принятием классификационных решений относительно уровня компетентности аттестуемых выпускников и выставления им оценки. На основании подтверждения своих способностей, соответствующих одному из установленных уровней – минимальному, базовому или высокому, выпускники получают одну из нормативно установленных оценок – «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно».

Внедрение компонентов многостадийного оценивания в практику аттестационных процедур высокой значимости осуществлено путем участия автора в апробации технологии, совокупности методик, создании методического обеспечения, приведенного в приложениях. Практическую полезность полученных результатов исследования удалось доказать путем их внедрения в деятельность Ивановского государственного университета, Методического центра аккредитации специалистов в области здравоохранения, Ивановского государственного медицинского университета.

Диссертация имеет практическую ценность как для исследователей, занимающихся теоретическим обоснованием эффективных методов оценивания в условиях развития когнитивных помощников и технологий искусственного интеллекта, увеличения объемов информации, формирования цифровых следов

обучающихся, так и для практиков, поскольку содержит готовые решения для организации аттестационных процедур различной значимости. Материалы данного исследования вносят вклад в совершенствование содержания и условий реализации образовательных программ в высшем образовании, повышение качества подготовки специалистов, могут быть применимы для проведения сравнительных исследований качества образования, функционирования внутренней и внешней систем оценки качества образования, развития аккредитационного мониторинга и независимой оценки квалификаций.

В числе перспективных направлений исследований, развивающих тему контрольно-оценочной деятельности в высшем образовании, повышение эффективности и обоснованности аттестационных процедур высокой значимости можно выделить: аксиологические основания оценивания в современных условиях; формирование единых банков калиброванных оценочных средств для промежуточной и итоговой аттестации в системе высшего образования, способствующих формированию единого образовательного пространства и общего понимания содержания подготовки тех или иных специалистов; развитие аутентичных форм и методов оценивания (таких, как демонстрационный экзамен) с доказательством их эффективности, объективности и сопоставимости получаемых результатов; аспекты, касающиеся автоматизированной проверки свободно конструируемых ответов обучающихся с соблюдением требований к надежности и валидности результатов оценки; а также этические вопросы, связанные со справедливостью и достоверностью оценок в условиях цифровой трансформации как сферы образования, так и общества в целом.

**СПИСОК ЛИТЕРАТУРЫ**

1. Азгальдов, Г. Г. О квалиметрии / Г. Г. Азгальдов, Э. П. Райхман. – Москва : Изд-во стандартов, 1973. – 198 с.
2. Альтбах, Ф. Дж. Глобальные перспективы высшего образования / Ф. Дж. Альтбах ; пер. с англ. Ю. Каптуревского ; под науч. ред. А. Рябова ; предисл. М. Юдкевич ; Нац. исслед. ун-т «Высшая школа экономики». – Москва : ВШЭ, 2018. – 548 с.
3. Амонашвили, Ш. А. Воспитательная и образовательная функция оценки учения школьников : экспериментально-педагогическое исследование / Ш. А. Амонашвили. – Москва : Педагогика, 1984. – 296 с.
4. Ананьев, Б. Г. Психология педагогической оценки / Б. Г. Ананьев // Избранные психологические труды : в 2 т. / Б. Г. Ананьев. – Москва : Педагогика, 1980. – Т. 2. – С. 128–267.
5. Анастаси, А. Психологическое тестирование / А. Анастаси, С. Урбина. – 7-е изд. – Санкт-Петербург : Питер, 2009. – 668 с.
6. Андерсен, П. Разработка тестов и анкет для национальной оценки учебных достижений / П. Андерсен, Д. Морган ; пер. с англ. В. Н. Симкина ; науч. ред. В. И. Звонников. – Москва : Логос, 2011. – 204 с.
7. Андреев, В. И. Проблемы педагогического мониторинга качества образования / В. И. Андреев // Вестник Российской академии образования. – 2001. – № 1. – С. 35–42.
8. Апробация инструментария оценки сформированности профессиональных компетенций у будущих педагогов / А. А. Марголис, М. А. Сафронова, А. С. Панфилова, Л. М. Шишлянникова // Психологическая наука и образование. – 2015. – № 5. – С. 77–91.
9. Артищева, Е. К. Методология изучения педагогической диагностики : моногр. / Е. К. Артищева. – Калининград : Изд-во БФУ им. И. Канта, 2013. – 184 с.

10. Бабанский, Ю. К. Избранные педагогические труды / Ю. К. Бабанский ; сост. М. Ю. Бабанский. – Москва : Педагогика, 1989. – 560 с.
11. Байденко, В. И. Модернизация профессионального образования: современный этап / В. И. Байденко, Дж. ван Зантворт. – Москва : Исслед. центр пробл. качества подгот. специалистов, 2003. – 674 с.
12. Балыхина, Т. М. Словарь терминов и понятий тестологии / Т. М. Балыхина. – Москва : Изд-во РУДН, 2000. – 164 с.
13. Батурин, Н. А. Психология оценки: общие представления, дифференциация понятий и области изучения / Н. А. Батурин // Вестник Южно-Уральского государственного университета. Сер.: Психология. – 2008. – № 31 (131). – С. 17–31.
14. Бахмутский, А. Е. Оценка качества школьного образования : дис. ... д-ра пед. наук / А. Е. Бахмутский. – Санкт-Петербург, 2004. – 343 с.
15. Бебенина, Е. В. Методология применения рейтингов для изучения образовательного пространства : дис. ... д-ра пед. наук / Е. В. Бебенина. – Москва, 2018. – 518 с.
16. Белова, С. Н. Теоретико-методические основания формирования компетентности преподавателей в области внутривузовского оценивания качества образовательного процесса : автореф. дис. ... д-ра пед. наук / С. Н. Белова. – Москва, 2017. – 50 с.
17. Берка, К. Измерения: понятия, теории, проблемы / К. Берка ; пер. с чеш. под ред. В. Б. Бирюкова. – Москва : Прогресс, 1987. – 320 с.
18. Бермус, А. Г. Модернизация образования: философия, политика, культура : моногр. / А. Г. Бермус. – Москва : Канон, 2008. – 384 с.
19. Бермус, А. Г. Управление качеством профессионально-педагогического образования : моногр. / А. Г. Бермус. – Ростов-на-Дону : РГПУ, 2003. – 288 с.
20. Бине, А. Измерение умственных способностей : пер. с фр. / А. Бине. – Санкт-Петербург : Союз, 1998. – 431 с.

21. Блинов, В. И. Профессиональные стандарты: от разработки к применению / В. И. Блинов, О. Ф. Батрова, Е. Ю. Есенина, А. А. Факторович // Высшее образование в России. – 2015. – № 4. – С. 5–14.
22. Блонский, П. П. Педология : учеб. для высш. пед. учеб. заведений / П. П. Блонский. – 2-е изд. – Москва : Гос. учеб.-пед. изд-во, 1936. – 336 с.
23. Болотов, В. А. Прошлое, настоящее и возможное будущее российской системы оценки качества образования / В. А. Болотов // Вопросы образования. – 2018. – № 3. – С. 287–297.
24. Болотов, В. А. Системы оценки качества образования : учеб. пособие / В. А. Болотов, Н. Ф. Ефремова. – Москва : Логос, 2007. – 192 с.
25. Болотов, В. А. Федеральный интернет-экзамен в сфере профессионального образования / В. А. Болотов, В. П. Киселева, В. Г. Наводнов // Высшее образование сегодня. – 2013. – № 12. – С. 2–6.
26. Болотов, В. А. Компетентностная модель: от идеи к образовательной программе / В. А. Болотов, В. В. Сериков // Педагогика. – 2003. – № 10. – С. 8–14.
27. Болотов, В. А. Развитие инструментальных технологий контроля качества образования: стандарты профессионализма и парадоксы роста / В. А. Болотов, А. Г. Шмелев // Высшее образование сегодня. – 2005. – № 4. – С. 16–21.
28. Болотова, Е. Л. Нормативно-правовое регулирование системы оценки достижений учащихся / Е. Л. Болотова // Педагогические измерения. – 2020. – № 2. – С. 88–101.
29. Бордовский, Г. А. Управление качеством образовательного процесса : моногр. / Г. А. Бордовский, А. А. Нестеров, С. Ю. Трапицын. – Санкт-Петербург : Изд-во РГПУ им. А. И. Герцена, 2001. – 359 с.
30. Боуэн, У. Г. Высшее образование в цифровую эпоху / У. Г. Боуэн ; пер. с англ. Д. Кралечкина ; под науч. ред. А. Смирнова ; Нац. исслед. ун-т «Высшая школа экономики». – Москва : Изд. дом ВШЭ, 2018. – 224 с.
31. Будущее высшего образования и академической профессии : страны БРИК и США / под ред. Ф. Альтбаха, Г. Андрущука, Я. Кузьминова, М. Юдкевич,



- Л. Райсберг ; пер. с англ. А. Гордеева, Г. Петренко ; под науч. ред. М. Юдкевич ; Нац. исслед. ун-т «Высшая школа экономики». – Москва : Изд. дом ВШЭ, 2013. – 247 с.
32. Булат, Р. Е. Управление качеством профессиональной подготовки в военно-технических вузах : дис. ... д-ра пед. наук / Р.Е. Булат. – Санкт-Петербург, 2010. – 354 с.
33. Бурлачук, Л. Ф. Психодиагностика : учеб. для вузов / Л. Ф. Бурлачук. – 2-е изд. – Санкт-Петербург : Питер, 2010. – 384 с.
34. Бусыгина, Н. П. Качественная методология и доказательные практики в психологии и образовании / Н. П. Бусыгина, А. В. Горобцова // Психологическая наука и образование. – 2021. – Т. 26, № 6. – С. 117–127.
35. Бусыгина, Н. П. Доказательный подход в образовании: критический анализ актуальных дискуссий / Н. П. Бусыгина, Т. Г. Подушкина, В. В. Станилевский // Психолого-педагогические исследования. – 2021. – Т. 13, № 4. – С. 162–176.
36. Вайнштейн, Ю. В. Подход к разработке адаптивной системы электронного обучения / Ю. В. Вайнштейн, Т. О. Кочеткова, В. А. Шершнева // Информатизация образования и методика электронного обучения : материалы II Междунар. науч. конф., Красноярск, 25–28 сентября 2018 г. : в 2 ч. / под. общ. ред. М. В. Носкова. – Красноярск : Сиб. федер. ун-т, 2018. – Ч. 1. – С. 124–128.
37. Вербицкий, А. А. Личностный и компетентностный подходы в образовании: проблемы интеграции / А. А. Вербицкий, О. Г. Ларионова. – Москва : Логос, 2010. – 328 с.
38. Виландеберк, А. А. Новые технологии оценки результатов обучения : (уровневое образование) : метод. пособие для преподавателей / А. А. Виландеберк, Н. Л. Шубина. – Санкт-Петербург : Изд-во РГПУ им. А. И. Герцена, 2008. – 168 с.
39. Воронова, Т. А. Актуальные направления в оценивании готовности выпускников к профессиональной педагогической деятельности /

- Т. А. Воронова, А. А. Малыгин // Высшее образование сегодня. – 2019. – № 12. – С. 10–16.
40. Воронова, Т. А. Изменения в контрольно-оценочной деятельности учителей и учащихся в соответствии с требованиями ФГОС ООО / Т. А. Воронова, А. А. Малыгин // Тенденции развития образования: кто и как использует и оценивает образовательные стандарты : материалы XIV ежегод. Междунар. науч.-практ. конф., Москва, 16–18 февраля 2017 г. – Москва : Дело, 2018. – С. 163–176.
41. Воронова, Т. А. Контрольно-оценочная система в вузе: традиции и требования современности / Т. А. Воронова, А. А. Малыгин // Научный поиск. – 2015. – № 3.1. – С. 32–34.
42. Воронова, Т. А. Независимая оценка качества условий образовательной деятельности: нормы и практики / Т. А. Воронова, Е. М. Дорожкин, А. А. Малыгин // Педагогический журнал Башкортостана. – 2018. – № 2 (75). – С. 49–56.
43. Воронова, Т. А. Первокурсники, выпускники и работодатели как акторы системы оценки качества высшего образования: интегративный мониторинг / Т. А. Воронова, А. А. Малыгин, И. Н. Смирнова // Тенденции развития образования. Эффективность образовательных институтов : материалы XVI ежегод. Междунар. науч.-практ. конф., Москва, 14–16 февраля 2019 г. – Москва : Дело, 2020. – С. 84–95.
44. Вроейнстийн, А. И. Оценка качества высшего образования : рекомендации по внешней оценке качества в вузах / А. И. Вроейнстийн. – Москва : Изд-во МНЭПУ, 2000. – 180 с.
45. Выготский, Л. С. Педагогическая психология / Л. С. Выготский. – Москва : Педагогика, 1991. – 412 с.
46. Галустьян, О. В. Система полифункционального контроля профессиональной подготовки компетентного специалиста в высшей школе : дис. ... д-ра пед. наук / О. В. Галустьян. – Воронеж, 2016. – 432 с.

47. Гласс, Дж. Статистические методы в педагогике и психологии / Дж. Гласс, Дж. Стенли. – Москва : Прогресс, 1976. – 495 с.
48. Глобализация и системы обеспечения качества высшего образования / С. А. Запрягаев, Е. В. Караваева, И. Г. Карелина, А. М. Салецкий. – Москва : Изд-во МГУ, 2007. – 292 с.
49. Горбачев, В. Т. Повышение эффективности адаптивного тестирования качества обучения студентов в вузах гуманитарного профиля : дис. ... канд. пед. наук / В. Т. Горбачев. – Москва, 2006. – 258 с.
50. Граничина, О. А. Контроль качества образовательного процесса в контексте управления вузом : автореф. дис. ... д-ра пед. наук / О. А. Граничина. – Санкт-Петербург, 2009. – 42 с.
51. Гриншкун, В. В. Иерархическая структура алгоритмов построения индивидуальных образовательных траекторий / В. В. Гриншкун, А. А. Заславский // Вестник МГПУ. Серия: Информатика и информатизация образования. – 2021. – № 4(58). – С. 15–20.
52. Гриншкун, В. В. Искусственный интеллект в образовательной деятельности и подготовке педагогов: необходимость исследований / В. В. Гриншкун, Л. А. Шунина // Современная {цифровая} дидактика. – Москва : ООО «А-Приор», 2023. – С. 49–55.
53. Гуськова, М. В. Основы эвалюации в совершенствовании управления образованием : моногр. / М. В. Гуськова. – Москва : ИНФРА-М, 2014. – 214 с.
54. Гуськова, М. В. Эвалюация в образовании : моногр. / М. В. Гуськова. – Москва : ИНФРА-М, 2012. – 153 с.
55. Гуськова, М. В. Эвалюация в управлении качеством образования в условиях его модернизации : дис. ... д-ра пед. наук / М. В. Гуськова. – Великий Новгород, 2014. – 438 с.
56. Демидова, М. Ю. Государственная итоговая аттестация учащихся по физике: эволюция экзаменационных материалов / М. Ю. Демидова // Педагогические измерения. – 2018. – № 2. – С. 57–66.

57. Демидова, М. Ю. Методическая система оценки учебных достижений учащихся по физике в условиях введения ФГОС (общее образование) : дис. ... д-ра пед. наук / М. Ю. Демидова. – Москва, 2014. – 438 с.
58. Демидова, М. Ю. Современные подходы к оценке качества естественнонаучного образования в международных и национальных исследованиях / М. Ю. Демидова // Естественнонаучное образование: проблемы оценки качества : сборник. Сер. «Методический ежегодник Химического факультета МГУ имени М. В. Ломоносова». Москва : Московский государственный университет имени М. В. Ломоносова (Химический факультет). Т 14, 2018. – С. 14–41.
59. Дорожкин, Е. М. Оценивание компетенций: проблемы, подходы, решения / Е. М. Дорожкин, А. А. Малыгин // Вестник Учебно-методического объединения по профессионально-педагогическому образованию. – 2013. – № 1 (47). – С. 12–17.
60. Дроботенко, Ю. Б. Структура и способы оценивания результатов высшего педагогического образования / Ю. Б. Дроботенко, Н. С. Макарова, Н. В. Чекалева // Ярославский педагогический вестник. – 2018. – № 5. – С. 174–182.
61. Дружинин, В. Н. Психология общих способностей / В. Н. Дружинин. – 3-е издание. – Санкт-Петербург : Питер, 2007. – 368 с.
62. Дятлова, К. Д. Система разработки и анализа педагогических тестов по биологии : автореф. дис. ... д-ра пед. наук / К. Д. Дятлова. – Санкт-Петербург, 2007. – 40 с.
63. Дятлова, К. Д. Опыт создания компетентностного итогового теста / К. Д. Дятлова, М. А. Гаврилова, И. А. Колпаков // Вестник Нижегородского университета им. Н. И. Лобачевского. – 2013. – № 2 (1). – С. 11–15.
64. Единый государственный экзамен. Научные основы, методология и практика организации эксперимента : сб. ст. / под ред. В. А. Болотова. – Москва : Логос, 2002. – 208 с.

65. Елина, Е. Г. Компетенции и результаты обучения: логика представления в образовательных программах / Е. Г. Елина, Е. Н. Ковтун, С. Е. Родионова // Высшее образование в России. – 2015. – № 1. – С. 10–20.
66. Елина, Е. Г. Российское высшее образование в условиях профессиональной стандартизации: опыт, вызовы, риски / Е. Г. Елина, Е. Н. Ковтун, С. Е. Родионова // Высшее образование в России. – 2019. – Т. 28, № 6. – С. 9–27.
67. Елисеев, И. Н. Математические модели и комплексы программ для автоматизированной оценки результатов обучения с использованием латентных переменных : дис. ... д-ра техн. наук / И. Н. Елисеев. – Новочеркасск, 2013. – 371 с.
68. Ерофеева, В. В. О подходах к совершенствованию компетентности выпускников творческих специальностей в условиях информатизации образования / В. В. Ерофеева, В. И. Звонников, А. А. Малыгин // Новые технологии оценки качества образования : сб. материалов XVI Форума Гильдии экспертов в сфере профессионального образования в рамках онлайн-конференций. – Москва, 2021. – С. 25–30.
69. Есенина, Е. Ю. Понятийная основа современного профессионального образования : моногр. / Е. Ю. Есенина. – Москва : Перо, 2015. – 191 с.
70. Ефимова, С. А. Модернизация системы аттестации обучающихся профессиональных образовательных организаций в условиях становления российской национальной системы квалификаций : дис. ... д-ра пед. наук / С. А. Ефимова. – Москва, 2017. – 337 с.
71. Ефремова, Н. Ф. Аргументации и доказательства надежности оценок компетенций студентов / Н. Ф. Ефремова // Известия Дагестанского государственного педагогического университета. Психолого-педагогические науки. – 2018. – Т. 12, № 2. – С. 43–50.
72. Ефремова, Н. Ф. Компетенции в образовании : формирование и оценивание / Н. Ф. Ефремова. – Москва : Нац. образование, 2012. – 416 с.

73. Ефремова, Н. Ф. Подходы к оцениванию компетенций в высшем образовании : учеб. пособие / Н. Ф. Ефремова. – Москва : Исслед. центр пробл. качества подгот. специалистов, 2010. – 216 с.
74. Ефремова, Н. Ф. Приемы доказательной аргументации оценки компетенций / Н. Ф. Ефремова // Инновационная наука: психология, педагогика, дефектология. – 2020. – Т. 3, № 2. – С. 112–124.
75. Ефремова, Н. Ф. Стратегии обеспечения справедливости в дизайне оценивания компетенций студентов / Н. Ф. Ефремова // Инновационная наука: психология, педагогика, дефектология. – 2023. – Т. 6, № 1. – С. 81–93.
76. Ефремова, Н. Ф. Тестовый контроль качества учебных достижений в образовании : дис. ... д-ра пед. наук / Н. Ф. Ефремова. – Ростов-на-Дону, 2003. – 458 с.
77. Ефремова, Н. Ф. Педагогические измерения в системе образования / Н. Ф. Ефремова, В. И. Звонников, М. Б. Челышкова // Педагогика. – 2006. – № 2. – С. 14–22.
78. Загвоздкин, В. К. Теория и практика применения стандартов в образовании / В. К. Загвоздкин. – Москва : Нар. образование : НИИ школьных технологий, 2011. – 344 с.
79. Загвязинский, В. И. Методология и методы дидактического исследования / В. И. Загвязинский. – Москва : Педагогика, 1982. – 158 с.
80. Звонников, В. И. Адаптивное тестирование в дистанционном обучении / В. И. Звонников, М. Б. Челышкова, А. А. Малыгин // Высшее образование сегодня. – 2012. – № 6. – С. 7–10.
81. Звонников, В. И. Адаптивное тестирование как эффективная форма контроля при дистанционном обучении / В. И. Звонников, М. Б. Челышкова, А. А. Малыгин // Вестник университета. – 2008. – № 2 (8). – С. 69–71.
82. Звонников, В. И. Адаптивное тестирование: вчера, сегодня, завтра / В. И. Звонников, М. Б. Челышкова // Высшее образование сегодня. – 2010. – № 8. – С. 14–17.

83. Звонников, В. И. Измерения и качество образования / В. И. Звонников. – Москва : Логос, 2006. – 312 с.
84. Звонников, В. И. Как поднять эффективность аккредитации профессиональных образовательных программ и снизить ее нагрузку на вузы? / В. И. Звонников // Высшее образование сегодня. – 2019. – № 3. – С. 11–16.
85. Звонников, В. И. О доказательном подходе и его видах в образовании / В. И. Звонников, А. А. Малыгин, М. Б. Челышкова // Вестник Ивановского государственного университета. Сер.: Естественные, общественные науки. – 2021. – Вып. 2. – С. 46–52.
86. Звонников, В. И. Оценивание в высшем образовании: от линейности к адаптивности / В. И. Звонников, А. А. Малыгин, М. Б. Челышкова // Известия высших учебных заведений. Сер.: Гуманитарные науки. – 2014. – Т. 5, вып. 2. – С. 166–171.
87. Звонников, В. И. Оценка качества результатов обучения при аттестации : (компетентностный подход) : учеб. пособие / В. И. Звонников, М. Б. Челышкова. – 2-е изд., перераб., доп. – Москва : Логос, 2012. – 280 с.
88. Звонников, В. И. Педагогические измерения в управлении качеством обучения : дис. ... д-ра пед. наук / В. И. Звонников. – Ростов-на-Дону, 2006. – 405 с.
89. Звонников, В. И. Принципы педагогических измерений результатов обучения / В. И. Звонников, Н. Н. Кузьмина, А. А. Малыгин // Социальная защита детей и молодежи в условиях глобализации: образование, экономика, право : сб. ст. в рамках реализации проекта программы Европейской комиссии Темпус IV «Создание транснациональных программ для подготовки магистров и аспирантов в области социального обеспечения детей и молодежи» (TACHUwe). – 2-е изд., перераб. и доп. – Ростов-на-Дону : ИП С. В. Беспмятнов, 2017. – С. 20–25.
90. Звонников, В. И. Современные подходы к оцениванию качества результатов высшего образования / В. И. Звонников, М. Б. Челышкова // Педагогические измерения. – 2016. – № 1. – С. 32–38.

91. Зимняя, И. А. Ключевые компетентности как результативно-целевая основа компетентностного подхода в образовании / И. А. Зимняя. – Москва : Исслед. центр пробл. качества подгот. специалистов, 2004. – 40 с.
92. Зимняя, И. А. Компетенция и компетентность в контексте компетентностного подхода в образовании / И. А. Зимняя // Ученые записки национального общества прикладной лингвистики. – 2013. – № 4 (4). – С. 16–31.
93. Иванов, Б. С. Теория и методология реализации педагогической системы качества образовательного процесса в вузе : дис. ... д-ра пед. наук / Б. И. Иванов. – СПб., 2004. – 454 с.
94. Иванов, Д. А. Компетентностный подход в образовании : проблемы, понятия, инструментарий : учеб.-метод. пособие / Д. И. Иванов, К. Г. Митрофанов, О. В. Соколова. – Москва : АПК и ПРО, 2003. – 101 с.
95. Иванова, О. Н. Организация государственных экзаменов для выпускников вузов в условиях пандемии COVID-19 / О. Н. Иванова // Вестник педагогических наук. – 2021. – № 3. – С. 54–58.
96. Иванова, С. В. Образование в организационно-гуманистическом измерении : моногр. / С. В. Иванова. – Москва : РУДН, 2007. – 236 с.
97. Иванова, С. В. Перспективы развития образования в условиях четвертой промышленной революции / С. В. Иванова, О. Б. Иванов // ЭТАП: экономическая теория, анализ, практика. – 2019. – № 6. – С. 7–30.
98. Иванова, С. В. Постмодернизм и качество образования : (постановка проблемы) / С. В. Иванова, И. М. Елкина // Ценности и смыслы. – 2016. – № 6 (46). – С. 115–124.
99. Иванова, С. В. Система оценивания результатов обучения при различных педагогических подходах в России и зарубежных странах / С. В. Иванова, И. М. Елкина // Вестник Академии педагогических наук Казахстана. – 2014. – № 3. – С. 19–31.
100. Иванова, С. В. Устойчивое развитие в неустойчивом мире: образовательный аспект / С. В. Иванова, О. Б. Иванов // Ценности и смыслы. – 2021. – № 3 (73). – С. 6–26.



101. Изменения в профессиональной деятельности педагога в контексте новой трудовой реальности : коллектив. моногр. / Н. В. Чекалева, А. П. Тряпицына, С. А. Писарева, Ю. Б. Дроботенко, Н. А. Дука, Е. И. Зарипова, Н. С. Макарова, Г. П. Сеницына, Т. О. Соловьева, И. В. Феттер, Е. А. Черкевич, Н. И. Чуркина ; под общ. ред. Н. В. Чекалевой. – Омск : Литера, 2015. – 228 с.
102. Измерение и оценка сформированности универсальных компетенций обучающихся при освоении образовательных программ бакалавриата, магистратуры, специалитета : коллектив. моногр. / под науч. ред. И. Ю. Тархановой. – Ярославль : РИО ЯГПУ, 2018. – 383 с.
103. Ингенкамп, К. Педагогическая диагностика : пер. с нем. / К. Ингенкамп. – Москва : Педагогика, 1991. – 240 с.
104. Инструментарий для оценки состояния образовательных систем и рекомендации по его использованию / В. А. Прудникова, С. Ю. Алашеев, Е. А. Карпухина, Д. Л. Константиновский, Н. Ю. Посталюк, В. Ф. Солдатов, Н. В. Тюрина, А. В. Фирсова, М. А. Шермет ; под общ. ред. В. А. Прудниковой. – Москва : Унив. кн., 2006. – 308 с.
105. Интерактивные множественные кейсы в аккредитации специалистов здравоохранения / Т. В. Семенова, Ж. М. Сизова, Т. Н. Малахова, А. А. Малыгин, М. Б. Чельшкова, С. А. Князева // Медицинский вестник Северного Кавказа. – 2019. – № 1. – С. 118–122.
106. Итин, Ю. К. Оптимизация управления качеством образовательного процесса в высшей школе : дис. ... д-ра пед. наук / Ю. К. Итин. – Москва, 2004. – 295 с.
107. Ишкова, Л. В. Теория и практика педагогической оценки в образовании старшеклассников и взрослых : моногр. / Л. В. Ишкова. – Санкт-Петербург : Изд-во РГПУ им. А. И. Герцена, 2000. – 187 с.
108. Кадневский, В. М. Генезис тестирования в истории отечественного образования : дис. ... д-ра пед. наук / В. М. Кадневский. – Омск, 2007. – 369 с.
109. Кадневский, В. М. История тестов / В. М. Кадневский. – Москва : Нар. образование, 2004. – 464 с.

110. Кадневский, В. М. Тесты в системе профессионального отбора и профессионального образования: история и современность / В. М. Кадневский, О. В. Могиль, Т. В. Ширшова // Вестник Омского университета. – 2014. – № 2. – С. 185–190.
111. Кадневский, В. М. Фрэнсис Гальтон: учёный-энциклопедист, один из первых создателей теории педагогических измерений / В. М. Кадневский, В. Лемиш, Т. В. Ширшова // Педагогические измерения. – 2012. – № 1. – С. 3–16.
112. Казакова, Е. И. Оценка универсальных компетенций студентов при освоении образовательных программ / Е. И. Казакова, И. Ю. Тарханова // Ярославский педагогический вестник. – 2018. – № 5. – С. 127–135.
113. Караваева, Е. В. Квалификации высшего образования и профессиональные квалификации: «сопряжение с напряжением» / Е. В. Караваева // Высшее образование в России. – 2017. – № 12 (218). – С. 5–12.
114. Карданова, Е. Ю. Моделирование и параметризация тестов: основы теории и приложения / Е. Ю. Карданова. – Москва : Федер. центр тестирования, 2008. – 292 с.
115. Качество образования в российских университетах: что мы поняли в пандемию : аналит. докл. / науч. ред. Е. А. Суханова, И. Д. Фруммин. – Томск : Изд-во Том. гос. ун-та, 2021. – 46 с.
116. Кейсы для оценки сформированности универсальных компетенций обучающихся при освоении образовательных программ высшего образования (программ бакалавриата, специалитета, магистратуры) : метод. рекомендации / под общ. ред. И. Ю. Тархановой. – Ярославль : РИО ЯГПУ, 2018. – 218 с.
117. Кеннеди, К. Как использовать анализ данных о добавленной стоимости для улучшения обучения школьников : руководство для школ и лидеров школьных округов / К. Кеннеди, М. Питерс, М. Томас ; пер. с англ. О. Лобачевой ; под науч. ред. Л. Любимова ; Нац. исслед. ун-т «Высшая школа экономики». – Москва : Изд. дом ВШЭ, 2017. – 264 с.

118. Клайн, П. Справочное руководство по конструированию тестов : введение в психометрическое проектирование / П. Клайн. – Киев : Ника-Центр, 1994. – 284 с.
119. Концепция оценивания квалификаций / В. И. Блинов, О. Ф. Батрова, Е. Ю. Есенина, А. А. Факторович // Образование и наука. – 2012. – № 10. – С. 46–66.
120. Концепция уровневой оценки компетенций учителя / Н. А. Бочарова, С. А. Писарева, М. Ю. Пучков, В. И. Снегурова, А. П. Тряпицына // Человек и образование. – 2017. – № 3 (52). – С. 164–171.
121. Краевский, В. В. Методология педагогики: новый этап : учеб. пособие для студентов высш. учеб. заведений / В. В. Краевский, Е. В. Бережнова. – Москва : Академия, 2006. – 400 с.
122. Краевский, В. В. Основы обучения : дидактика и методика : учеб. пособие для студентов высш. учеб. заведений / В. В. Краевский, А. В. Хуторской. – Москва : Академия, 2007. – 352 с.
123. Крамер, Д. Математическая обработка данных в социальных науках : современные методы : учеб. пособие для студентов высш. учеб. заведений / Д. Крамер ; пер. с англ. И. В. Тимофеева, Я. И. Киселевой ; науч. ред. О. В. Митина. – Москва : Академия, 2007. – 288 с.
124. Красильникова, В. А. Теория и технологии компьютерного обучения и тестирования : моногр. / В. А. Красильникова. – Москва : Дом педагогики, 2009. – 339 с.
125. Крокер, Л. Введение в классическую и современную теорию тестов : учеб. / Л. Крокер, Дж. Алгина ; пер. с англ. Н. Н. Найденовой, В. Н. Симкина, М. Б. Чельшковой ; под общ. ред. В. И. Звонникова, М. Б. Чельшковой. – Москва : Логос, 2010. – 668 с.
126. Кузнецова, М. И. Система контроля и оценки образовательных достижений младших школьников как фактор повышения качества образования : дис. ... д-ра пед. наук / М. И. Кузнецова. – Москва, 2017. – 488 с.

127. Кузьминов, Я. И. Университеты России : как это работает / Я. И. Кузьминов, М. М. Юдкевич. – Москва : Изд. дом ВШЭ, 2021. – 616 с.
128. Кузьмина, Н. В. Акмеологическая теория качества подготовки специалистов образования / Н. В. Кузьмина. – Москва : Исслед. центр пробл. качества подгот. специалистов, 2001. – 104 с.
129. Куравский, Л. С. Новый подход к компьютеризированному адаптивному тестированию / Л. С. Куравский, С. Л. Артеменков, Г. А. Юрьев, Е. Л. Григоренко // Экспериментальная психология. – 2017. – Т. 10, №. 3. – С. 33–45.
130. Левина, Е. Ю. Система управления развитием высшего образования на основе информационно-когнитивного подхода : дис. ... д-ра пед. наук / Е. Ю. Левина. – Казань, 2018. – 414 с.
131. Левицкий, М. Л. Влияние интеграции информационных технологий на цифровую трансформацию образования и формирование ценностных ориентаций обучающихся / М. Л. Левицкий, В. В. Гриншкун, О. Ю. Заславская // Ценностные основы развития российского образования: теория и практика : 80-летию Российской академии образования посвящается: монография / Российская академия образования. – Москва : ООО «МАКС Пресс», 2023. – С. 62–71.
132. Левицкий, М. Л. Многокомпонентная модель унификации и интеграции цифровых ресурсов вуза / М. Л. Левицкий, В. В. Гриншкун, О. Ю. Заславская // Вестник МГПУ. Серия: Информатика и информатизация образования. – 2023. – № 1 (63). – С. 7–23.
133. Левицкий, М. Л. Тенденции и особенности современного этапа информатизации высшей школы / М. Л. Левицкий, В. В. Гриншкун, О. Ю. Заславская // Вестник Российского университета дружбы народов. Серия: Информатизация образования. – 2022. – Т. 19, № 4. – С. 285–299.
134. Ледовская, Т. В. Основные подходы к оценке результатов освоения студентами основных образовательных программ / Т. В. Ледовская,

- Н. Э. Солянин // Ярославский педагогический вестник. – 2019. – № 1 (106). – С. 49–55.
135. Логвиненко, А. Д. Измерения в психологии: математические основы : учеб. пособие / А. Д. Логвиненко. – Москва : Изд-во МГУ, 1993. – 480 с.
136. Ломакина, Т. Ю. Современные тенденции развития системы управления качеством профессионального образования / Т. Ю. Ломакина // Отечественная и зарубежная педагогика. – 2011. – № 1. – С. 165–173.
137. Майоров, А. Н. Теория и практика создания тестов для системы образования : как выбирать, создавать и использовать тесты для целей образования / А. Н. Майоров. – Москва : Нар. образование, 2000. – 352 с.
138. Малахова, Т. Н. Проектирование инструментария для аккредитации выпускников медицинских вузов : дис. ... канд. пед. наук / Т. Н. Малахова. – Москва, 2021. – 226 с.
139. Малыгин, А. А. Адаптивное оценивание в аттестационных процедурах студентов и выпускников / А. А. Малыгин // Alma Mater : вестник высшей школы. – 2023. – № 8. – С. 39–49.
140. Малыгин, А. А. Адаптивное оценивание в аттестационных процедурах студентов современного вуза / А. А. Малыгин // Новые дидактические решения для развития высшего образования в условиях цифровой трансформации : сб. науч.-метод. материалов / под науч. ред. М. Л. Левицкого [и др.]. – Ярославль : РИО ЯГПУ, 2023. – С. 73–92.
141. Малыгин, А. А. Адаптивное тестирование в дистанционном обучении : моногр. / А. А. Малыгин. – Иваново : Иван. гос. хим.-технол. ун-т, 2012. – 138 с.
142. Малыгин, А. А. Адаптивное тестирование учебных достижений студентов в дистанционном обучении : дис. ... канд. пед. наук / А. А. Малыгин. – Москва, 2011. – 183 с.
143. Малыгин, А. А. Адаптивные измерения в образовательном оценивании / А. А. Малыгин // Теория и методика профессионального образования : сб. науч. ст. – Вып. 4 : в 2 ч. – Минск : РИПО, 2017. – Ч. 1. – С. 216–231.

144. Малыгин, А. А. Аттестация студентов в современных условиях: от повышения эффективности оценивания к доверию оценкам / А. А. Малыгин // Высшее образование сегодня. – 2023. – № 6. – С. 8–16.
145. Малыгин, А. А. Компьютерное адаптивное тестирование в обеспечении качества дистанционного обучения / А. А. Малыгин // Вестник университета. – 2010. – № 4. – С. 166–169.
146. Малыгин, А. А. Контрольно-оценочная деятельность педагога в высшем образовании : нормативные требования / А. А. Малыгин, Н. Ю. Сорокин // Проблемы высшего образования. – 2017. – Т. 1. – С. 59–62.
147. Малыгин, А. А. Метод критических ситуаций в системе оценивания компетенций / А. А. Малыгин, Е. А. Соловьева // Качество педагогического образования в условиях современных вызовов [Электронный ресурс]: Сб. науч. тр. IX Межд. форума по педагогическому образованию (Казань, 24–26 мая 2023 г.). – Электронные текстовые данные (1 файл: 7,07 Мб). – Казань: – Казань: Издательство Казанского университета, 2023. – Ч. II. – 555 с. С. 288–295.
148. Малыгин, А. А. Многостадийное адаптивное тестирование в системе оценивания результатов электронного обучения / А. А. Малыгин // Непрерывное образование: теория и практика реализации : материалы II Междунар. науч.-практ. конф. – Екатеринбург : Рос. гос. проф.-пед. ун-т, 2019. – С. 245–248.
149. Малыгин, А. А. Обеспечение качества оценок студентов в итоговой аттестации / А. А. Малыгин, М. Б. Челышкова // Отечественная и зарубежная педагогика. – 2023. – Т. 1, № 1 (89). – С. 7–23.
150. Малыгин, А. А. Обеспечение надежности и валидности при оценивании результатов обучения / А. А. Малыгин, Н. Ю. Сорокин // Проблемы высшего образования. – Хабаровск : ТОГУ, 2015. – С. 60–63.
151. Малыгин, А. А. Образовательное оценивание: от отметок в форме цифр до измерения компетенций / А. А. Малыгин // Отечественная и зарубежная педагогика. – 2023. – Т. 2, № 2 (92). – С. 71–89.

152. Малыгин, А. А. Образовательный процесс в магистратуре : оценка студентами ИГХТУ / А. А. Малыгин, М. Б. Клейман, Н. Ю. Котвина // Известия вузов. Сер.: Гуманитарные науки. – 2013. – Т. 4, вып. 3. – С. 192–197.
153. Малыгин, А. А. Оценивание правовой компетентности будущих педагогов-психологов / А. А. Малыгин, С. А. Слепнев // Психолого-педагогическая магистратура: теоретические исследования и практические результаты : моногр. / О.Н. Масленникова, Т. А. Воронова, А. А. Малыгин, С. В. Кузьмин [и др.]; под науч. ред. А. А. Малыгина, О. Н. Масленниковой. – Иваново : Иван. гос. ун-т, 2022. – С. 155–179.
154. Малыгин, А. А. Оценивание при аттестации студентов в современном высшем образовании: цели и подходы / А. А. Малыгин // Мир университетской науки: культура, образование. – 2023. – № 9. – С. 80–94.
155. Малыгин, А. А. Оценивание при аттестации студентов: адаптивный подход / А. А. Малыгин // Педагогика. – 2023. – № 9 (87). – С. 86–98.
156. Малыгин, А. А. Оценивание результатов обучения на основе образовательных измерений : учеб. пособие / А. А. Малыгин, Е. А. Соловьева. – Иваново : Иван. гос. ун-т, 2023. – 200 с.
157. Малыгин, А. А. Оценивание результатов электронного обучения на основе современной теории / А. А. Малыгин, М. Б. Челышкова // Университетский комплекс как региональный центр образования, науки и культуры : материалы Всерос. науч.-метод. конф., Оренбург. гос. ун-т. – Оренбург : ОГУ, 2019. – С. 5193–5201.
158. Малыгин, А. А. Педагогические измерения: разработка инструментария : учеб.-метод. пособие / А. А. Малыгин, М. Б. Челышкова. – Иваново : Иван. гос. ун-т, 2016. – 92 с.
159. Малыгин, А. А. Педагогические измерения: шкалирование результатов : учеб.-метод. пособие / В. И. Звонников, А. А. Малыгин. – Иваново : Иван. гос. ун-т, 2017. – 82 с.
160. Малыгин, А. А. Современная теория тестов как теоретическая основа современных подходов к оцениванию результатов обучения / А. А. Малыгин,

- С. В. Щаницина // Известия вузов. Сер.: Гуманитарные науки. – 2012. – Т. 3, вып. 4. – С. 324–327.
161. Малыгин, А. А. Современные практики оценивания результатов обучения в высшем образовании / А. А. Малыгин, Е. А. Соловьева, Ю. Ю. Травина // Научный поиск: личность, образование, культура. – 2021. – № 4 (42). – С. 3–13.
162. Малыгин, А. А. Современные форматы образовательного тестирования / А. А. Малыгин // Высшее образование сегодня. – 2018. – № 6. – С. 15–18.
163. Малыгин, А. А. Стратегии и алгоритмы реализации адаптивных технологий педагогических измерений / А. А. Малыгин // Вестник университета. – 2013. – № 15. – С. 393–402.
164. Малыгин, А. А. Теория педагогических измерений как научная основа для повышения эффективности контрольно-оценочных процессов в образовании / А. А. Малыгин // Известия вузов. Сер.: Гуманитарные науки. – 2010. – Т. 1, вып. 2. – С. 154–158.
165. Маслак, А. А. Измерение латентных переменных в социально-экономических системах : моногр. / А. А. Маслак. – Славянск-на-Кубани : Изд. центр СГПИ, 2006. – 333 с.
166. Массовое высшее образование триумф БРИК? / пер. с англ. М. С. Добряковой, Л. Ф. Пирожковой ; под науч. ред. М. С. Добряковой. – Москва : Изд. дом ВШЭ, 2014. – 528 с.
167. Минко, Н. Т. Педагогическое сопровождение компьютерного адаптивного тестирования в контексте персонального образования : дис. ... канд. пед. наук / Н. Т. Минко. – Улан-Удэ, 2010. – 209 с.
168. Митина, О. В. Математические методы в психологии : практикум / О. В. Митина. – Москва : Аспект Пресс, 2008. – 238 с.
169. Михайлычев, Е. А. Дидактическая тестология / Е. А. Михайлычев. – Москва : Нар. образование, 2001. – 432 с.
170. Михеев, В. И. Методы теории измерения в педагогике : учеб. пособие / В. И. Михеев. – Москва : Логос, 2003. – 64 с.



171. Модель уровневой оценки профессиональной компетентности учителя / С. А. Писарева, М. Ю. Пучков, С. В. Ривкина, А. П. Тряпицына // *Science for Education Today*. – 2019. – Т. 9, № 3. – С. 151-168.
172. Мясников, В. А. Измерение качества образования / В. А. Мясников, Н. Н. Найденова // *Педагогическая диагностика*. – 2007. – № 2. – С. 42–57.
173. Мясников, В. А. Компетенции и педагогические измерения / В. А. Мясников, Н. Н. Найденова // *Педагогическая диагностика*. – 2007. – № 5. – С. 29–56.
174. Мясников, В. А. Образование в глобальном измерении / В. А. Мясников, Н. Н. Найденова, И. А. Тагунова. – Москва : ИТИП РАО, 2009. – 512 с.
175. Найденова, Н. Н. Количественная методология педагогических измерений в междисциплинарных исследованиях непрерывного образования / Н. Н. Найденова // *Методология профессионального образования: сб. науч. ст. Междунар. науч.-практ. конф., посвящен. науч. вкладу акад. РАО Александра Михайловича Новикова (30 января 2018 г.)* / сост. М. А. Аксенова, С. И. Гудилина, М. Б. Яковлева ; под науч. ред. проф. Т. Ю. Ломакиной, М. В. Никитина. – Москва : РАО, 2018. – С. 205–209.
176. Найденова, Н. Н. Оценивание результатов обучения в процессе стандартизации обновленного содержания общего образования в зарубежных странах / Н. Н. Найденова // *Отечественная и зарубежная педагогика*. – 2022. – Т. 1, № 6 (88). – С. 140–157.
177. Найденова, Н. Н. Формирование репрезентативной выборки : учеб. пособие / Н. Н. Найденова. – Москва : Логос, 2003. – 176 с.
178. Наследов, А. Д. Компьютерный анализ данных в психологии и социальных науках / А. Д. Наследов. – 2-е изд. – Санкт-Петербург : Питер, 2007. – 416 с.
179. Насс, О. В. Теоретико-методические основания формирования компетентности преподавателей в области создания электронных образовательных ресурсов : (на базе адаптивных инструментальных комплексов) : дис. ... д-ра пед. наук / О. В. Насс. – Москва, 2013. – 374 с.

180. Нейман, Ю. М. Введение в теорию моделирования и параметризации педагогических тестов / Ю. М. Нейман, В. А. Хлебников. – Москва : Прометей, 2000. – 168 с.
181. Новиков, А. М. Методология учебной деятельности: контроль, оценка, рефлексия / А. М. Новиков // Педагогическая диагностика. – 2015. – № 2. – С. 16–23.
182. Новиков, А. М. Методология : словарь системы основных понятий / А. М. Новиков, Д. А. Новиков. – Москва : ЛЕНАНД, 2021. – 208 с.
183. Новиков, А. М. Методология / А. М. Новиков, Д. А. Новиков. – 2-е изд., испр. – Москва : КРАСАНД, 2014. – 632 с.
184. Новое качество высшего образования в современной России : (содержание, механизмы реализации, долгосрочные и ближайшие перспективы) : концептуально-программный подход : тр. исслед. центра / под науч. ред. д-ра техн. наук Н. А. Селезневой, д-ра экон. наук А. И. Субетто. – Москва : Исслед. центр пробл. качества подгот. специалистов, 1995. – 199 с.
185. Новые результаты образования : технологии проектирования, измерения и оценки качества моногр. / Н. А. Трубицына, Н. А. Баранова, Т. М. Банникова, А. В. Глазкова. – Ижевск : Изд-во «Удмуртский университет», 2011. – 214 с.
186. О независимой оценке квалификации : Федер. закон № 238-ФЗ от 3 июля 2016 г. – URL: <http://consultant.ru/> (дата обращения: 15.11.2023).
187. О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы : Указ Президента РФ от 9 мая 2017 г. № 203. – URL: <http://consultant.ru/> (дата обращения: 15.11.2023).
188. Об образовании в Российской Федерации : Федер. закон № 273-ФЗ от 29 декабря 2012 г. : с изменениями 2023 г. (включает все изменения до 1 января 2024 г.). – URL: <http://consultant.ru/> (дата обращения: 12.12.2023).
189. Об утверждении Порядка проведения государственной итоговой аттестации по образовательным программам высшего образования – программам бакалавриата, программам специалитета и программам магистратуры : приказ

- Минобрнауки России № 636 от 29 июня 2015 г. – URL: <http://consultant.ru/> (дата обращения: 10.09.2023).
190. Об утверждении программы «Цифровая экономика Российской Федерации» : распоряжение Правительства РФ от 28.07.2017 г. № 1632 р. – URL: <http://consultant.ru/> (дата обращения: 12.12.2018).
191. Образование для сложного общества : докл. Global Education Futures. – URL: [https://futuref.org/educationfutures\\_ru](https://futuref.org/educationfutures_ru) (дата обращения: 10.10.2018).
192. Образовательная аналитика: управление образовательной организацией и создание контента на основе данных / М. Б. Свердлов (науч. ред.), Е. В. Вербицкий, А. В. Конобеев, А. И. Крецу, В. Д. Стриканов ; Нац. исслед. ун-т «Высшая школа экономики», Ин-т образования. – Москва : НИУ ВШЭ, 2021. – 65 с. – (Современная аналитика образования ; № 9 (58)).
193. Осмоловская, И. М. Дидактика : от классики к современности : моногр. / И. М. Осмоловская. – Москва ; Санкт-Петербург : Нестор-История, 2020. – 248 с.
194. Остапенко, Р. И. Основы структурного моделирования в психологии и педагогике : учеб. пособие для студентов и аспирантов психол. и пед. спец. вузов / Р. И. Остапенко. – Воронеж : ВГПУ, 2012. – 120 с.
195. Отечественные университеты в динамике золотого века русской культуры / Е. В. Олесеюк (отв. ред.) [и др.] ; Федер. центр образоват. законодательства. – Санкт-Петербург : Союз, 2005. – 370 с.
196. Оценивание: образовательные возможности : сб. науч.-метод. ст. / редкол. : Т. И. Краснова (отв. ред.) [и др.] ; под общ. ред. М. А. Гусаковского. – Минск : БГУ, 2006. – Вып. 4. – 257 с.
197. Оценка информационно-коммуникационной компетентности учащихся: подходы, инструмент, валидность и надежность результатов / С. М. Авдеева, М. Г. Руднев, Г. М. Васин, К. В. Тарасова, Д. М. Панова // Вопросы образования. – 2017. – № 4. – С. 104–132.
198. Оценка программ: методология и практика / под ред. А. И. Кузьмина, Р. О’Салливан, Н. А. Кошелевой. – Москва : Престо-РК, 2009. – 396 с.

199. Оценка профессиональной готовности специалистов в системе здравоохранения / под ред. Т. В. Семеновой. Москва : ГЭОТАР-Медиа, 2019. – 272 с.
200. Панасюк, В. П. Системное управление качеством образования в школе / В. П. Панасюк. – Санкт-Петербург ; Москва : ИЦПКПС, 2000. – 239 с.
201. Педагогическое образование в современной России: стратегические ориентиры развития : моногр. / Юж. федер. ун-т ; науч. ред. Ю. П. Зинченко. – Ростов-на-Дону ; Таганрог : Изд-во Юж. федер. ун-та, 2020. – 612 с.
202. Педагогическое образование в университете: контекстно-биографический подход : моногр. / под ред. А. Л. Гаврикова, М. Н. Певзнера. – Великий Новгород : НовГУ им. Ярослава Мудрого, 2001. – 200 с.
203. Переверзев, В. Ю. Критериально-ориентированное педагогическое тестирование : учеб. пособие / В. Ю. Переверзев. – Москва : Логос, 2003. – 120 с.
204. Переосмысление инженерного образования. Подход CDIO / Э. Ф. Кроули [и др.] ; пер. с англ. С. Рыбушкиной ; науч. ред. А. Чучалин. – 2-е изд. – Москва : ВШЭ, 2015. – 497 с.
205. Персонализация в образовании: от программируемого к адаптивному обучению / Д. А. Кравченко, И. А. Блескина, Е. Н. Каляева, Е. А. Землякова, Д. Ф. Аббакумов // Современная зарубежная психология. – 2020. – Т. 9, № 3. – С. 34–46.
206. Перспективные подходы к оцениванию компетентности обучающихся в современном вузе / Ж. М. Сизова, Т. В. Семенова, Н. Н. Найденова, В. И. Звонников, М. Б. Челышкова, А. А. Малыгин // Образовательное пространство в информационную эпоху = Education Environment for the Information Age – 2019 : сб. науч. тр. Междунар. науч.-практ. конф. / под ред. С. В. Ивановой. – Москва : ИСРО РАО, 2019. – С. 559–572.
207. Пещеров, Д. В. Проблема оценивания и рейтингования учебных достижений в высшей школе России в конце XIX – начале XX века: опыт изучения архивных материалов / Д. В. Пещеров // Наука и школа. – 2022. – № 2. – С. 106–113.

208. Поливанова, К. Н. Национальные и международные программы оценки качества образования : учеб. пособие к курсу / К. Н. Поливанова. – Москва : Моск. высш. шк. соц. и экон. наук, 2012. – 154 с.
209. Полонский, В. М. Оценка качества научно-педагогических исследований : учеб. пособие / В. М. Полонский. – 2-е изд., перераб. и доп. – Москва : ИНФРА-М, 2023. – 220 с.
210. Попова, Ю. Б. От LMS к адаптивным обучающим системам / Ю. Б. Попова // Системный анализ и прикладная информатика. – 2019. – № 2. – С. 58–64.
211. Поташник, М. М. Качество образования: проблемы и технология управления : (в вопросах и ответах) : учеб. пособие / М. М. Поташник. – Москва : Пед. о-во России, 2002. – 352 с.
212. Поташник, М. М. Освоение ФГОС : метод. материалы для учителя : метод. пособие / М. М. Поташник, М. В. Левит. – Москва : Пед. о-во России, 2016. – 208 с.
213. Правила для желающих поступить в университеты // Сборник распоряжений по Министерству народного просвещения. – Санкт-Петербург : Тип. Императ. акад. наук, 1866. – Т. 2 : 1835–1849 гг. – Стб. 173.
214. Равен, Д. Педагогическое тестирование: проблемы, заблуждения, перспективы : пер. с англ. / Д. Равен. – Москва : Когито-Центр, 2001. – 144 с.
215. Резаев, А. В. ChatGPT и искусственный интеллект в университетах: какое будущее нам ожидать? / А. В. Резаев, Н. Д. Трегубова // Высшее образование в России. – 2023. – Т. 32, № 6. – С. 19–37.
216. Рекомендации по применению профессиональных стандартов при разработке образовательных программ / О. М. Зайцева, П. Н. Новиков, Ю. Э. Волошановская, А. А. Гончарова, Н. М. Савина. – Москва : Перо, 2023. – 39 с.
217. Решетникова, О. А. Банк заданий для оценки читательской грамотности как инструмент формирования блока метапредметных результатов, связанных с работой с информацией: общие подходы / О. А. Решетникова, М. Ю. Демидова // Педагогические измерения. – 2023. – № 1. – С. 4–12.

218. Рудинский, И. Д. Компетенция. Компетентность. Компетентностный подход / И. Д. Рудинский, Н. А. Давыдова, С. В. Петров ; под ред. И. Д. Рудинского. – Москва : Горячая линия – Телеком, 2018. – 240 с.
219. Руководство по разработке и применению оценочных средств для аттестации и аккредитации : учеб. пособие для авт. оценоч. средств на прим. группы мед. спец. / Т. В. Семенова, Ж. М. Сизова, В. И. Звонников, М. Б. Чельшкова. – Санкт-Петербург : Санкт-Петербург. гос. педиатр. мед. ун-т М-ва здравоохранения РФ, 2021. – 400 с.
220. Рыжов, А. Н. Генезис подходов к оцениванию учебной деятельности в среднем общем образовании в России (XVIII – начало XX в.) / А. Н. Рыжов // Казанская наука. – 2017. – № 12. – С. 194–197.
221. Рыжов, А. Н. Подходы к оцениванию учебных достижений в среднем общем образовании России в конце XIX – 80-х гг. XX в. / А. Н. Рыжов // Школа будущего. – 2017. – № 6. – С. 40–48.
222. Сабельникова, Е. В. Результаты обучения в высшем образовании на национальном уровне на примере проекта «Оценка результатов обучения в высшем образовании» (CLA) / Е. В. Сабельникова, Н. Л. Хмелева // Психологическая наука и образование. – 2015. – Т. 20, № 2. – С. 16–23.
223. Савинков, В. И. Востребованность и оценка населением качества образования / В. И. Савинков, П. А. Бакланов ; под ред. акад. Г. В. Осипова. Москва : ИСПИ РАН, 2017. – 296 с.
224. Сагитов, Е. Б. Как педагогические практики связаны со списыванием среди студентов российских вузов / Е. Б. Сагитов, Е. Д. Шмелева // Вопросы образования = Educational studies Moscow. – 2022. – № 1. – С. 138–159.
225. Сафонцев, С. А. Образовательная квалиметрия как фактор повышения эффективности контроля качества обучения : дис. ... д-ра пед. наук / С. А. Сафонцев. – Ростов-на-Дону, 2004. – 395 с.
226. Сердюков, В. И. Направления совершенствования автоматизированных систем контроля результатов обучения / В. И. Сердюков, Н. А. Сердюкова // Информатизация образования и науки. – 2014. – № 3 (23). – С. 75–85.

227. Сизова, Ж. М. Оценка профессиональной готовности специалистов здравоохранения при аккредитации / Ж. М. Сизова, Т. В. Семенова, М. Б. Челышкова // Медицинский вестник Северного Кавказа. – 2017. – Т. 12, № 4. – С. 459–463.
228. Сизова, Ж. М. Совершенствование качества оценочного инструментария для аккредитации специалистов здравоохранения / Ж. М. Сизова, М. Б. Челышкова // Медицинское образование и вузовская наука. – 2018. – № 1 (11). – С. 19–23.
229. Смыслова, С. Проектирование образовательного опыта / С. Смыслова. – Москва : Универс. ун-т, 2022. – 320 с.
230. Спенсер, Л. Компетенции на работе. Модели максимальной эффективности работы : пер. с англ. / Л. Спенсер, С. Спенсер. – Москва : НИРРО, 2005. – 384 с.
231. Справедливость оценок в аккредитации специалистов как проблема / В. И. Звонников, А. А. Малыгин, Т. В. Семенова, Ж. М. Сизова, М. Б. Челышкова // Ценности и смыслы. – 2023. – № 2 (84). – С. 53–71.
232. Стариченко, Б. Е. Особенности проведения демонстрационного экзамена в вузе / Б. Е. Стариченко, Л. В. Сардак // Педагогическое образование в России. – 2023. – № 2. – С. 123–132.
233. Стрекалова, Н. Б. Управление качеством самостоятельной работы студентов в открытой информационно-образовательной среде : автореф. дис. ... д-ра пед. наук / Н. Б. Стрекалова. – Самара, 2017. – 53 с.
234. Субетто, А. И. Онтология и эпистемология компетентностного подхода, классификация и квалиметрия компетентностей / А. И. Субетто. – Санкт-Петербург : Исслед. центр пробл. качества подгот. специалистов, 2006. – 72 с.
235. Субетто, А. И. Оценочные средства и технологии аттестации качества подготовки специалистов в вузах: методология, методика, практика : моногр. / А. И. Субетто. – Москва : Исслед. центр пробл. качества подгот. специалистов, 2005. – 284 с.
236. Тарасова, К. В. Измерение критического мышления студентов в открытой онлайн-среде: концептуальная рамка и типология заданий / К. В. Тарасова,

- Е. А. Орел // Вопросы образования = Educational studies Moscow. – 2022. – № 3. – С. 187–212.
237. Тарханова, И. Ю. Современные регуляторы становления новой дидактики высшего образования / И. Ю. Тарханова // Ярославский педагогический вестник. – 2019. – № 2. – С. 45–52.
238. Татур, Ю. Г. Образовательный процесс в вузе : методология и опыт проектирования : учеб. пособие для слушателей системы подгот. и повышения квалификации преподавателей / Ю. Г. Татур. – Москва : Изд-во МГТУ им. Н. Э. Баумана, 2005. – 224 с.
239. Терюха, Р. В. Технология адаптивного компьютерного тестирования в профессиональной подготовке инженеров : дис. ... канд. пед. наук / Р. В. Терюха. Краснодар, 2006. – 261 с.
240. Томильцев, А. В. Проблемы оценки профессиональной подготовки: методологические подходы / А. В. Томильцев, А. В. Мальцев // Образование и наука. – 2018. – Т. 20, № 4. – С. 9–33.
241. Трайнев, В. А. Дистанционное обучение и его развитие: обобщение методологии и практики использования / В. А. Трайнев, В. Ф. Гуркин, О. В. Трайнев. – Москва : Дашков и Ко, 2007. – 294 с.
242. Трапицын, С. Ю. Теоретические основы управления качеством образовательного процесса в военном вузе : дис. ... д-ра пед. наук / С. Ю. Трапицын. – Санкт-Петербург, 2000. – 370 с.
243. Третьякова, Т. В. Система оценки качества образования и ее построение в регионах с территориальными и национальными особенностями : дис. ... д-ра пед. наук / Т. В. Третьякова. – Москва, 2010. – 464 с.
244. Трудности и перспективы цифровой трансформации образования / А. Ю. Уваров, Э. Гейбл, И. В. Дворецкая [и др.] ; под ред. А. Ю. Уварова, И. Д. Фрумина; Нац. исслед. ун-т «Высшая школа экономики», Ин-т образования. – Москва : Изд. дом ВШЭ, 2019. – 343 с. – (Российское образование: достижения, вызовы, перспективы).



245. Туулик, М. А. Оценивание как социально-воспитательное явление : дис. ... д-ра пед. наук / М. А. Туулик. – Таллинн, 1991. – 495 с.
246. Угланова, И. Л. Методология Evidence-Centered Design для измерения комплексных психологических конструктов / И. Л. Угланова, И. В. Брун, Г. М. Васин // Современная зарубежная психология. – 2018. – Т. 7, № 3. – С. 18–27.
247. Унт, И. К. Индивидуализация и дифференциация обучения / И. К. Унт. – Москва : Педагогика, 1990. – 192 с.
248. Управление качеством образования : практ.-ориентиров. моногр. и метод. пособие / М. М. Поташник, Е. А. Ямбург, Д. Ш. Матрос, А. Моисеев ; под ред. М. М. Поташника. – 2-е изд., перераб. и доп. – Москва : Пед. о-во России, 2004. – 448 с. – (Менеджмент в образовании).
249. Ушинский, К. Д. Несколько слов об училищных отметках (баллах), употребляемых в наше время во всей России / К. Д. Ушинский // Архив К. Д. Ушинского. – Москва : Изд. АПН РСФСР, 1960. – Т. 2. – С. 225–271.
250. Факторович, А. А. Методология ценностно-мотивационного управления качеством образования в вузе : дис. ... д-ра пед. наук / А. А. Факторович. – Москва, 2012. – 401 с.
251. Факторович, А. А. Принципы управления качеством образования в вузе в условиях реализации ФГОС / А. А. Факторович // Высшее образование в России. – 2010. – № 12. – С. 42–51.
252. Федерякин, Д. А. Новые источники информации в компьютерном тестировании / Д. А. Федерякин, И. Л. Угланова, М. А. Скрябин // Вестник Томского государственного университета. – 2021. – № 465. – С. 179–187.
253. Фер, Р. М. Психометрика : введение / Р. Майкл Фер, Верн Р. Бакарак ; пер. с англ. [А. С. Науменко, А. Ю. Попов] ; под ред. Н. А. Батурина, Е. В. Эйдмана. – Челябинск : Изд. центр ЮУрГУ, 2010. – 445 с.
254. Фомина, Н. Б. Оценка качества образования. Ч. 4 : Новые способы оценивания учащихся : метод. пособие / Н. Б. Фомина. – Москва : Перспектива, 2009. – 48 с.

255. Хей, Дж. Введение в методы байесовского статистического вывода / Дж. Хей. – Москва : Финансы и статистика, 1987. – 335 с.
256. Хлопотов, М. В. Применение байесовской сети при построении моделей обучающихся для оценки уровня сформированности компетенций / М. В. Хлопотов // Наукоеведение : интернет-журнал. – 2014. – № 5 (24). – URL: <https://naukovedenie.ru/PDF/20TVN514.pdf> (дата обращения: 12.12.2018).
257. Хлопотов, М. В. Методы интеллектуального анализа данных для мониторинга и диагностики качества образования / М. В. Хлопотов, И. Ю. Коцюба // Дистанционное и виртуальное обучение. – 2014. – № 5 (83). – С. 18–25.
258. Хуторской, А. В. Методологические основания применения компетентностного подхода к проектированию образования / А. В. Хуторской // Высшее образование в России. – 2007. – № 12 (218). – С. 85–91.
259. Хуторской, А. В. Современная дидактика : учеб. / А. В. Хуторской. – Санкт-Петербург : Питер, 2001. – 544 с.
260. Чекалева, Н. В. Проблемы оценки компетенций студентов вуза / Н. В. Чекалева // Философия образования. – 2015. – № 1 (58). – С. 146–156.
261. Чельшкова, М. Б. Аттестация выпускников вузов в рамках компетентностного подхода / М. Б. Чельшкова // Вестник Костромского государственного университета им. Н. А. Некрасова. – 2012. – Т. 18, № 6. – С. 270–273.
262. Чельшкова, М. Б. Теоретико-методологические и технологические основы адаптивного тестирования в образовании : дис. ... д-ра пед. наук / М. Б. Чельшкова. – Москва, 2001. – 324 с.
263. Чельшкова, М. Б. Теория и практика конструирования педагогических тестов : учеб. пособие / М. Б. Чельшкова. – Москва : Логос, 2002. – 432 с.
264. Шил, Д. Анализ данных национальной оценки учебных достижений / Д. Шил, Ф. Картрайт ; ред. сер. В. Грини, Т. Келлаган ; пер. с англ. под науч. ред. В. И. Звонникова, М. Б. Чельшковой. – Москва : Логос, 2015. – 336 с. – (Национальная оценка учебных достижений).

265. Шихов, Ю. А. Проектирование и реализация комплекса квалиметрического мониторинга подготовки обучающихся в системе «профильная школа – вуз» : автореф. дис. ... д-ра пед. наук / Ю. А. Шихов. – Ижевск, 2008. – 32 с.
266. Шкалирование и выравнивание результатов педагогических измерений / В. И. Звонников, Н. Н. Найденова, С. В. Никифоров, М. Б. Чельшкова. – Москва : Логос, 2003. – 96 с.
267. Шмелев, А. Г. Основы психодиагностики : учеб. пособие для студентов педвузов / А. Г. Шмелев. – Москва ; Ростов на-Дону : Феникс, 1996. – 544 с.
268. Шмелев, А. Г. Практическая тестология : тестирование в образовании, прикладной психологии и управлении персоналом / А. Г. Шмелев. – Москва : Маска, 2013. – 668 с.
269. Шмелев, А. Г. Тестология – это не только психометрические тесты / А. Г. Шмелев // Вестник по педагогике и психологии Южной Сибири. – 2019. – № 4. – С. 61–75.
270. Яковлев, Е. В. Теория и практика внутривузовского управления качеством образования : дис. ... д-ра пед. наук / Е. В. Яковлев. – Челябинск, 2000. – 418 с.
271. Ярославцев, В. Эффективность и качество компьютерного тестирования / В. Ярославцев // Стандарты и качество. – 2002. – № 9. – С. 83–85.
272. *Advances in Educational and Psychological Testing : Theory and Applications* / ed. by R. K. Hambleton, J. Zaal. – 4th ed. – Boston : Kluwer, 2000. – 458 p.
273. Aiken, L. R. *Tests and examinations : measuring abilities and performance* / L. R. Aiken. – New York : John Wiley & Sons, 1998. – 302 p.
274. Andrich, D. *Rasch Models for Measurement* / D. Andrich. – Newbury Park (CA) : Sage Publ., 1988. – 96 p.
275. Angoff, W. A. *Scales, Norm and Equating Scores* / W. A. Angoff. – Princeton (NJ) : Educational Testing Service, 1984. – 144 p.
276. *Assessment, Learning and Judgement in Higher Education* / ed. by G. Joughin. – New York : Springer Science : Business Media B.V., 2009. – 233 p.

277. Automated Scoring of Complex Tasks in Computer-Based Testing / ed. by D. M. Williamson, R. J. Mislevy, I. I. Bejar. – Mahwah (NJ) : Lawrence Erlbaum Associates, 2006. – 440 p.
278. Automatic Item Generation : Theory and Practice / ed. by M. J. Gierl, T. M. Haladyna. – New York : Teylor & Francis, 2013. – 246 p.
279. Bandalos, D. L. Measurement Theory and Applications for the Social Sciences / D. L. Bandalos. – New York : Guilford, 2018. – 661 p.
280. Bartram, D. Computer-based Testing and the Internet : Issues and Advances / D. Bartram, R. K. Hambleton. – London : Wiley, 2006. – 263 p.
281. Bayes nets in educational assessment : where do the numbers come from? / R. J. Mislevy, R. G. Almond, D. Yan, L. S. Steinberg // Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence / ed. by K. B. Laskey, H. Prade. – San Francisco : Kaufmann, 1999. – P. 437–446.
282. Becker, K. A. Test administration models / K. A. Becker, B. A. Bergstrom // Practical Assessment, Research & Evaluation. – 2013. – Vol. 18. – URL: <http://pareonline.net/getvn.asp?v=18&n=14> (дата обращения: 14.11.2018).
283. Berk, R. A. Criterion-Referenced Measurement: The State of the Art / R. A. Berk. – Baltimor, MD: Johns Hopkins University Press, 1980. – 234 p.
284. Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability / A. Birnbaum // Statistical Theories of Mental Test Scores / F. M. Lord, M. R. Novick. – Reading (MA) : Addison-Wesly, 1968. – P. 397–479.
285. Bloom, B. S. Taxonomy of Educational Objectives / B. S. Bloom. Handbook 1 : Cognitive Domain. – New York : Longmans Green, 1956. – 207 p.
286. Bock, R. D. Item Response Theory / R. D Bock, R. D. Gibbons; University of Chicago. – Hoboken : Wiley, 2021. – 384 p.
287. Brennan, R. Generalizability Theory / R. Brennan. – New York : Springer, 2001. – 538 p.
288. Carlson, J. E. Operational Best Practices for Statewide Large-Scale Assessment Programs / J. E. Carlson. – Washington : Council of Chief State School Officers : The Association of Test Publishers, 2013. – 216 p.

289. Chatterji, M. Designing and using tools for educational assessment / M. Chatterji. – Boston : Allyn & Bacon, 2003. – 498 p.
290. Cizek, G. J. Standards Setting / G. J. Cizek, M. B. Bunch. – New York : Sage Publications, 2010. – 318 p.
291. Competence Assessment in Education Research, Models and Instruments / D. Leutner, J. Fleischer, J. Grünkorn, E. Klieme. – New York : Springer International Publishing AG, 2017. – 503 p.
292. Competence Development and Assessment in TVET : Theoretical Framework and Empirical Results / F. Rauner, L. Heinemann, A. Maurer, B. Haasler, B. Erdwien, T. Martens. – New York : Springer Science + Business Media Dordrecht, 2013. – 169 p.
293. Computer-based Testing: Building the Foundation for Future Assessments / C. N. Mills, M. T. Potenza, J. J. Fremer, W. C. Ward. – Mahwah (NJ) : Erlbaum Associates, 2002. – 326 p.
294. Crocker, L. Introduction to Classical and Modern Test Theory / L. Crocker, J. Algina. – Pacific Grove (CA) : Wadsworth, 2006. – 527 p.
295. Cross-analysis of big data in accreditation of health specialists / M. B. Chelyshkova, T. V. Semenova, N. N. Naydenova, E. M. Dorozhkin, A. A. Malygin, V. V. Akhunov // Electronic Journal of General Medicine. –2018. – Vol. 15, iss. 5. – <https://doi.org/10.29333/ejgm/93469>.
296. Cultural Validity in Assessment / ed. by M. Basterra, E. Trumbull, G. Solano-Flores. – New York : Routledge, 2011. – 300 p.
297. Data Analytics and Psychometrics: Informing Assessment Practices / ed. by H. Jiao, R. W. Lissitz, A. Van Wie. – Charlotte (NC) : Information Age Publishing, 2018. – 268 p.
298. Davier, A. A. von. Innovative Assessment of Collaboration / A. F. von Davier, M. Zhu, P. C. Kyllonen. – New York : Springer International Publishing Switzerland, 2017. – 343 p.

299. Desjardins, C. D. Handbook of Educational Measurement and Psychometrics Using R / C. D. Desjardins, O. Bulut. – Boca Raton (Florida) : CRC Press Taylor & Francis Group, 2018. – 302 p.
300. Drasgow, F. Technology and Testing: Improving Educational and Psychological Measurement / F. Drasgow. – New York : Routledge, 2015. – 376 p.
301. Educational and Psychological Measurement / ed. by W. H. Finch, B. F. French. – New York : Routledge, 2018. – 468 p.
302. Educational Measurement / ed. by R. L. Linn. – New York : Macmillan, 1989. – 609 p.
303. Essays on Contemporary Psychometrics / ed. by L. A. van der Ark, W. H. M. Emons, R. R. Meijer. – Cham : Springer, 2023. – 456 p.
304. Essays on item response theory / ed. by A. Boomsma, M. A. J. van Duijn, T. A. B. Snijders. – New York : Springer, 2001. – 464 p.
305. ETS Standards for Quality and Fairness / Educational Testing Service. – Princeton, 2015. – 63 p.
306. Fairness and quality of data in healthcare professionals' accreditation / T. V. Semenova, Zh. M. Sizova, M. B. Chelyshkova, E. M. Dorozhkin, A. A. Malygin // Modern Journal of Language Teaching Methods. – 2017. – Vol. 7, iss. 9. – P. 13–25.
307. Fairness in Educational Assessment and Measurement / ed. by N. J. Dorans, L. L. Cook. – New York : Routledge, 2016. – 328 p.
308. Garden, J. Assessment and Learning : an Introduction / J. Garden. – New York : Sage Publication Ltd., 2006. – 229 p.
309. Gill, J. Bayesian Methods : a Social and Behavioral Sciences Approach. – 3rd ed. – New York : CRC Press Taylor & Francis Group, 2015. – 689 p.
310. Glaser, R. Instructional technology and the measurement of learning outcomes // American Psychologist. – 1963. – Vol. 18. – P. 519–521.
311. Griffin, P. Assessment and Teaching of 21st Century Skills / P. Griffin, B. McGaw, E. Care. – Dordrecht : Springer Science + Business Media B.V., 2012. – 345 p.

312. Gronlund, N. E. *How to Construct Achievement Tests* / N. E. Gronlund. – New Jersey : Prentice Hall, 1988. – 220 p.
313. Gruijter, D. N. de. *Statistical Test Theory for the Behavioral Sciences* / D. N. de Gruijter, L. J. Th. van der Kamp. – New York : Chapman & Hall / CRC, 2008. – 264 p.
314. Hambleton, R. K. *Fundamentals of Item Response Theory* / R. K. Hambleton, H. Swaminathan, H. J. Rogers. – New York : Sage Publications, 1991. – 174 p.
315. *Handbook of Human and Social Conditions in Assessment* / ed. by G. T. Brown, L. R. Harris. – New York : Routledge, 2016. – 587 p.
316. *Handbook of Test Development* / ed. by S. M. Downing, T. M. Haladyna. – New York : Lawrence Erlbaum Associates, Inc., 2006. – 778 p.
317. Individual educational paths in training of masters in pedagogical universities / A. A. Malygin, M. R. Moskalenko, M. V. Ozhiganova, E. M. Dorozhkin, N. V. Chekaleva // *Astra Salvensis. Supplement.* – 2018. – № 2. – P. 125–139.
318. Innovative approaches to increasing the student assessment procedures effectiveness / E. M. Dorozhkin, M. B. Chelyshkova, A. A. Malygin [et al.] // *International Journal of Environmental and Science Education.* – 2016. – Vol. 11, iss. 14. – P. 7129–7144.
319. Integrating data-based decision making, assessment for learning and diagnostic testing in formative assessment / F. M. van der Kleij, J. A. Vermeulen, K. Schildkamp, T. J. H. M. Eggen // *Assessment in Education : Principles, Policy & Practice.* – 2015. – Vol. 22, № 3. – P. 324–343.
320. *Item Generation for Test Development* / ed. by S. H. Irvine, P. C. Kyllonen. – London : Lawrence Erlbaum Associates, Inc., 2002. – 412 p.
321. Keeves, J. P. *Educational Research Methodology and Measurement : an International Handbook* / J. P. Keeves. – New York : Perg. Press, 1988. – 832 p.
322. Kingsbury, G. Procedures for selecting items for computerized adaptive tests / G. Kingsbury, A. Zara // *Applied Psychological Measurement.* – 1989. – № 2. – P. 359–375.

323. Kline, R. B. Principles and Practice of Structural Equation Modeling / R. B. Kline. – 4th ed. – New York : The Guilford Press, 2016. – 554 p.
324. Kolen, M. J. Test Equating, Scaling, and Linking : Methods and Practices / M. J. Kolen, R. L. Brennan. – 3rd ed. – New York : Springer Science + Business Media, 2014. – 582 p. – (Statistics for Social and Behavioral Sciences).
325. Large-scale assessment programs for all students : validity, technical adequacy, and implementation / ed. by G. Tindal, T. M. Haladyna. – Mahwah : Lawrence Erlbaum Associates, 2002. – 536 p.
326. Lee, G. Incorporating the testlet concept in test score analyses / G. Lee, R. L. Brennan, D. A. Frisbie // Educational Measurement: Issues and Practice. – 2000. – Vol. 19, № 4. – P. 9–15.
327. Levy, R. Bayesian Psychometric Modeling / R. Levy, R. J. Misley. – New York : CRC Press, 2016. – 480 p.
328. Linden, W. J. van der. Elements of Adaptive Testing, Statistical for Social and Behavioral Sciences / W. J. van der Linden, C. A. W. Glas. – New York : Springer Science + Business Media, LLC, 2010. – 437 p.
329. Linden, W. J. van der. Handbook of Item Response Theory : Models / W. J. van der Linden. – New York : CRC Press, 2016. – 624 p.
330. Lord, F. M. Application of Item Response Theory to Practical Testing Problems / F. M. Lord. – Hillsdale (NJ) : Lawrence Erlbaum Ass., Publ., 1980. – 274 p.
331. Lord, F. M. Statistical Theories of Mental Test Scores / F. M. Lord, M. R. Novick. – Reading (MA) : Addison-Wesley, 1968. – 568 p.
332. Magis, D. Computerized Adaptive and Multistage Testing with R : Using Packages catR and mstR / D. Magis, D. Yan, A. A. von Davier. – New York : Springer International Publishing AG, 2017. – 183 p.
333. McDonald, R. P. Test theory : a unified treatment / R. P. McDonald. – Mahwah : Lawrence Erlbaum Assotiates, 1999. – 486 p.
334. Mertens, D. M. Research and evaluation in education and psychology : integrating diversity with quantitative, qualitative, and mixed methods / D. M. Mertens. – Thousand Oaks : Sage Publications, 2005. – 507 p.



335. Miller, M. D. *Measurement and Assessment in Teaching* / M. D. Miller, R. L. Linn, N. E. Gronlund. – 10th ed. – New York : Library of Congress Cataloging in Publication Data, 2009. – 552 p.
336. *Modeling and Measuring Competencies in Higher Education: Tasks and Challenges* / S. Blömeke, O. Zlatkin-Troitschankaia, C. Kuhn, J. Fege. Rotterdam : Sense Publishers, 2013. – 236 p.
337. *Multidimensional Scaling : History, Theory and Applications* / ed. by F. W. Young, R. M. Hammer. – Hillsdale ; New York : Lawrence Erlbaum Associates, 1987. – 318 p.
338. Netemeyer, R. G. *Scaling Procedures: Issues and applications* / R. G. Netemeyer, W. O. Bearden, S. Sharma. – New York : Sage Publ., 2003. – 206 p.
339. *New Horizons in Testing : Latent Trait Test Theory and Computerized Adaptive Testing* / ed. by D. J. Weiss. – New York : Academic Press, 1983. – 345 p.
340. *Objective Measurement : Theory into Practice* / ed. by M. Wilson. – Norwood (NJ) : Ablex Publ. Corp., 1992. – Vol. 1. – 360 p.
341. Owen, R. J. *A Bayesian Approach to Tailored Testing* / R. J. Owen. – Princeton (NJ) : Educational Testing Service, 1969. – 25 p. – (Research Report ; № 69–92).
342. Petersen, N. S. *Scaling, norming, and equating* / N. S. Petersen // *Educational Measurement* / N. S. Petersen, M. J. Kolen, H. D. Hoover ; ed. by R. L. Linn. – 3rd ed. – New York : American Council on Education and Macmillan, 1989. – P. 221–262.
343. Popham, W. J. *Criterion-referenced measurement* / W. J. Popham. – Englewood Cliffs. New York : Prentice Hall, 1978.
344. *Practical Considerations in Computer-based Testing* / C. G. Parshall, J. A. Spray, J. C. Kalohn, T. C. Davey. – New York : Springer, 2002. – 221 p.
345. Rasch, G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago ; London : The Univ. of Chicago Press, 1980. 199 p.
346. Reise, S. P. *Handbook of Item Response Theory Modeling : Applications to Typical Performance Assessment* / S. P. Reise, D. A. Revicki. – New York : Taylor & Francis, 2015. – 554 p.

347. Revisiting development of evaluation tools fund during realization of competence-oriented educational programs / A. A. Malygin, E. D. Kolegova, O. B. Akimova, E. M. Dorozhkin, N. V. Chekaleva // *Astra Salvensis. Supplement.* – 2018. – № 2. – P. 286–313.
348. Rosen, M. *Cognitive Abilities and Educational Outcomes* / M. Rosen, K. Y. Hansen, U. Wolff. – New York : Springer International Publishing AG, 2017. – 327 p.
349. Rutkowski, L. *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis* / L. Rutkowski, M. von Davier, D. Rutkowski. – New York : CRC Press Taylor & Francis Group, 2014. – 623 p.
350. Stevens, S. S. On the theory of scales of measurement / S. S. Stevens // *Science. New Series.* – 1946. – Vol. 103. – P. 677–680.
351. *Theoretical and Practical Advances in Computer-based Educational Measurement* / ed. by B. P. Veldkamp, C. Sluijter. – Cham : Springer Nature, 2019. – 399 p.
352. *The Routledge Handbook of Language Testing* / ed. by G. Fulcher, F. Davidson. – New York : Taylor & Francis Books, 2012. – 551 p.
353. *The Sage handbook of quantitative methodology for the social sciences* / ed. by D. Kaplan. – Thousand Oaks : Sage Publications, 2004. – 512 p.
354. Thompson, N. A. A framework for the development of computerized adaptive tests / N. A. Thompson, D. J. Weiss // *Practical Assessment, Research & Evaluation.* – 2011. – № 16. – URL: <http://pareonline.net/getvn.asp?v=16&n=1> (дата обращения: 05.05.2018).
355. Tierney, R. D. Fairness in educational assessment / R. D. Tierney // *Encyclopedia of Educational Philosophy and Theory* / ed. by M. A. Peters. – Singapore : Springer Science + Business Media, 2016. – P. 793–798.
356. *Validity Generalization : a Critical Review* / ed. by K. Murphy. – New York : Lawrence Erlbaum Associates, 2003. – 443 p.
357. Viswanathan, M. *Measurement error and research design* / M. Viswanathan. – Thousand Oaks : Sage Publications, 2005. – 454 p.
358. Wainer, H. *Computerized Adaptive Testing : a Primer* / H. Wainer. – 2nd ed. – Mahwah (NJ) : Lawrence Erlbaum Associates, 2000. – 278 p.

359. Walsh, W. B. Test and Assessment / W. B. Walsh, N. E. Betz. – 4th ed. – Upper Saddle River : Prentice-Hall, 2001. – 481 p.
360. Weiss, D. J. Computerized adaptive testing for effective and efficient measurement in counseling and education / D. J. Weiss // Measurement and Evaluation in Counseling and Development. – 2004. – Vol. 37, № 2. – P. 70–84.
361. Willet, J. B. Questions and answers in the measurement of change / J. B. Willet // Review of Research in Education / ed. by E. Z. Rothkopf. – Washington (DC) : AERA, 1989. – Vol. 15. – P. 345–422.
362. Wilson, M. Constructing Measures : an Item Response Modeling Approach / M. Wilson. – Mahwah : Erlbaum, 2005. – 228 p.
363. Wilson, M. From principles to practice: an embedded assessment system / M. Wilson, K. Sloane // Applied Measurement in Education. – 2000. – Vol. 13, № 2. – P. 181–208.
364. Wools, S. All About Validity: an Evaluation System for the Quality of Educational Assessment / S. Wools. – 2015. – URL: [https://www.researchgate.net/publication/281434844\\_All\\_About\\_VValidity\\_An\\_evaluation\\_system\\_for\\_the\\_quality\\_of\\_educational\\_assessment](https://www.researchgate.net/publication/281434844_All_About_VValidity_An_evaluation_system_for_the_quality_of_educational_assessment) (дата обращения: 20.12.2019).
365. Wright, B. D. Best Test Design / B. D. Wright, M. H. Stone. – Chicago : Messa Press, 1979. – 245 p.

## ПРИЛОЖЕНИЯ

### Приложение 1

#### Виды моделей современной теории тестов (IRT)

Число параметров, входящих в аналитическое задание функций, является основанием для подразделения семейства IRF на классы.

Для дихотомической шкалы, когда результаты выполнения заданий испытуемым оцениваются 0 или 1 (не выполнил или выполнил соответственно), используются модели Раша, Бирнбаума и трехпараметрическая модель. Ниже приводятся указанные модели.

*Однопараметрическая модель (1PL) или модель Г. Раша:*

$$P_j(\theta) = \frac{e^{1,7(\theta - \beta_j)}}{1 + e^{1,7(\theta - \beta_j)}}, \quad (1)$$

$$P_i(\beta) = \frac{e^{1,7(\theta_i - \beta)}}{1 + e^{1,7(\theta_i - \beta)}}, \quad (2)$$

где  $\theta$  и  $\beta$  – независимые переменные для первой и второй функции соответственно.

*Двухпараметрическая модель (2PL) А. Бирнбаума:*

$$P_j(\theta) = \frac{e^{1,7a_j(\theta - \beta_j)}}{1 + e^{a_j 1,7(\theta - \beta_j)}}, \quad (3)$$

$$P_i(\beta) = \frac{e^{1,7a_i(\theta_i - \beta)}}{1 + e^{1,7a_j(\theta_i - \beta)}}, \quad (4)$$

где  $a_j$  – параметр дифференцирующей способности задания, указывающий дифференцирующую способность задания при измерении различных значений уровня подготовленности  $\theta$  и принимающий значения в интервале (0,5; 2,5),  $a_i$  – параметр, характеризующий структуру знаний испытуемого.

*Трехпараметрическая модель (3PL):*

$$P_j(\theta) = c_j + (1 - c_j) \frac{e^{1,7a_j(\theta - \beta_j)}}{1 + e^{a_j 1,7(\theta - \beta_j)}}, \quad (5)$$

где все обозначения прежние, а  $c_j$  – параметр, характеризующий вероятность правильного ответа на  $j$  задание в том случае, если этот ответ угадан.

В случае политомической шкалы, предусматривающей категорирование оценок за выполнение заданий с конструируемым ответом, в модели вводятся дополнительные параметры. Они получили названия модели с градуированными ответами (*Graded Response Model – GRM*), модели частичного кредитования (*Partial Credit Model – PCM*) и обобщенной модели частичного кредитования (*Generalized Partial Credit Model – GPCM*).

*Модель с градуированными ответами (GRM):*

$$P^*(x_{ij} \geq k | \theta_i, a_j, \beta_{jk}) = \frac{e^{1,7a_j(\theta_i - \beta_{jk})}}{1 + e^{1,7a_j(\theta_i - \beta_{jk})}}, \quad (6)$$

где  $\theta_i$  – параметр подготовленности  $i$ -го испытуемого,  $a_j$  – параметр дифференцирующей способности  $j$ -го задания,  $\beta_{jk}$  – параметр пороговой трудности для ответной категории  $k$  на задание  $j$ .

*Модель частичного кредитования (PCM):*

$$\pi_{ijk} = \frac{e^{\sum_{f=0}^k [\theta_i - (\delta_j + \tau_f)]}}{\sum_{l=0}^m e^{\sum_{f=0}^l [\theta_i - (\delta_j + \tau_f)]}}, \quad (7)$$

где  $\pi_{ijk}$  – вероятность выполнения  $i$ -м испытуемым  $k$ -го шага в задании  $j$ ,  $\theta_i$  – параметр подготовленности  $i$ -го испытуемого,  $\delta_j$  – параметр трудности для  $k$ -го шага в задании  $j$ ,  $\tau_f$  – параметр трудности выполнения отдельного шага.

*Обобщенная модель частичного кредитования (GPCM):*

$$P_{jx} = \frac{e^{\sum_{k=0}^x [a_j(\theta - \delta_{jk})]}}{\sum_{l=0}^m e^{\sum_{k=0}^l [a_j(\theta - \delta_{jk})]}}, \quad (8)$$

где  $m$  – количество возможных категорий,  $x$  – балл испытуемого по заданию,  $j$  – номер задания, остальные обозначения прежние.

**Макет паспорта компетенции****Министерство науки и высшего образования Российской Федерации****ИВАНОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ***(наименование)* факультет (институт)

ОДОБРЕНО

Ученым советом факультета / института

Протокол заседания

№ \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_ г.

**Паспорт компетенции:**наименование компетенции из ФГОС ВО, ПООП или дополнительно определенной  
(код)

Уровень высшего образования:

Квалификация выпускника:

Направление подготовки:

XX.XX.XX Наименование направления подготовки

Направленность (профиль)  
образовательной программы:

## 1. Основные характеристики компетенции и индикаторы достижения компетенции

*(Выбирается и заполняется соответствующая таблица.*

*Детализация индикаторов достижения компетенций 1.1.1, 1.1.2 и т.п. может быть использована при необходимости.*

*Кодирование индикаторов может быть иным в соответствии с ПООП)*

Категория УК	Код и наименование универсальной компетенции	Код и наименование индикатора достижения компетенции
Командная работа и лидерство	УК-3. Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде	УК-3.1. ... УК-3.2. Определяет свою роль в социальном взаимодействии и командной работе, исходя из стратегии сотрудничества для достижения поставленной цели УК-3.3.... ...

Категория ОПК (при наличии)	Код и наименование общепрофессиональной компетенции	Код и наименование индикатора достижения компетенции
Нормативные основания профессиональной деятельности...	ОПК-1. Способен осуществлять и оптимизировать профессиональную деятельность...	ОПК-1.1. Знать... 1.1.1... 1.1.2... ... ОПК-1.2. Уметь... 1.2.1... 1.2.2... ... ОПК-1.3. Иметь практический опыт.../Иметь навыки... 1.3.1... 1.3.2... ...

Тип задач или задача профессиональной деятельности	Код и наименование профессиональной компетенции	Код и наименование индикатора достижения компетенции	Основания включения в перечень планируемых результатов освоения ОП и установления индикаторов достижения ПК
	ПК-1...	ПК-1.1. Знать... 1.1.1... 1.1.2... ПК-1.2. Уметь... 1.2.1..... 1.2.2... ПК-1.3. Иметь практический опыт.../Иметь навыки... 1.3.1... 1.3.2...	<i>(Фиксируется одно или несколько оснований)</i> 1. Является обязательной или рекомендована Примерной образовательной программой 2. На основе анализа профессионального стандарта (наименование, дата утверждения)

			ОТФ (код и наименование) ТФ (код и наименование) 3. Анализ иных требований, предъявляемых к выпускникам на рынке труда 4. Обобщение отечественного и зарубежного опыта 5. Консультации с ведущими работодателями, объединениями работодателей отрасли, в которой востребованы выпускники

## 2. Оценочное(ые) средство(а) для определения уровня сформированности компетенции

**2.1. Оценочное средство:** (например, компетентностно-ориентированное задание или ситуационное задание (кейс) или др.)

*(Далее приводится формулировка задания, в т.ч. указываются требования к его выполнению, необходимые информационные материалы (при наличии))*

### Критерии и шкала оценки:

Индикаторы достижения компетенции (наблюдаемые при выполнении задания действия студента)	Показатели сформированности компетенции (варианты выполнения задания)	Уровень сформированности компетенции
<i>Могут быть указаны коды и(или) наименования индикатора</i>		Компетенция сформирована
		Компетенция сформирована в основном
		Компетенция не сформирована

### Методические указания по организации и процедуре оценивания:

*(Данный раздел включается при необходимости: представляются указания для преподавателей по организации контрольного мероприятия с использованием данного оценочного средства: продолжительность, возможность использования студентами информационных материалов и др., рекомендации по анализу и оценке результатов выполнения студентом, эталон выполнения задания, при наличии.)*

*(Далее представлен пример оценочного средства для УК-3 на основе источника:*

*Кейсы для оценки сформированности универсальных компетенций обучающихся при освоении образовательных программ высшего образования 9 программ бакалавриата, специалитета, магистратуры): методические рекомендации/коллектив авторов под общ. ред. И.Ю. Тархановой. – Ярославль: РИО ЯГПУ, 2018. - 218с.)*

**2.2. Оценочное средство:** ситуационное задание (кейс) «Выбор руководителя».



**Инструкция:** Внимательно прочитайте текст и изучите описание ситуации, ее контекст и дополнительные материалы. После этого приступайте к выполнению заданий кейса.

**Описание ситуации:** Руководителю отдела поручено важное задание по разработке брендбука одной компании. Срок выполнения данного заказа – 1 месяц. Показателями выполнения проекта являются: одобрение заказчиков, представление не менее 3-х вариантов брендбука, отработка вариантов на фокус-группах с обязательным участием представителей заказчика. Неожиданно меняются сроки давно запланированной командировки руководителя, в связи с чем он лично лишается возможности непосредственного участия в реализации проекта. Так и не решившись поручить руководство выполнением проекта одному из своих подчиненных, руководитель отправляется в командировку и контролирует выполнение проекта дистанционно, определяя поручения каждому сотруднику и требуя от всех ежедневного отчета. В результате сроки выполнения проекта сорваны, ни один из планируемых вариантов так и не достиг завершающей стадии, потому не был представлен заказчику.

**Контекст ситуации:** В отделе работает несколько сотрудников. Один из них – опытный сотрудник, с высоким уровнем мотивации, не раз выполнявший подобные задачи, умеет договариваться с людьми, но часто не идет на поиск новых решений и возможностей, полагаясь на собственный опыт. В хороших отношениях с руководителем отдела.

Другой сотрудник, хорошо зарекомендовавший себя в работе, не имеет длительного опыта реализации подобных проектов. Хорошо выполняет поставленные задачи. Иногда может проявить креативность и творчество, но больше закрыт в общении.

Третий сотрудник находится на испытательном сроке, с отличным релевантным образованием, стремится закрепиться в компании и зарекомендовать себя, фонтанирует идеями, понимает новые тренды, импонирует руководителю отдела как сотрудник.

**Дополнительные материалы:** В отделе существуют примерные инструкции для реализации проектов.

Смежные отделы фирмы не сотрудничают друг с другом напрямую, только посредством деловой коммуникации (переписки), личные отношения не смешиваются с рабочими.

В фирме документально закреплена финансовая стимуляция работников за качественную работу, соблюдение показателей и временного отрезка выполнения работ.

**Задание к кейсу:** Проведите анализ сложившейся ситуации. Представьте себя в качестве руководителя отдела и выполните следующие задания:

- 1) сформулируйте главную проблему сложившейся ситуации;
- 2) определите сотрудника, которого бы Вы назначили ответственным на время своего отсутствия;
- 3) сформулируйте не менее 3-х аргументов выбора данного претендента;
- 4) какие функции по руководству проектом Вы делегировали ответственному, какие оставили бы за собой.

### 3.2. Критерии и шкала оценки:

Индикаторы достижения компетенции (наблюдаемые при выполнении задания действия студента)	Показатели сформированности компетенции (варианты выполнения задания)	Уровень сформированности компетенции
УК -3.2. Определяет свою роль в социальном взаимодействии и командной работе, исходя из стратегии сотрудничества для достижения поставленной цели	Выделена основная проблема: некомпетентность руководителя отдела в осуществлении обоснованного выбора ответственного, делегировании полномочий другим сотрудникам.  Выбран конкретный сотрудник в качестве ответственного.	Компетенция сформирована

	<p>Выделено не менее 3-х аргументов выбора, отражающих каждую из следующих позиций: характеристика профессиональных качеств, оценка опыта, оценка возможностей претендентов.</p> <p>Делегированные функции соотносятся с «сильными» сторонами претендента, выделенными при аргументации его выбора</p>	
	<p>Выделены составляющие проблемы: сложности в оценке компетенций претендентов, проблема выбора одного из претендентов, страх перед командной работой в непривычных обстоятельствах и др.</p> <p>Выбран конкретный сотрудник или два сотрудника в качестве кандидатов на роль ответственного.</p> <p>Выделено не менее 3-х аргументов выбора, отражающих не менее 2-х из позиций: характеристика профессиональных качеств, оценка опыта, оценка возможностей претендентов.</p> <p>Делегированные функции соотносятся с «сильными» сторонами претендента, выделенными при аргументации его выбора</p>	<p>Компетенция сформирована в основном</p>
	<p>Предлагаемые студентом формулировки не связаны с проблемой.</p> <p>Выбран конкретный сотрудник или два сотрудника в качестве кандидатов на роль ответственного.</p> <p>Выделено менее 3-х аргументов выбора ответственного, аргументы носят случайный характер.</p> <p>Функции не распределены или делегированные функции не соотносятся с «сильными» сторонами претендента</p>	<p>Компетенция не сформирована</p>

**Методические указания по организации и процедуре оценивания:**

Продолжительность выполнения задания – 30 минут. Дополнительных информационных материалов не требуется. Приоритетным при анализе и оценке выполнения задания студентом

является констатация ожидаемых действий (в соответствии с выделенными показателями) и их обоснованность, а не установление правильности (неправильности) предложенных студентом решений.

Документ составлен в соответствии с требованиями ФГОС ВО по направлению подготовки (*код и наименование*) и (или) (*при наличии*) профессиональных стандартов (*наименование, дата утверждения*).

**Автор(ы) - составители:** (*Указываются должность, ученая степень, ученое звание, ФИО.*)

Документ обновлен

протокол заседания Ученого совета факультета № \_\_\_\_\_ от «\_\_\_\_\_» \_\_\_\_\_ 20\_\_ г.

Декан факультета \_\_\_\_\_  
(подпись)

Руководитель ОП \_\_\_\_\_  
(подпись)

## Различные таксономии для операционализации конструктов

### *Таксономия Б. Блума*

Уровень	Что описывает
<b>Знание</b>	Эта категория обозначает запоминание и воспроизведение изученного материала. Подразумевается, что студент может извлечь из памяти ту или иную информацию: как ее частичные составляющие (факты, определения), так и целостные теории.
<b>Понимание</b>	Понимание – более сложный уровень познания, чем просто запоминание. Показателем понимания изученного может служить способность ученика преобразовать информацию из одной формы в другую. Например, изобразить в пантомиме мы можем только то слово, которое понимаем, то есть мы переводим смысл термина из словесной формы в пантомиму. Также показателем понимания является способность учащегося интерпретировать материал, что называется, объяснять своими словами.
<b>Применение</b>	Эта категория обозначает умение использовать изученный материал в конкретных условиях и в новых ситуациях. Одно дело — знать правила «в теории», другое же — использовать их на практике, в деятельности.
<b>Анализ</b>	Эта категория обозначает умение разбить материал на составляющие части, декомпозировать понятие или объект так, чтобы ясно выступала его структура. Сюда относятся: выявление частей целого, определение взаимосвязей между ними, осознание принципов организации целого. Учебные результаты характеризуются более высоким познавательным уровнем, чем понимание и применение: они требуют не просто понять материал, но и разобраться в его строении и структуре.
<b>Синтез</b>	Эта категория обозначает умение комбинировать элементы так, чтобы получить целое, обладающее новизной. Например, это какой-то продукт, артефакт, доклад, выступление, презентация, иллюстрация — то, что авторски (творчески) упорядочивает имеющийся материал или его части.
<b>Оценка</b>	Здесь подразумевается не просто умение создать новое, но и вынести суждение, оценить созданное или существующее. Эта категория включает способность оценить значимость, аргументировать ее, опираясь на конкретные критерии (как сформулированные самостоятельно, так и заданные учителем). Оценивание

предполагает достижение учебных результатов всех предшествующих категорий.

***Версия таксономии Блума, переосмысленная Кратводем и Андерсоном***

<b>Уровень</b>	<b>Что описывает</b>	<b>Глаголы для операционализации результатов обучения</b>
<b>Создает</b>	Может синтезировать имеющиеся знания для создания чего-то нового.	Создает, планирует и реализует оригинальную деятельность.
<b>Оценивает</b>	Может определить критерии и по ним вынести суждения, оценить	Проверяет, выражает аргументированное мнение (дает обратную связь).
<b>Анализирует</b>	Может декомпозировать материал на составные части, определять взаимосвязи между различными частями материала.	Дифференцирует (определяет отличия и сходства), организует (в категории, структуры), может атрибутировать (соотнести теорию с примерами ее реализации на практике).
<b>Применяет</b>	Понимает, как применить на практике, использует полученное знание в практической ситуации.	Выполняет, применяет на практике.
<b>Понимает</b>	Может конструировать значения из полученного знания, интерпретировать, рассказать своими словами.	Интерпретирует, классифицирует, суммирует, сравнивает, объясняет, может привести пример, может сделать вывод.
<b>Знает</b>	Может распознавать («вытащить из памяти»).	Узнает, вспоминает (припоминает).

***Расширенная таксономия когнитивных способностей Б. Блума***

<b>Уровень</b>	<b>Действия обучающегося</b>	<b>Действия преподавателя</b>	<b>Возможная формулировка задания</b>
<b>Знание</b>	Отвечает Воспринимает Запоминает Распознает Излагает Характеризует Описывает	Руководит Управляет Направляет Рассказывает Сообщает Показывает Проверяет	Сделайте сообщение, перечислите, опишите, воспроизведите, установите (что это, где это), сформулируйте, узнайте, запомните, расскажите, изложите факты, повторите, определите
<b>Понимание</b>	Объясняет Истолковывает	Слушает Спрашивает	Измените, превратите, переформулируйте,

Уровень	Действия обучающегося	Действия преподавателя	Возможная формулировка задания
	Интерпретирует Доказывает Раскрывает Отождествляет	Сравнивает Сопоставляет Опрашивает Проверяет	опишите, объясните, сделайте обзор, расскажите, суммируйте, свяжите воедино, разьясните, сделайте заключение, вывод, изложите основную идею
<b>Применение</b>	Решает новые проблемы Доказывает Отбирает Организует Иницирует Вырабатывает Конструирует	Фасилитирует Оказывает содействие Поддерживает Наблюдает Обсуждает Критикует	Примените, попробуйте на практике, используйте, употребите, решите, докажите, покажите, проиллюстрируйте, сделайте
<b>Анализ</b>	Обсуждает Раскрывает Перечисляет Анализирует Разделяет на части Разбирает критически	Исследует Стимулирует Наблюдает Предоставляет ресурсы	Проанализируйте, разделите на части, разыщите, найдите, определите, различите, проверьте, сравните, классифицируйте, опровергните
<b>Синтез</b>	Обсуждает Обобщает Связывает Сопоставляет Резюмирует Суммирует	Организует обратную связь (рефлексию) Расширяет Оценивает Развивает идею Дискутирует	Создайте, изобретите, предскажите, сконструируйте, оформите, вообразите, улучшите, предложите
<b>Оценка</b>	Судит Оценивает Обсуждает Подвергает сомнению Формирует Составляет Высказывает свое мнение	Принимает Допускает Соглашается Признает Гармонизирует Согласовывает	Составьте мнение, придите к выводу, отберите, выберите, оцените, разберите, обсудите, проверьте, аргументируйте, рекомендуйте, определите, оправдайте, убедите

**SOLO – таксономия**

*Structure of the Observed Learning Outcomes /*

*Структура наблюдаемых результатов обучения*

<b>Уровень расширенных абстракций</b>	Ответ студента включает не только владение терминологическим аппаратом, но также связи между понятиями, теориями, фактами,	Создавать, выдвигать гипотезы
---------------------------------------	--	-------------------------------

	<p>концепциями, он соотносит понятийный аппарат, общается до уровня новой темы или понятия.</p> <p><i>Пример: «У меня есть несколько идей насчет...и я могу связать их в единое целое...и могу посмотреть под другим углом, связать с чем-то еще...»</i></p>	
<b>Уровень отношений</b>	<p>Ответ студента включает упоминание разных ключевых аспектов, а главное связей между ними. Интеграция частей одного понятия.</p> <p><i>Пример: «У меня есть несколько идей насчет... и я могу связать их в единое целое»</i></p>	<p>Анализировать, применять, сравнивать, объяснять причинно-следственные связи, критиковать, подтверждать</p>
<b>Мульти-структурный</b>	<p>Ответ студента затрагивает несколько значимых аспектов, но они рассматриваются изолированно, связи не установлены. Поверхностное понимание.</p> <p><i>Пример: «У меня есть несколько идей насчет...»</i></p>	<p>Описывать, сочетать, перечислять</p>
<b>Одноструктурный</b>	<p>Ответ студента очень простой и касается только одного значимого аспекта. Поверхностное понимание.</p> <p><i>Пример: «У меня есть одна идея насчет...»</i></p>	<p>Определять, называть</p>
<b>Преструктурный</b>	<p>Студент не может ничего ответить, так как ничего не понял.</p> <p><i>Пример: «Я не уверен насчет...»</i></p>	

### Пример структуры спецификации измерителя для аттестационного испытания

1. Цель создания измерителя. Обоснование подхода к его созданию.
2. Целевая аудитория, для которой разрабатывается измеритель и её характеристики.
3. Документы, определяющие содержание измерителя.
4. Перечень индикаторов компетенций, которые измеряются в рамках аттестационного испытания.

Таблица 1

#### Компетенции, подлежащие измерению в рамках аттестационного испытания, и их индикаторы

Компетенция	Степень сформированности	Индикаторы – основные поведенческие признаки сформированности компетенции	Действие, в котором проявляется индикатор

5. Описание измерителя: общее количество заданий, тип заданий с указанием количества заданий каждой формы и номера заданий.

Таблица 2

#### Распределение заданий по типам и индикаторам

Индикаторы	Число заданий	Тип задания				
		МВО <sup>5</sup>	МВН <sup>6</sup>	УС <sup>7</sup>	УП <sup>8</sup>	ОФ <sup>9</sup>
Всего						

6. Рекомендуемая стратегия расположения заданий (композиция измерителя).

<sup>5</sup> МВО – задания множественного выбора с одним верным ответом.

<sup>6</sup> МВН – задания множественного выбора с несколькими верными ответами.

<sup>7</sup> УС – задания на установление соответствия.

<sup>8</sup> УП – задания на установление последовательности.

<sup>9</sup> ОФ – задания открытой формы.



**7. Рекомендуемое время выполнения***Таблица 3***Время выполнения заданий и теста в целом**

Тип задания	Среднее время выполнения одного задания, мин.	Число заданий данного типа, шт.	Время выполнения заданий данного типа, мин

Время на инструктаж / чтение инструкций:

Общее время выполнение с учетом инструктажа:

**8. Рекомендации по оцениванию заданий (дихотомическая или политомическая оценка) и измерителя в целом.**

Максимальный итоговый балл:

Пороговые баллы (при наличии):

*Таблица 4***Оценивание заданий**

№ задания	Тип задания	Балл за задание

**9. Обобщенный план измерителя.***Таблица 5*

№ задания	Компетенции	Индикаторы	Действие, в котором проявляется индикатор	Степень сформированности	Тип задания	Время выполнения задания (мин.)	Макс. балл за выполнение задания

**10. Предоставление результатов измерений аттестуемым.****11. Условия использования и распространения измерителя и оценочных заданий.**

## Требования к заданиям разного типа

### *1. Общие требования*

1.1. Задания должны быть направлены на оценку значимого содержания, а не того содержания, для которого проще и легче формулируются задания.

1.2. Текст должен исключать всякую двусмысленность и неясность формулировок.

1.3. Текст задания формулируется предельно кратко, т. е. освобождается от всякого постороннего для данной проблемы материала. Текст задания должен иметь предельно простую синтаксическую конструкцию.

1.4. В задании не используются слова, вызывающее различное понимание у испытуемых, а также слова, являющиеся подсказкой, например, «иногда», «часто», «всегда», «все», «никогда».

1.5. Если в задании базового уровня изложена ситуация или обстоятельства, то они должны быть знакомы испытуемым.

### *2. Дополнительные рекомендации к качеству ответов*

#### *для заданий с выбором ответа*

2.1. Формулировка основной части задания, при соблюдении условий, изложенных в п. 1.3., должна быть законченной, т. е. испытуемый должен из ее содержания понять, какую задачу ему предстоит выполнить до анализа предложенных вариантов ответа.

2.2. В заданиях базового уровня в ответы целесообразно включать не более двух- трех важных, ключевых слов, при этом правильный ответ не должен отличаться какими-либо формальными признаками от неправильных ответов (дистракторов).

2.3. В заданиях базового уровня не должно требоваться выбрать один неправильный ответ при наличии нескольких правильных.

2.4. Из текста задания исключаются все вербальные ассоциации, способствующие выбору правильного ответа с помощью догадки.

2.5. Задания должны иметь дистракторы, одинаково правдоподобные и привлекательные для выбора.

2.6. Исключаются ответы, вытекающие один из другого или дополняющие друг друга.

2.7. При формулировке дистракторов не следует использовать выражения «ни один из перечисленных», «все перечисленные» и т. д., так как они способствуют угадыванию правильного ответа.

2.8. Из ответов, как правило, исключаются повторяющиеся слова путем ввода их в основной текст задания.

2.9. Все ответы, по возможности, должны быть приблизительно одной длины.

2.10. Все ответы должны быть грамматически согласованными с основной частью задания.

2.11. Если ответ выражен в виде числа, то, как правило, эти числа располагаются от меньшего к большему или наоборот. При этом если в ответе имеются числа 1, 2, 3, 4, то они должны стоять под соответствующими номерами.

2.12. Все ответы должны быть перечислены либо строго друг под другом, либо в одну строку, чтобы испытуемые не тратили времени на ориентировку в разнообразии возможных способов пространственного размещения ответов. Исключение могут составлять лишь ответы в виде рисунков.

### *3. Дополнительные рекомендации к заданиям на установление последовательности*

3.1. Задание начинается со слов «Установите последовательность ...».

3.2. В условии перечисляются под буквенными обозначениями (следующими в алфавитном порядке) все элементы.

3.3. Формулируется критерий упорядочивания.

3.4. Для всех предметов используется одинаковый формат ответов, представленный ниже.

При оформлении ответа испытуемый вначале заносит ответ в таблицу, а потом переносит в бланк ответов. Используется следующая инструкция при оформлении задания: «Запишите в таблицу выбранные буквы в правильной последовательности, а затем перенесите их в бланк ответов (без пробелов и других символов)».

#### *4. Дополнительные рекомендации к заданиям на установление соответствия*

4.1. Задания на установление однозначного или множественного соответствия начинаются со слов «Установите соответствие...».

4.2. Задание формулируется так, чтобы все содержание можно было выразить в виде двух множеств с соответствующими названиями.

4.3. Элементы первого столбца обозначаются цифрами и располагаются слева, а элементы второго столбца обозначаются буквами русского алфавита и располагаются справа. При этом желательно, чтобы количество элементов в первом и втором столбцах не превышало 7 в каждом (1234567 – слева и АБВГДЕЖ – справа).

4.4. Для каждого столбца вводится определенное название, обобщающее все элементы столбца. Название столбца записывается заглавными буквами.

4.5. Элементы столбцов должны быть выбраны по одному основанию.

4.6. Для заданий на однозначное соответствие необходимо, чтобы второй столбец содержал не менее одного дистрактора.

4.7. В заданиях на однозначное соответствие каждому элементу первого столбца ставится в соответствие ровно один элемент второго. При этом один элемент второго столбца может соответствовать не более чем одному элементу первого столбца.

4.8. В заданиях на множественное соответствие каждому элементу первого столбца ставится в соответствие не менее одного элемента второго. При этом один элемент второго столбца может соответствовать не более чем одному элементу первого столбца.

### *5. Дополнительные рекомендации к заданиям*

#### *с краткими ответами*

5.1. Каждое задание с кратким ответом на дополнение должно быть нацелено только на одно дополнение, место для которого обозначается подчеркиком длиной не менее трех символов или многоточием. Вид подчеркика – «\_\_\_\_\_».

5.2. Подчерк ставится на месте ключевого элемента, знание которого является наиболее существенным для контролируемого материала.

5.3. Все подчеркики в открытых заданиях для одного теста должны быть равной длины.

5.4. Дополнения лучше ставить в конце задания или как можно ближе к концу.

5.5. После подчеркика, если это необходимо, указываются единицы измерения.

5.6. Ответы могут записываться ТОЛЬКО в виде сочетания:

- 1) символов из набора «0123456789» и символа «-» (дефис);
- 2) символов кириллицы и символа «-» (дефис);
- 3) символов латиницы и символа «-» (дефис).

5.7. Использование символов «.» (точка), «/» (косая черта) в ответах не допускается.

5.8. В эталонах ответа указываются все возможные варианты верного ответа, разделенные символами <или>, но не более пяти. Указывается порядок следования слов, если ответ состоит из нескольких слов.

5.9. Желательно использовать не более 17 символов для одного варианта верного ответа.

### *6. Дополнительные рекомендации к заданиям*

#### *со свободными развернутыми ответами*

6.1. Задания с развернутыми ответами сопровождаются системой оценивания его выполнения, которая должна включать критерии оценивания, вариант

(варианты, образцы) правильных ответов (элементов ответа) и, по ряду дисциплин, решения.

6.2. Формулировка задания должна соответствовать критериям оценивания. Это означает, что испытуемый после прочтения задания должен понять, какую задачу ему предстоит выполнить, и с какой полнотой он должен дать ответ для получения максимального балла. Например, сколько привести аргументов, фактов или примеров, нужно ли привести чертеж или диаграмму, нужно ли привести полное решение с пояснениями.

6.3. Основой для определения системы кодировки и выставления баллов за выполнение задания должен быть анализ из спектра ответов испытуемых (на выборке апробации) и соотнесение данного спектра ответов с экспертными критериями оценки.

6.4. В заданиях со свободным ответом, проверяющих предметные знания, умения, а не коммуникативные умения и навыки, от испытуемого не должно требоваться написания длинного текста. В задании должны быть даны рекомендации о предполагаемой длине ответа (возможное число предложений, определенная часть страницы и т. д.).

6.5. Максимальная оценка эксперта для одного простого задания по одной категории (одного аспекта оценивания) не должна превышать 5. Желательно, чтобы он равен либо 2, либо 3.

## Приложение 6

### Конструктор заданий

1. Ознакомление	2. Понимание	3. Применение	4. Анализ	5. Оценка	6. Создание
1. Назовите основные части...	8. Объясните причины того, что...	15. Изобразите информацию о... графически	22. Раскройте особенности...	29. Ранжируйте ... и обоснуйте...	36. Предложите новый (иной) вариант...
2. Сгруппируйте вместе все...	9. Обрисуйте в общих чертах шаги, необходимые для того, чтобы...	16. Предложите способ, позволяющий...	23. Проанализируйте структуру ... с точки зрения...	30. Определите, какое из решений является оптимальным для... (на ваш взгляд), какие еще существуют решения...	37. Разработайте план, позволяющий (препятствующий)...
3. Составьте список понятий, касающихся...	10. Покажите связи, которые направлены на..	17. Сделайте эскиз рисунка (схемы), который показывает...	24. Составьте перечень основных свойств ..., характеризующих ... с точки зрения...	31. Оцените значимость ... для...	38. Найдите необычный способ, позволяющий...
4. Расположите в определенном порядке...	11. Постройте прогноз развития...	18. Сравните ... и ..., а затем обоснуйте...	25. Постройте классификацию ... на основании...	32. Определите возможные критерии оценки...	39. Придумайте игру, которая...
5. Изложите в форме текста...	12. Прокомментируйте положение о том, что...	19. Проведите (разработайте) эксперимент, подтверждающий, что...	26. Найдите в тексте (модели, схеме и т. п.) то, что...	33. Выскажите критические суждения о...	40. Предложите новую (свою) классификацию...
6. Вспомните и напишите...	13. Изложите иначе (переформулируйте) идею о том, что...	20. Проведите презентацию...	27. Сравните точки зрения ... и ... на...	34. Оцените возможности ... для...	41. Напишите возможный (наиболее вероятный) сценарий развития...
7. Прочитайте самостоятельно..	14. Приведите пример, что (как, где)...	21. Рассчитайте на основании данных о...	28. Выявите принципы, лежащие в основе...	35. Проведите экспертизу состояния...	42. Изложите в форме ... свое мнение(понимание)...

**Примеры оценочных заданий для различных стадий оценивания  
(направление подготовки 44.03.02 Психолого-педагогическое  
образование, образовательная программа «Психология образования»)**

*Первая стадия оценивания – количественный уровень образовательных измерений, задания с выбором одного правильного ответа*

*Дисциплина обязательной части образовательной программы «Педагога и педагогическая психология».*

*Справочно (из рабочей программы дисциплины). Основная цель состоит в создании условий для освоения студентами понимания педагогики как науки, педагогической психологии как науки, а также как составной части общей культуры человека, основы для конструирования педагогической практики, стимулировании интереса студентов к изучению предметов психолого-педагогического цикла, формированию основных общекультурных и общепрофессиональных компетенций, предусмотренных моделью выпускника бакалавриата по данному направлению подготовки. Место дисциплины в структуре образовательной программы: Успешное освоение данной дисциплины будет способствовать формированию готовности студентов к осуществлению компетентной психолого-педагогической деятельности. Учебная дисциплина направлена на формирование у студентов педагогической культуры, современного педагогического менталитета, системных знаний в области педагогики и педагогической психологии, формирование профессионально-понятийного языка, овладения комплексом умений применять психолого-педагогическое знание в практической деятельности педагога-психолога. При изучении курса студенты опираются на знания, умения и опыт, полученные на учебной дисциплине «Введение в психолого-педагогическую деятельность», «Психология», учебная, ознакомительная практика. Данная учебная дисциплина является базовой для изучения всех последующих дисциплин, конкретизирующих основные разделы данного учебного курса: «Методология и методы психолого-педагогических исследований», «Теория и методика воспитательной деятельности», «Психолого-педагогическое сопровождение учебной деятельности», «Образовательные технологии», «Педагогика и психология профессионального образования и обучения», служит теоретической базой для различных видов практики.*

**1. Кто из ученых является основателем культурно-исторического подхода (концепции) в психологии?**

- 1) Давыдов В.В.
- 2) Гальперин П.Я.
- 3) Зимняя И.А.
- 4) Выготский Л.С.

**2. Что составляет предмет изучения педагогической психологии?**

- 1) законы и закономерности образовательного и педагогического процессов
- 2) законы и закономерности развития личности детей разного возраста
- 3) закономерности интеллектуального и личностного развития ребенка как субъекта учебной деятельности под руководством педагога
- 4) закономерности развития психических процессов у детей разного возраста



**3. Кто из ученых разработал понятие «педагогическая система» и определил ее компоненты?**

- 1) Рубинштейн С.Л.
- 2) Выготский Л.С.
- 3) Кузьмина Н.В.
- 4) Гальперин П.Я.

**4. В чем сущность понятия «обучаемость»?**

- 1) это наличный уровень знаний, умений и навыков ученика
- 2) это восприимчивость к обучению, готовность учиться
- 3) это комплекс предметных умений по усвоению определенной области знаний
- 4) это самостоятельное освоение знаний

**5. В чем сущность понятия «обученность»?**

- 1) это реально усвоенный уровень знаний, умений и навыков ученика
- 2) это восприимчивость к обучению, готовность учиться
- 3) это комплекс предметных умений по усвоению определенной области знаний
- 4) это самостоятельное освоение знаний

**6. В каком возрасте учебная деятельность является ведущим видом деятельности?**

- 1) в подростковом
- 2) в старшем школьном возрасте
- 3) в младшем школьном возрасте (учащиеся начальных классов)
- 4) у дошкольников

**7. Как называется система развивающего обучения, приоритетной целью которой является общее развитие учащихся начальной школы: интеллектуальное, эмоциональное, волевое и т.д. ?**

- 1) Система Д.Б. Эльконина – В.В. Давыдова
- 2) Система Л.В. Занкова
- 3) Система Л.Г. Петерсон
- 4) Система «Школа диалога культур»

**8. В какой системе развивающего обучения центральными идеями являются: развитие учащихся как субъектов учебной деятельности и приоритетное развитие теоретического мышления?**

- 1) Система Д.Б. Эльконина – В.В. Давыдова
- 2) Система Л.В. Занкова
- 3) Система Л.Г. Петерсон
- 4) Система «Школа диалога культур»

**9. В какой теории обучения раскрыт процесс формирования умственных действий учащихся?**

- 1) Теория бихевиоризма
- 2) Ассоциативно-рефлекторная теория
- 3) Гештальт-психология
- 4) Теория обучения П.Я Гальперина

**10. Кем из ученых было сформулировано положение о ведущих видах деятельности учащихся?**

- 1) Выготским Л.С.
- 2) Занковым Л.В..
- 3) Леонтьевым А.А.
- 4) Божович Л. И.

**11. Учащийся как субъект учебной деятельности характеризуется**

- 1) самооценкой, самопознанием, самосовершенствованием
- 2) осознанностью, активностью, самостоятельностью
- 3) обученностью, воспитанностью, обучаемостью
- 4) честностью, принципиальностью, ответственностью

**12. Каковы компоненты учебной деятельности?**

- 1) когнитивный, творческий, оценочный
- 2) контрольный, проверочный, оценивающий
- 3) мотивационный, операционный, аналитический
- 4) целевой, содержательный, результативный

**13. Каковы основные характеристики педагогического сотрудничества?**

- 1) контроль, оценка, проверка
  - 2) взаимодействие, помощь, контроль
  - 3) взаимопроверка, оценка, отметка
  - 4) взаимопонимание, взаимодействие, взаимоотношение
14. Какие взаимоотношения характерны для педагогики сотрудничества?
- 1) отношения на основе доверия и ответственности
  - 2) отношения на основе контроля и оценки
  - 3) отношения на основе проверки и требований
  - 4) отношения на основе требований и контроля за их выполнением

**15. Ведущим принципом отечественной педагогической психологии является**

- 1) принцип социального моделирования
- 2) принцип трансформации знаний, их расширение и приспособление к решению новых задач
- 3) принцип единства сознания и деятельности
- 4) принцип установления связи между стимулами и реакциями

**16. Выготский Л.С. рассматривая проблему соотношения обучения и развития:**

- 1) отождествляет процессы обучения и развития
- 2) полагает, что обучение должно опираться на зону актуального развития ребенка
- 3) полагает, что обучение должно забегать вперед развития и вести его за собой
- 4) полагает, что при обучении надо опираться на уже достигнутый уровень развития

**17. Согласно теории поэтапного формирования умственных действий, организация процесса обучения в первую очередь должна опираться на:**

- 1) материальное действие
- 2) создание ориентировочной основы действия
- 3) речевую форму выполнения действия
- 4) внутреннюю речь

**18. Основной функцией педагогической оценки является:**

- 1) определение уровня фактического исполнения учебного действия
- 2) осуществление подкрепления в виде наказания-поощрения
- 3) развитие мотивационной сферы учащегося
- 4) выполнение требований педагога

**19. Уровень (зону) актуального развития характеризует:**

- 1) обученность, воспитанность, развитость
- 2) обучаемость, воспитуемость, развиваемость
- 3) самообучаемость, саморазвиваемость, самовоспитуемость
- 4) обученность, обучаемость.

**20. Деятельность человека направлена на изменение своей личности в соответствии с сознательно поставленными целями, сложившимися идеалами и убеждениями – это...:**

- 1) воспитание
- 2) педагогические закономерности воспитания
- 3) самовоспитание
- 4) самообразование.

**21. Кем был предложен термин « педагогическая психология»?**

- 1) Лазурским А.Ф.
- 2) Каптеревым П.Ф.
- 3) Ушинским
- 4) Макаренко А.С.

**22. Отметкой в дидактике называют**

- 1) количественный показатель оценки учебных достижений учащихся
- 2) обеспечение обратной связи с учащимися
- 3) качественный показатель уровня и глубины знаний учащихся
- 4) метод устного контроля

**23. Основаниями для внутренней дифференциации учащихся не является**

- 1) материальная обеспеченность школы
- 2) способности учащихся
- 3) психологические особенности детей
- 4) интересы детей

**24. Что является ведущей деятельностью в подростковом возрасте?**

- 1) учебная деятельность
- 2) личностно-значимое общение
- 3) трудовая деятельность
- 4) деятельность по самообслуживанию

**25. В каком документе не раскрываются нормативно-правовые основы образования в РФ?**

- 1) Конституция Российской Федерации
- 2) 273-ФЗ «Об образовании в Российской Федерации»
- 3) ФГОС СОО
- 4) календарно-тематическое планирование учителя

**26. Кто из педагогов прошлого первым сформулировал теоретические основы дидактики (теории обучения) как науки?**

- 1) Руссо Ж.Ж.
- 2) Дистервег А.
- 3) Коменский Я.А.
- 4) Песталоцци И.Г.

**27. Укажите две взаимосвязанные стороны процесса обучения.**

- 1) обучение и развитие
- 2) преподавание и учение
- 3) умения и навыки
- 4) образование и учение

**28. Авторами какой теории содержания образования являются отечественные дидакты М.Н.Скаткин и И.Я. Лернер?**

- 1) теория материального образования
- 2) теория формального образования

- 3) системно-культурологическая теория
- 4) педоцентрическая теория

**29. На основании какого документа проектируется содержание образования в 5-9 классах современной школы?**

- 1) ФКГОС (БУП 2004)
- 2) ФГОС ООО
- 3) Учебников по предмету
- 4) Методических рекомендаций учителю

**30. На основе какого документа составляется учителем календарно - тематический план по предмету?**

- 1) на основе учебного плана школы
- 2) на основе программы по учебному предмету
- 3) на основе расписания уроков
- 4) на основе стандарта по предмету

**31. Что такое обучение?**

- 1) это процесс передачи знаний от учителя к учащимся
- 2) это усвоение учебного материала учащимися, их самостоятельная работа
- 3) это целенаправленный, управляемый и контролируемый двусторонний процесс, направленный на достижение результатов обучения.
- 4) это руководство учителем учебной деятельностью учащихся

**32. Укажите основные звенья (этапы) процесса обучения:**

- 1) познание, обучение, преподавание, учение, воспитание;
- 2) образование, обучение, воспитание, развитие, коррекция, социализация;
- 3) живое созерцание, абстрактное мышление, практика
- 4) целеполагание, восприятие, осмысление, закрепление, применение, анализ результатов, коррекция.

**33. Каковы уровни общего образования в соответствии с 273 ФЗ??**

- 1) начальное, основное, дошкольное
- 2) начальное, основное, среднее полное образование
- 3) дошкольное, начальное, основное, среднее общее образование
- 4) дошкольное, основное, полное среднее образование

**34. В чем сущность преподавания в контексте личностно-деятельностного подхода?**

- 1) в сообщении учащимся знаний
- 2) в установлении дисциплины на уроке
- 3) в психолого-педагогической поддержке учебной деятельности учащихся
- 4) в создании условий для развития личности и формировании учебной деятельности учащегося средствами содержания образования

**35. Укажите методы обучения, в основе классификации которых лежит характер познавательной деятельности учащихся по усвоению различных элементов содержания образования:**

- 1) словесные, наглядные, практические
- 2) индуктивные и дедуктивные
- 3) объяснительно-иллюстративные, репродуктивные, проблемные, исследовательские
- 4) методы стимулирования и мотивации, организации и осуществления учебно-познавательной деятельности, контроля и самоконтроля

**36. Какие государственные документы являются основанием для проектирования содержания образования в конкретном образовательном учреждении?**

- 1) приказы, инструкции, методические материалы
- 2) федеральные государственные образовательные стандарты
- 3) учебные планы образовательных учреждений
- 4) указания министерства образования и науки РФ, методические письма.

**37. К какой группе методов можно отнести тестирование, устный опрос, контрольную работу?**

- 1) к методам организации и осуществления учебно-познавательной деятельности
- 2) к методам стимулирования и мотивации учебно-познавательной деятельности
- 3) к методам контроля и самоконтроля
- 4) к объяснительно-иллюстративным

**38. Кто первым дал теоретическое основание классно-урочной системы?**

- 1) Песталоцци И.Г.
- 2) Дистервег А.
- 3) Коменский Я.А.
- 4) Ушинский К.Д.

**39. Учение – это (выберите наиболее полное определение):**

- 1) деятельность преподавателя по передаче знаний
- 2) деятельность учащихся по усвоению основ морали
- 3) взаимосвязанная деятельность учителя и ученика по усвоению содержания образования
- 4) деятельность учащегося по усвоению содержания образования

**40. Каков главный критерий отбора содержания образования в теории материального образования?**

- 1) соответствие интересам учащихся
- 2) степень влияния на совершенствование форм мышления, воображения, памяти
- 3) степень пригодности для жизни, для практической деятельности
- 4) отражение разных сторон социально-культурного опыта

**41. Какую сторону процесса обучения представляет собой педагогическое руководство учебной деятельностью учащихся?**

- 1) учение
- 2) преподавание
- 3) обучение
- 4) образование

**42. Учебный план образовательной организации – это:**

- 1) государственный документ, в котором определено, какие учебные предметы и в каком количестве часов в неделю должны изучаться в школе
- 2) документ образовательной организации, в котором дано распределение образовательных областей и учебных предметов по годам обучения, их общая трудоемкость, количество часов на вариативную часть содержания образования и т.д.
- 3) государственный документ, определяющий порядок изучения учебного материала по конкретному учебному предмету
- 4) документ, отражающий порядок изучения учебных предметов по годам обучения в конкретной школе

**43. В какой теории содержания образования отражена идея, что сам учащийся должен определять «как качество, так и количество обучения»?**

- 1) теория материального образования
- 2) теория формального образования
- 3) системно-культурологическая теория
- 4) педоцентрическая теория

**44. Что лежит в основе следующей классификации уроков: уроки изучения нового материала, уроки закрепления, уроки повторения и обобщения, уроки контроля и т.д.?**

- 1) источник знания
- 2) дидактическая задача
- 3) характер познавательной деятельности учащихся
- 4) особенности процесса обучения

**45. Образование - это**

- 1) единый процесс обучения и развития
- 2) единый процесс обучения и воспитания
- 3) единый процесс воспитания и развития
- 4) единый процесс формирования и развития

**46. Что такое методы обучения?**

- 1) те исходные позиции, которыми руководствуется учитель в своей работе
- 2) способы передачи учителем знаний, умений и навыков
- 3) основные идеи, на которых проектируется и организуется процесс обучения
- 4) способы взаимосвязанной деятельности учителя и учащихся, при помощи которых достигается решение основных задач обучения

**47. Как соотносятся понятия «отметка» и «оценка»?**

- 1) отметка шире, чем оценка
- 2) оценка более широкое понятие, чем отметка
- 3) это антонимы
- 4) это синонимы

**48. Укажите классификацию методов обучения по компонентам учебной деятельности учащихся:**

- 1) словесные, наглядные, практические
- 2) методы стимулирования и мотивации, организации и осуществления, контроля и самоконтроля
- 3) объяснительно-иллюстративные, репродуктивные, проблемные, частично-поисковые, исследовательские
- 4) методы первичного освоения материала, закрепления, выработки умений и навыков, проверки и оценки

**49. В какой теории содержания образования главное внимание уделяется развитию умственных способностей учащихся?**

- 1) теория материального образования
- 2) теория формального образования
- 3) системно-культурологическая теория
- 4) педоцентрическая теория

**50. Какие виды результатов отражены во ФГОС ОО?**

- 1) предметные, образовательные, развивающие;
- 2) предметные, метапредметные, личностные
- 3) образовательные, личностные, дидактические;

4) развивающие, воспитательные, образовательные

**Вторая стадия оценивания – количественный уровень образовательных измерений, профессионально-ориентированные (ситуационные) задания**

### **Задача 1**

**Мальчик Саша 8 лет постоянно дерется с одноклассниками. Все попытки разговора классного руководителя с родителями заканчиваются словами: «Наш сын дома хорошо себя ведет. Это все ваши наговоры и выдумки. В ваши задачи входит воспитывать детей, вот и воспитывайте».**

#### Вопросы:

1. Чем, по Вашему мнению, можно объяснить поведение родителей?
2. Оцените роль взаимодействия родителей, классного руководителя, учителей-предметников и администрации образовательной организации в воспитании ребенка.
3. Покажите, как можно продуктивно решить ситуацию.

### **Задача 2.**

**На консультацию пришли родители подростка. Сыну 16 лет. В интерпретации он ведет себя странно: стал нервным, скрытным. Часто у него «блестят глаза». Ухудшились успеваемость и поведение.**

#### Вопросы:

1. Сформулируйте возникшие у вас гипотезы о возможных причинах трудностей родителей.
2. По каким признакам родители могут определить, что их ребенок стал употреблять спиртное (другие психоактивные вещества)?
3. Определите основные этапы работы над проблемой.

Перед вами - ряд затруднительных педагогических ситуаций. Познакомившись с содержанием каждой из них, Вам надо выбрать **два варианта** из числа предложенных вариантов реагирования на данную ситуацию, но таких, которые с психолого-педагогической точки зрения наиболее правильны, по Вашему мнению.

**С 1. Ситуация. «Вы приступили к проведению урока, все учащиеся успокоились, настала тишина, и вдруг в классе кто-то громко засмеялся. Когда вы, не успев ничего сказать, вопросительно и удивленно посмотрели на учащегося, который засмеялся, он, смотря вам прямо в глаза, заявил: «Мне всегда смешно смотреть на вас, и хочется смеяться, когда вы начинаете вести занятия».**

Как вы отреагируете на это?

1. «Вот тебе и на!»
2. «А что тебе смешно?»
3. «Ну, и ради бога!»
4. «Ты что, дурачок?»
5. «Люблю веселых людей».

6. «Я рад (а), что создаю у тебя веселое настроение».

**С 2. Ситуация.** *«В самом начале занятия или уже после того, как вы провели несколько занятий, учащийся заявляет вам: «Я не думаю, что вы, как педагог, сможете нас чему-то научить»».*

Какой должна быть реакция учителя?

1. «Твое дело - учиться, а не учить учителя».
2. «Таких, как ты, я, конечно, ничему не смогу научить».
3. «Может быть, тебе лучше перейти в другой класс или учиться у другого учителя?»
4. «Тебе просто не хочется учиться».
5. «Мне интересно знать, почему ты так думаешь».
6. «Давай поговорим об этом подробнее. В моем поведении, наверное, есть что-то такое, что наводит тебя на подобную мысль».

**Третья стадия оценивания - качественный уровень образовательных измерений, структурированное собеседование по заданной теме / проблеме**

#### **Задача**

К психологу обратилась руководитель танцевального кружка и рассказала про девочку Дашу. Она сообщила, что девочка (6 лет) с трех лет посещает хореографический кружок, с 5-ти – вокальную и театральную студии. Очень часто выступает на сцене, участвует в разных конкурсах. И вот перед очередным выступлением театрального кружка Даша пыталась руководить своими сверстницами: «Я лучше вас знаю, я на сцене уже много раз выступала, а вы нет. Поэтому я буду играть роль лисы». Девочки пытаются ей не подчиниться, и между ними разгорелась ссора.

*[Проанализируйте сложившуюся ситуацию. Какую ошибку совершила девочка Даша при выборе модели общения со сверстницами? Какие могут быть последствия, если оставить ситуацию как есть? Какие действия руководителя танцевального кружка вы считаете оптимальными (какие рекомендации Вы дадите руководителю кружка при его обращении к вам)?]*

**1. Оценка ситуации:**

**2. Прогнозирование результатов:**

**3. Варианты решения:**

**4. Выбор и обоснование решения:**

#### **Опорные материалы.**

*Одаренный ребенок* - это ребенок, который выделяется яркими, очевидными, иногда выдающимися достижениями (или имеет внутренние предпосылки для таких достижений) в том или ином виде деятельности. На сегодняшний день большинство



психологов признает, что уровень, качественное своеобразие и характер развития одаренности - это всегда результат сложного взаимодействия наследственности (природных задатков) и социальной среды, опосредованного деятельностью ребенка (игровой, учебной, трудовой). При этом особое значение имеют собственная активность ребенка, а также психологические механизмы саморазвития личности, лежащие в основе формирования и реализации индивидуального дарования.

**Одаренным и талантливым детям свойственны следующие черты:**

- сильно развитое чувство справедливости, проявляющееся очень рано. Личные системы ценностей у одаренных детей очень широки.
- остро воспринимают общественную несправедливость. Устанавливают высокие требования к себе и к окружающим и живо откликаются на правду, справедливость, гармонию и природу.
- остро воспринимают общественную несправедливость. Устанавливают высокие требования к себе и к окружающим и живо откликаются на правду, справедливость, гармонию и природу, и т.д.

Обычно одаренные дети проявляют повышенную требовательность к себе и другим, нетерпимость к нарушителям собственных канонов. Такие особенности поведения и сознания позволяют одаренному ребенку отказаться от кумиров и авторитетов, что, с одной стороны, затрудняет процесс обучения, с другой - помогает сформировать свой собственный стиль. На ранних стадиях работы с одаренными детьми можно наблюдать и другую неприятную особенность - поверхностность знаний. Это объясняется множественностью интересов ребенка, его желанием заниматься всем, к чему возникает интерес.

Следует сказать и о том, что одаренные дети доставляют неудобства не только другим, но, зачастую, и себе самим. Наиболее ярко это проявляется в общении, то есть возникают проблемы межличностной коммуникации одаренных детей. Беря на себя роль организатора, руководителя в раннем возрасте, они, тем самым, вызывают недовольство со стороны остальных участников общения или игры. Это недовольство тем сильнее, чем меньше понимание неординарности человека, взявшего власть в свои руки. Позже одаренные дети бывают склонны к командованию, управлению другими, становятся более жесткими и нетерпимыми. Этот факт можно рассматривать с различных точек зрения: если талантливый ребенок приложит максимум усилий к привлечению внимания к своей личности, то он будет иметь высокий авторитет и уважение группы, в которой развивается; и, напротив, неостребованные управленческие таланты приводят к тому, что человек отвергается коллективом. В первом случае создается благоприятная психологическая обстановка для дальнейшего развития личности, во втором - конфликты могут привести к полной потере интереса к дальнейшему развитию.

### Расчеты показателей надежности по различным методикам

Расчет коэффициентов надежности по двум формулам (7) и (8), представленным в параграфе 3.3 главы 3.

На основании двухкратного применения оценочного инструментария (теста) составляется таблица сопряженности

		Измерение 1	
		Не аттестован	Аттестован
Измерение 2	Не аттестован	0,2	0,02
	Аттестован	0,03	0,7

Вычислим коэффициент надежности  $\varphi$  по формуле (7):

$$\begin{aligned} \varphi &= \frac{P_{(--)} \cdot P_{(++)} - P_{(+-)} \cdot P_{(-+)}}{\sqrt{(P_{(--)} + P_{(-+)}) \cdot (P_{(+-)} + P_{(++)}) \cdot (P_{(--)} + P_{(+-)}) \cdot (P_{(-+) + P_{(++)})}} \\ &= \frac{0,2 * 0,7 - 0,03 * 0,02}{\sqrt{(0,2 + 0,03) \cdot (0,02 + 0,7) \cdot (0,2 + 0,02) \cdot (0,03 + 0,7)}} \\ &= \frac{0,14 - 0,0006}{\sqrt{(0,23) \cdot (0,72) \cdot (0,22) \cdot (0,73)}} = 0,875 \end{aligned}$$

Рассчитаем капшу Коэна ( $\kappa$ -коэффициент согласованности) по формуле (8):

$$\begin{aligned} \kappa &= \frac{(P_{(--)} + P_{(++)}) - (P_{(++) + P_{(\mp)}}) \cdot (P_{(++) + P_{(\pm)}}) - (P_{(--)} + P_{(\pm)}) \cdot (P_{(--)} + P_{(\mp)})}{1 - (P_{(++) + P_{(\mp)}}) \cdot (P_{(++) + P_{(\pm)}}) - (P_{(--)} + P_{(\pm)}) \cdot (P_{(--)} + P_{(\mp)})} \\ &= \frac{(0,2 + 0,7) - (0,7 + 0,03) \cdot (0,7 + 0,02) - (0,2 + 0,02) \cdot (0,2 + 0,03)}{1 - (0,7 + 0,03) \cdot (0,2 + 0,02) - (0,2 + 0,02) \cdot (0,2 + 0,03)} \\ &= \frac{0,3238}{0,4238} = 0,76 \end{aligned}$$

Полученные значения двух коэффициентов оказались достаточно близки к другу другу и к значению 0,8, что свидетельствует об удовлетворительной надежности полученных результатов с помощью используемого оценочного инструментария.

## Практические рекомендации по математико-статистической обработке и интерпретации результатов тестирования

Данные практические рекомендации являются руководством пользователя при работе с [комплексом программ для математической обработки результатов тестирования и их интерпретации](#). Комплекс предназначен для преподавателей, интересующихся современными средствами оценки результатов обучения обучающихся, а также для неспециалистов в области тестологии, стремящихся к овладению методами создания контрольно-оценочных средств и совершенствованию используемых в учебном процессе педагогических тестов. Руководство содержит сведения, необходимые для понимания технологии математической обработки результатов тестирования в рамках классической теории. В тексте имеются все формулы, по которым ведётся расчёт, а также советы по интерпретации полученных результатов.

Работа с комплексом позволяет пользователю, не имеющему специальной подготовки, быстро и легко выполнить расчёт основных характеристик теста и заданий.

Минимальные системные требования: наличие *Microsoft Excel*.

Для удобства перемещения по тексту используйте имеющиеся в нём ссылки. Для быстрого перехода из [Содержания](#) к какому-либо разделу нажмите клавишу CTRL и щёлкните ссылку. Для возврата к первой странице нажмите одновременно клавиши CTRL и HOME.

### Общие сведения о комплексе

Комплекс программ для математической обработки и интерпретации результатов тестирования создан средствами *Microsoft Office* и представляет собой книгу *Excel* с именем МАТЕМ\_ОБ.xls, на листах которой отображаются результаты расчётов, производимых с помощью макросов, содержащихся в книге.

### Содержание комплекса

#### Листы

Название	Представленная информация
Инструкция *	Краткая инструкция по работе с файлом
Данные для отладки *	Данные четырёх примеров, которые можно использовать для знакомства с работой
Данные тестирования **	Данные о тестировании: <ul style="list-style-type: none"> <li>– количество испытуемых и заданий;</li> <li>– варианты ответов испытуемых;</li> <li>– эталон для проверки;</li> <li>– вес заданий;</li> <li>– обозначение пропуска задания;</li> <li>– символы, используемые для обозначения вариантов ответа.</li> </ul>
Результаты тестирования ***	Основные характеристики теста: <ul style="list-style-type: none"> <li>– количество испытуемых, правильно выполнивших задание;</li> <li>– индивидуальные баллы испытуемых, средний балл;</li> <li>– размах, дисперсия и стандартное отклонение индивидуальных баллов;</li> <li>– результат проверки распределения баллов на нормальность;</li> <li>– коэффициент надежности теста (<i>имеется возможность расчёта количества заданий, необходимого для повышения надежности</i>).</li> </ul>
Диаграмма *	Гистограмма распределения индивидуальных баллов

Характеристики заданий ***	Основные характеристики заданий: – трудность; – валидность; – коэффициент дифференцирующей способности ( <i>возможен пересчёт для разных значений процента наполнения контрастных групп</i> ); – рекомендации по интерпретации.
Дистракторный анализ *	Коэффициенты корреляции среднего балла по тесту со средним баллом испытуемых, выбиравших дистрактор, и рекомендации по их интерпретации, а также доработке дистракторов.
Данные дистракт_ан *	Количество испытуемых, выбиравших дистрактор. Результат проверки условия равномерной привлекательности дистракторов
IRT *	Результаты расчетов в рамках современной теории создания тестов
ICC_diagr2-3 ***	Характеристические кривые заданий в 1-, 2- или 3-параметрической модели
IFC_diagr2-3 ***	Информационные кривые заданий и теста в 2- и 3-параметрической модели

\* пользователю запрещены любые изменения

\*\* заполняется пользователем

\*\*\* пользователю разрешено заполнять отдельные ячейки для перерасчётов

Перед началом ввода новых данных некоторые листы скрыты от пользователя; они появляются в процессе работы по мере готовности результатов. Некоторые листы необходимы для хранения промежуточных результатов и скрыты постоянно.

### Кнопки и макросы

<i>Название кнопки</i>	<i>Располагается на листе</i>	<i>Что выполняет связанный с ней макрос</i>
<a href="#">Очистить</a>	Данные тестирования	Удаляет всю информацию предыдущего сеанса работы. При необходимости пользователь должен сам позаботиться о сохранении информации, скопировав её в отдельный файл.
<a href="#">Подготовить таблицу</a>	Данные тестирования (видна только после нажатия кнопки <b>Очистить</b> )	Подготавливает сетку таблицы для ввода ответов испытуемых. Количество строк и столбцов соответствует количеству испытуемых и заданий (если эти значения не введены, при выполнении макроса возникнет ошибка).
<a href="#">Перейти к результатам</a>	Данные тестирования (не видна после нажатия кнопки <b>Очистить</b> ; появляется после подготовки таблицы ввода данных)	Делает видимыми листы <b>Результаты тестирования</b> и <b>Диаграмма</b> . Производит расчёт основных характеристик теста, формирует диаграмму; подготавливает информацию для обработки на следующих этапах. Для корректной работы необходимо заполнить все ячейки, выделенные жёлтым цветом на листе <b>Данные тестирования</b> (кроме ячеек <i>Пропуск</i> и <i>Варианты выбора</i> ).
<a href="#">Вычислить характеристики заданий</a>	Результаты тестирования	Делает видимым лист <b>Характеристики заданий</b> . Производит расчёт основных характеристик заданий. Корректно работает, если пользователь не вносил изменений на имеющихся листах.

<a href="#">Дистракторный анализ</a>	Результаты тестирования	Делает видимыми листы <b>Дистракторный анализ</b> и <b>Данные дистракт_ан</b> . Производит расчёт данных для дистракторного анализа. Для корректной работы необходимы все данные с листа <b>Данные тестирования</b> (включая ячейки <i>Пропуск</i> и <i>Варианты выбора</i> ). Результат не имеет смысла, если в таблице данных сведения о правильности выполнения заданий, а не варианты выбора.
<a href="#">IRT</a>	Характеристики заданий	Делает видимым лист <b>IRT</b> . Производит расчёт параметров в рамках современной теории создания тестов. Корректно работает, если пользователь не вносил изменений на имеющихся листах.
<a href="#">Пересчитать диф.способность</a>	Характеристики заданий	Удаляет коэффициенты дифференцирующей способности для прежнего процента наполнения контрастных групп и производит расчёт для нового значения процента. Корректно работает, если пользователь не вносил изменений на имеющихся листах.

На листе **Результаты тестирования** при вводе желаемого значения коэффициента надёжности происходит автоматический расчёт (перерасчёт) количества заданий, необходимого для повышения надёжности.

На листе **Характеристики заданий** при выборе нового уровня значимости происходит изменение выделения цветом коэффициентов валидности в соответствии с изменением уровня значимости.

### Особенности работы книги с макросами

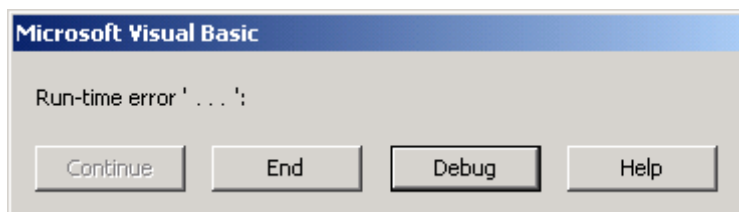
Для корректной работы макросов пользователю запрещается:

- 1) добавлять, удалять или переименовывать листы;
- 2) добавлять и удалять ячейки, строки и столбцы на листах книги;
- 3) самостоятельно заполнять ячейки, не предназначенные для ввода данных.

При необходимости внесения изменений завершите расчёты и скопируйте листы с результатами в новую книгу.

Действия, выполняемые макросами невозможно отменить; а информация, потерянная во время работы макроса, не восстанавливается. Поэтому перед запуском макроса на выполнение ознакомьтесь с тем, какие данные нужны для его корректной работы, какие действия он производит и какую информацию удаляет.

При отсутствии данных, которые обрабатывает макрос, возникают ошибки времени выполнения программы:



Для устранения ошибки выполните следующие действия:

- 1) нажмите кнопку **End** в окне сообщения об ошибке;
- 2) вернитесь к листу, на котором должны содержаться данные, обрабатываемые макросом; проверьте, заполнены ли все ячейки, выделенные жёлтым цветом (при необходимости заполните пустые);

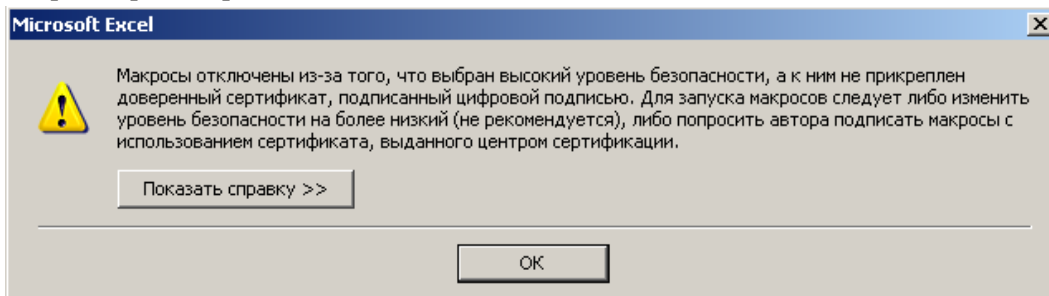
3) ещё раз запустите макрос, нажав соответствующую кнопку.

### Начало работы

Откройте файл [МАТЕМ\\_ОВ.xls](#).

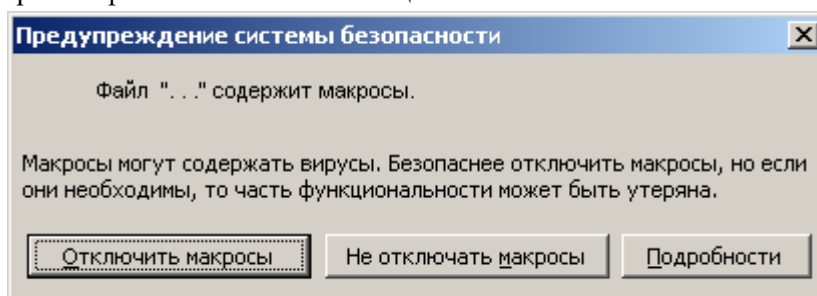
При открытии файла могут появляться предупреждения системы безопасности об отключении макросов. Для их подключения необходимо произвести следующие действия:

- При открытии файла появляется сообщение



В этом случае после нажатия кнопки **ОК** в пункте меню **Сервис** выберите команду **Параметры...** В появившемся окне на вкладке **Безопасность** нажмите кнопку **Безопасность макросов...** и выберите средний уровень безопасности. После этого закройте файл МАТЕМ\_ОВ.xls и снова откройте его.

- При открытии файла появляется сообщение



В этом случае достаточно нажать кнопку «Не отключать макросы».

На листе **Инструкция** даны краткие указания по работе с файлом. При несоблюдении указанной последовательности действий возможна некорректная работа макросов, осуществляющих расчёты. Перейдите на лист **Данные тестирования**. На нём уже имеются четыре набора данных, на примере которых Вы можете ознакомиться с работой программного комплекса.

### Ввод данных тестирования

Для обработки новых данных перейдите на лист **Данные тестирования** и нажмите кнопку **Очистить**, после чего будут удалены прежняя информация с листа данных и все результаты предыдущих расчётов. При необходимости сохраните эти результаты в отдельный файл.

Введите количество испытуемых и число заданий в тесте в соответствующие ячейки, выделенные жёлтым цветом. Если какое-то задание в Вашем тесте имело разбиение на подвопросы, то каждый из них будет обрабатываться отдельно и иметь свой номер; учитывайте это обстоятельство при вводе количества заданий.

После нажатия кнопки **Подготовить таблицу** появятся сетка и шапка таблицы для ввода данных. Дистракторный анализ будет проводиться для заданий с выбором, поэтому в строке **С выбором** введите информацию о типе задания: 1, если это задание с выбором, и 0 в противном случае. Для дистракторного анализа потребуется информация о том, какие символы использовались для обозначения предлагаемых вариантов ответов в заданиях с выбором. Каждый из использовавшихся

хотя бы в одном задании символов должен быть записан в жёлтой ячейке рамки **варианты выбора**; при этом ВСЕ используемые в тесте символы введите В ОДНУ ячейку. При формировании данных дистракторного анализа повторяющиеся символы этой ячейки будут удалены, поэтому при её заполнении не меняйте язык ввода (некоторые русские и латинские буквы одинаковы по начертанию, но различны при обработке; не используйте также строчную и прописную буквы для обозначения различных вариантов ответа).

Для более полного дистракторного анализа требуется отметить задания, которые испытуемым были пропущены. Для этого Вы можете выбрать любой символ (или набор символов), не используемых для обозначения вариантов выбора. Внесите выбранный символ в ячейку рамки **пропуск** листа **Данные тестирования**.

Заполнение таблицы данных начинайте со строки **Эталон**, в которой должны содержаться варианты правильных ответов.

**Максимальный балл** за задание по умолчанию назначен равным единице. Если правильное выполнение задания должно вносить больший вклад в результат испытуемого по тесту, то введите максимальный балл за задание в соответствующую ячейку.

Далее заполните основную часть таблицы.

Существует два варианта представления данных тестирования:

- 1) для каждого испытуемого имеется информация о том, какой вариант ответа он выбирал при ответе на каждое задание с выбором;
- 2) результаты тестирования представляют собой набор нулей и единиц, обозначающих неверно и правильно выполненные задания соответственно (при заполнении эталона в этом случае в каждую ячейку введите единицу).

Рекомендуется первый вариант представления, поскольку второй не даёт информации для дистракторного анализа. И в том и в другом варианте рекомендуется определённым символом отмечать задания, которые испытуемый пропустил.

Обратите внимание, что номера испытуемых и заданий заполняются автоматически. Не изменяйте содержимого этих ячеек во избежание путаницы при дальнейшей обработке. Если Вам необходимо соотнести результаты с конкретными испытуемыми, скопируйте необходимую информацию в отдельную книгу и дополните номера испытуемых фамилиями и именами.

После заполнения таблицы нажмите кнопку [Перейти к результатам](#) (расчёт и заполнение ячеек с результатами сопровождается миганием экрана, что не является сбоем в работе компьютера). Этот этап завершается появлением в рабочей книге двух новых листов – **Результаты тестирования** и **Диаграмма**.

### Анализ результатов тестирования

#### Таблица результатов

На листе **Результаты тестирования** сформирована таблица результатов, в которой содержится следующая информация:

- 1) В каждой строке представлен результат проверки правильности выполнения заданий испытуемым: если выбранный ответ соответствует эталону, то результат выполнения задания равен весу задания, в противном случае (задание выполнено неправильно или пропущено) – нулю.
- 2) Столбец  $X_i$  содержит «сырые» баллы испытуемых, равные сумме баллов за выполнение отдельных заданий ( $i = 1, 2, \dots, N$ ).
- 3) Строка  $R_j$  содержит количества испытуемых, правильно выполнивших соответствующее задание ( $j = 1, 2, \dots, K$ ).
- 4) Столбец  $X_i^2$  и строки  $p_j$  и  $p_j q_j$  содержат значения, необходимые для дальнейших расчётов ( $p_j = R_j / K$ ,  $q_j = 1 - p_j$ ).



## Основные характеристики теста

Программы представленного комплекса позволяют произвести расчёт следующих характеристик:

- 1) [размах](#) индивидуальных баллов;
- 2) [мода](#) распределения индивидуальных баллов;
- 3) [среднее](#) выборочное значение индивидуальных баллов;
- 4) [дисперсия](#) и [стандартное отклонение](#) индивидуальных баллов;
- 5) [оценка близости распределения к нормальному](#);
- 6) [оценка надёжности](#) теста;
- 7) данные для построения [диаграммы](#) распределения индивидуальных баллов.

Ниже представлены формулы, используемые программным комплексом для расчёта этих характеристик, а также рекомендации по интерпретации получаемых результатов.

### **Размах индивидуальных баллов**

*Размах*  $X_{\max} - X_{\min}$  измеряет расстояние, в пределах которого изменяются все значения показателя в распределении (индивидуальных баллов).

### **Среднее выборочное**

*Среднее выборочное (среднее арифметическое)* для совокупности индивидуальных баллов  $X_1, X_2, \dots, X_K$  группы  $K$  испытуемых вычисляется по формуле

$$x_{cp} = \frac{\sum_{i=1}^K X_i}{K} = \frac{X_1 + X_2 + \dots + X_K}{K}.$$

### **Дисперсия**

Подсчет *дисперсии* основан на вычислении отклонений каждого значения показателя от среднего арифметического в распределении:

$$s_x^2 = \frac{\sum_{i=1}^K (X_i - \bar{X})^2}{K - 1} = \frac{K \sum_{i=1}^K X_i^2 - \left( \sum_{i=1}^K X_i \right)^2}{K(K - 1)}.$$

Низкая дисперсия указывает на низкое качество теста, поскольку слабая вариация результатов говорит о слабой дифференциации испытуемых по уровню подготовки. Излишне высокая дисперсия характерна для случая, когда все учащиеся различны по числу выполненных заданий, что также требует переработки теста.

### **Стандартное отклонение**

*Стандартное отклонение* равно корню квадратному из дисперсии  $s_x = \sqrt{s_x^2}$ . Оно характеризует степень рассеяния отдельных значений вокруг среднего. Совокупность с большей неоднородностью будет иметь большие по модулю отклонения, наоборот, для однородных распределений отклонения должны быть близки к нулю.

### **Проверка близости распределения к нормальному**

Проверка близости распределения баллов к нормальному осуществляется с помощью критерия согласия Пирсона. Расчёт производится для выборок объемом не менее 50 испытуемых. В процессе расчёта происходит автоматическое объединение данных для исключения вариантов с нулевыми частотами; если такое объединение невозможно, то как и в случае малой выборки выдаётся сообщение, что проверка распределения баллов на нормальность невозможна. Если же расчёт осуществлен, то результат (ИСТИНА или ЛОЖЬ) представлен на листе **Результаты тестирования**. Уровень значимости (вероятность отклонения гипотезы о незначимости отличия распределения от нормального) можно менять.

### **Оценка надёжности теста**

Завершает расчёт характеристик оценка надёжности теста. Для вычисления коэффициента надёжности используются две формулы:

- *коэффициент Кьюдера-Ричардсона* (вычисляется только в случае, когда все веса заданий равны единице) [6]:



$$KR_{20} = \frac{N}{N-1} \left( 1 - \frac{\sum_{i=1}^N p_i q_i}{s_x^2} \right);$$

– коэффициент надежности Гутмана [5]

$$RG = 1 - \frac{\sum_{i=1}^K E_i}{N \cdot K},$$

где  $E_i$  – число ошибок в профиле  $i$ -ого испытуемого (если учащийся верно выполнил  $n_0$  заданий теста, то ошибками тестирования являются нули на местах с номерами, не превосходящими  $n_0$ , и единицы на местах с номерами, большими  $n_0$ ).

Чтобы по значению коэффициента дать качественную оценку надежности теста, используют следующую таблицу:

Значение коэффициента надежности	Оценка надежности
менее 0,69	неудовлетворительная
0,7-0,79	удовлетворительная
0,8-0,89	хорошая
более 0,9	отличная

Одним из способов *повышения надёжности* является увеличение количества заданий в тесте.

Если обозначить через  $R_{\text{над}}$  – имеющееся значение коэффициента надежности, а через  $R_{\text{план}}$  – желаемое, то число раз, в которое должно возрасти количество заданий в тесте определяется формулой:

$$l = \frac{R_{\text{план}} (1 - R_{\text{над}})}{R_{\text{над}} (1 - R_{\text{план}})}.$$

На листе **Результаты тестирования** производится расчёт нового количества заданий в тесте для любого желаемого значения коэффициента надёжности (которое нужно ввести в ячейку, выделенную жёлтым цветом). Чем меньше имеющееся значение коэффициента, тем большее количество заданий нужно добавить для повышения надёжности. Поэтому если вычисляются оба коэффициента KR-20 и RG, то при расчёте берётся больший из них.

### **Построение диаграммы**

Для построения диаграммы распределения индивидуальных баллов необходимо построить *частотный ряд*, – для каждого возможного значения индивидуального балла подсчитывается количество испытуемых, получивших этот балл. Частотный ряд представлен на листе **Результаты тестирования** под таблицей результатов испытуемых справа от характеристик теста. В первой строке этого ряда перечислены все баллы от минимального до максимального (включая также те баллы, которые не получены ни одним испытуемым, но содержатся в указанных пределах), во второй строке – частоты получения баллов. Гистограмма распределения, построенная по этим данным, расположена на отдельном листе **Диаграмма**. Рядом с частотным рядом имеется значение моды. *Мода* – это значение, которое встречается наиболее часто среди результатов выполнения теста (на гистограмме распределения мода – это значение, соответствующее самому высокому столбику). Значения баллов, которые встречаются наиболее часто (их частота равна моде), выделены в частотном ряду синим цветом и жирным шрифтом.

Если Вы случайно внесли изменения на листе **Результаты тестирования** (в ячейках, не предназначенных для ввода, а особенно выделенных красным цветом), то вся информация может быть восстановлена. Для этого вернитесь на лист **Данные тестирования** и нажмите кнопку [Перейти к результатам](#).

В верхней части листа **Результаты тестирования** имеются три кнопки – [Вычислить характеристики заданий](#), [Дистракторный анализ](#) и [IRT](#) которые позволяют перейти к следующим этапам математической обработки и интерпретации результатов.

### Расчёт характеристик заданий

На листе **Результаты тестирования** нажмите кнопку [Вычислить характеристики заданий](#) (расчёт и заполнение ячеек с результатами сопровождается миганием экрана, что не является сбоем в работе компьютера).

На вновь появившемся листе **Характеристики заданий** представлена таблица, содержащая информацию о [трудности](#), [валидности](#) и [дифференцирующей способности](#) каждого задания.

### Оценка трудности заданий

Оценка трудности  $j$ -го задания рассчитывается по формуле [4, 6]

$$p_j = \frac{R_j}{K}.$$

Заметим, что чем легче задание, тем больше будет доля правильных ответов на него ( $p_j$ ), поэтому естественнее было бы интерпретировать эту долю как легкость задания.

В хорошо сбалансированном по трудности тесте должно быть несколько трудных заданий, несколько легких, но основная масса заданий должна иметь трудность от 0,3 до 0,7; при этом желательно, чтобы задания располагались в порядке возрастания их трудности.

В таблице для наглядности ячейки со значениями трудностей, меньшими 0,3 и большими 0,7, выделены цветом (см. обозначения под таблицей).

### Оценка валидности заданий

Валидность заданий теста определяется степенью соответствия задания цели дифференциации испытуемых. Для этого определяются коэффициенты корреляции оценки за задание с баллом по всему тесту. Это делается с помощью коэффициента корреляции по формуле

$$r = \frac{\sum_{i=1}^K X_i Y_i - \frac{1}{K} \sum_{i=1}^K X_i \sum_{i=1}^K Y_i}{\sqrt{\sum_{i=1}^K X_i^2 - \frac{1}{K} \left( \sum_{i=1}^K X_i \right)^2} \cdot \sqrt{\sum_{i=1}^K Y_i^2 - \frac{1}{K} \left( \sum_{i=1}^K Y_i \right)^2}},$$

где  $X_i$  – тестовый балл  $i$ -го испытуемого,  $Y_i$  – балл  $i$ -го испытуемого за задание. Заметим, что в случае дихотомического оценивания задания расчёт коэффициента несколько упрощается [6].

Если  $r < 0$ , то задание следует удалить из теста, т. к. в нем побеждают слабые ученики, а сильные выбирают неверный ответ либо пропускают задание при выполнении теста. Положительные значения, но близкие к нулю (незначимые), указывают на низкую прогностическую способность задания теста; такие задания требуют доработки содержания. В таблице для наглядности ячейки с незначимыми коэффициентами валидности и значениями, меньшими нуля, выделены цветом (см. обозначения под таблицей в файле МАТЕМ\_ОБ.xls).

Значимость коэффициентов валидности определяется при заданном уровне  $\alpha$ . Если коэффициент значим, то с вероятностью  $1-\alpha$  тестовый результат и балл за задание связаны линейной зависимостью. Если коэффициент незначим, то с вероятностью  $1-\alpha$  тестовый результат и балл за задание не связаны линейной зависимостью. Уровень значимости можно менять после расчёта характеристик, – изменения автоматически отражаются на листе **Характеристики заданий**.

### **Оценка дифференцирующей способности заданий**

Способность дифференцировать испытуемых на лучших и худших показывает *коэффициент дифференцирующей способности* (или *индекс дискриминативности*) задания. Самый простой способ вычисления такого индекса называется методом контрастных групп и заключается в следующем [1, 3]. Из всей группы испытуемых выделяется некоторая часть лучших по результатам выполнения теста испытуемых (будем называть их сильной подгруппой) и столько же худших (слабая подгруппа). Затем для каждой из этих подгрупп вычисляется доля правильных ответов в подгруппе. Обозначим через  $p_{1j}$  долю правильных ответов на  $j$ -е задание в сильной подгруппе, а через  $p_{0j}$  – долю правильных ответов в слабой подгруппе. Тогда индекс дискриминативности  $i$ -го задания определяется по формуле:

$$(r_{\text{дис}})_j = p_{1j} - p_{0j}.$$

Для задания, с которым справились все сильные испытуемые, и не справился ни один слабый, индекс дискриминативности  $r_{\text{дис}}$  будет равен 1; в этом случае задание обладает максимальным дифференцирующим эффектом. Для задания, с которым справились все слабые испытуемые, и не справился ни один сильный, индекс дискриминативности будет равен  $-1$ . В остальных случаях индекс будет принимать значения между  $-1$  и  $1$ . Задания с нулевым и отрицательным значением индекса дискриминативности плохо дифференцируют учащихся, поэтому должны быть удалены из теста. Если индекс положителен, но меньше  $0,2$ , то такое задание требует тщательного анализа содержания. В таблице для наглядности ячейки со значениями индекса дифференцирующей способности, меньшими нуля и находящимися в интервале  $0-0,2$ , выделены цветом (см. обозначения под таблицей в файле МАТЕМ\_ОВ.xls).

Процент наполнения контрастных групп можно варьировать. Для этого введите желаемое значение процента в ячейку, выделенную жёлтым цветом, в верхней части листа **Характеристики заданий**, после чего нажмите кнопку [Пересчитать диф. способность](#). Если введено отрицательное значение процента, то установится значение по умолчанию (27 %, рекомендуется некоторыми авторами [3]). Если значение процента превышает 50 %, то одни и те же испытуемые будут включены в разные контрастные группы, чего быть не должно, поэтому в этом случае процент устанавливается в максимально допустимое значение (50 %). Пересчёт дифференцирующей способности можно производить неоднократно, выполняя каждый раз описанные выше действия.

### **Дистракторный анализ**

На листе **Результаты тестирования** нажмите кнопку [Дистракторный анализ](#). Расчёт занимает относительно длительное время (до нескольких минут; зависит от количества испытуемых, числа заданий и дистракторов), в течение которого идёт считывание и обработка информации и заполнение ячеек, сопровождающееся миганием экрана, что не является сбоем в работе компьютера.

Предлагаемый вариант дистракторного анализа построен на подсчете значения точечно-бисериального коэффициента корреляции для каждого дистрактора в заданиях теста. Формула для подсчета значений коэффициента и методика расчета аналогичны расчету  $r_{pbis}$ , описанному при определении валидности заданий [6]. Отличие заключается в том, что вместо испытуемых, верно (неверно) выполнивших задания, рассматриваются испытуемые, выбравшие (не выбравшие)

данный дистрактор. Точечно-бисериальный коэффициент в данном случае является характеристикой, в некотором смысле обобщающей количественную оценку привлекательности варианта ответа и его качественную оценку (сильные или слабые испытуемые его выбирают).

При анализе качества дистракторов и заданий с выбором особого внимания требуют следующие значения показателя (в таблице все они выделены определённым цветом):

- 1) Отрицательные значения коэффициента корреляции указывают на ситуацию, когда хорошо выполнившие тест испытуемые не будут выбирать данный дистрактор в качестве правильного ответа. *Положительные значения коэффициента для дистракторов* указывают на то, что эти неверные ответы выбирают в качестве правильных сильные ученики, что противоречит представлению преподавателя о хороших заданиях теста. Поэтому подобные дистракторы следует проанализировать и, если просчет в содержании не поддается коррекции, удалить из теста.
- 2) При правильном положении вещей значения коэффициентов точечно-бисериальной корреляции для дистракторов должны быть отрицательными и превышающими по модулю 0,2. *Близкие к нулю отрицательные значения коэффициентов корреляции для дистракторов* (от -0,2 до 0) говорят о необходимости их переделки или исключения.
- 3) В отличие от дистракторов в хороших заданиях правильные ответы выбираются сильными учениками. Правильным ответам должны соответствовать значения коэффициентов, превышающие 0,5. *Отрицательная или нулевая корреляция для верного ответа* говорит о некорректности формулировки задания. Если просчет в содержании таких заданий не поддается коррекции, то их необходимо удалить из теста.
- 4) *Положительные значения коэффициентов, меньшие 0,5*, для правильных ответов могут отражать случайный характер ответов учеников, наличие систематических проблем в усвоении проверяемого материала, вызванных дефектами преподавания, либо некорректную формулировку задания.
- 5) Некоторые задания испытуемые в процессе работы могут пропускать. Нормальной можно считать ситуацию, когда пропущенные задания появляются ближе к концу (особенно для достаточно длинного теста), – в этом случае можно говорить о том, что некоторые испытуемые из-за недостатка времени просто не добрались до этих заданий. Если при этом средний балл пропустивших задание ниже среднего, то это говорит лишь о том, что задание было трудным для слабых испытуемых, и они пропускали его, а не пытались угадать правильный ответ. Высокий средний балл пропускавших задание (*положительные значения коэффициентов в строке пропустили*) может свидетельствовать о том, что в формулировке задания содержится некорректность, которая сбивает с толку сильного испытуемого, поэтому такие задания требуют тщательного анализа содержания.
- 6) Дистракторы, которые выбирают *менее 5% испытуемых* являются настолько неправдоподобными, что в данном задании просто создают иллюзию существования нескольких ответов. Подобные дистракторы должны быть либо удалены из задания, либо изменены. Если же такие значения будут среди правильных ответов, это будет говорить о существенном пробеле в знаниях испытуемых либо о крайне некорректной постановке задания.
- 7) Неверно ответившие ученики должны распределяться почти равномерно между всеми предложенными дистракторами (*условие равномерной привлекательности дистракторов*). Например, если на задание правильно ответили 70 % испытуемых, и в задании имеется три дистрактора, то каждый из дистракторов должны выбирать примерно  $(100\% - 70\%)/3 = 10\%$  испытуемых.

Заметим, что нулевые коэффициенты корреляции при дистракторном анализе не означают, что соответствующий дистрактор не выбирался. Поэтому в качестве дополнительной информации для дистракторного анализа на листе **Данные дистракт\_ан** представлены количества испытуемых,

выбравших дистракторы или пропустивших задание, а также результат проверки условия равномерной привлекательности дистракторов.

В случае, когда данные тестирования имеют вид 0/1 (правильное/неправильное выполнение задания), проводится анализ информации об испытуемых, правильно выполнивших, не справившихся и пропустивших задание.

### Расчёт параметров в рамках современной теории тестов (IRT)

Современная теория тестов (IRT) не отрицает классическую, а развивает ее [6]. Поэтому разработку теста следует разбить на два этапа: сначала эмпирические данные обрабатываются с помощью математико-статистического аппарата классической теории тестов, после чего для углубленного анализа качества заданий и объективной оценки их параметров привлекается аппарат IRT. Современная теория тестов имеет ряд ограничений в применении. Поскольку в основе ее применения лежат абсолютные математические модели, она не работает на маленьких выборках (желательно иметь не менее 1000 человек), а также на результатах «сырых» тестов. Одно из важных условий – распределение результатов по нормальному закону.

#### Характеристические кривые заданий

В современной теории создания тестов значения уровня подготовленности испытуемых (обозначается  $\theta$ ) и трудности заданий теста ( $\beta$ ) выражены в одних и тех же единицах измерения, а значит, могут быть помещены на одну стандартную шкалу, что позволяет соотнести уровень знаний любого испытуемого с мерой трудности каждого задания теста. IRT позволяет определить вероятность правильного ответа на задание теста, которая выражается как функция одной переменной. Можно рассмотреть условную вероятность  $P_j(\theta)$  правильного выполнения  $j$ -го задания с трудностью  $\beta_j$  испытуемыми с различными уровнями подготовки  $\theta$ . График функции  $P_j(\theta)$  называется характеристической кривой задания (ICC). Для выражения функции  $P_j(\theta)$  существуют различные модели.

Однопараметрическая модель, которая часто называется простой логистической моделью, является одной из семейства логистических кривых, описанных Г. Рашем. Аналитическое задание характеристической кривой задания в однопараметрической модели представлено формулой

$$P_j(\theta) = \frac{e^{1,7(\theta-\beta_j)}}{1 + e^{1,7(\theta-\beta_j)}},$$

где в качестве параметра выбирается трудность задания  $\beta_j$ .

Формула в двухпараметрической модели имеет вид

$$P_j(\theta) = \frac{e^{1,7a_j(\theta-\beta_j)}}{1 + e^{1,7a_j(\theta-\beta_j)}},$$

где второй параметр  $a_j$  представляет собой дифференцирующую способность задания.

Для того чтобы учесть фактор угадывания, А.Бирнбаум предложил трехпараметрическую логистическую модель

$$P_j(\theta) = c_j + (1 - c_j) \cdot \frac{e^{1,7a_j(\theta-\beta_j)}}{1 + e^{1,7a_j(\theta-\beta_j)}}.$$

Третий параметр  $c_j$  – вероятность правильного ответа испытуемым на  $j$ -тое задание теста при полном отсутствии знаний у тестируемых. Этот параметр имеет ненулевое значение для заданий с выбором ответа. Его величина определяется количеством предлагаемых для выбора ответов: например, для задания с пятью ответами по классическому определению вероятности  $c = 0,2$ , при четырех предложенных ответах  $c = 0,25$ . В представленном варианте расчета учитывается количество реально работающих вариантов ответа (то есть неработающие дистракторы повышают вероятность угадывания), определяемое результатами дистракторного анализа.

#### Информационные кривые заданий

Наиболее сильный аргумент в пользу современной теории создания тестов связан с введением информационной функции, которая используется для оценки эффективности теста. Чем больше значение информационной функции в некотором интервале оси латентной переменной  $\theta$ , тем лучше работает задание в этом интервале.

Для построения информационной кривой задания (*IFC*) используется информационная функция, заданная формулой:

- в однопараметрической модели

$$I_j(\theta) = 2,89 \frac{e^{1,7(\theta-\beta_j)}}{(1 + e^{1,7(\theta-\beta_j)})^2},$$

- в двухпараметрической модели

$$I_j(\theta) = \frac{2,89 \cdot a_j^2 \cdot e^{1,7a_j(\theta-\beta_j)}}{(1 + e^{1,7a_j(\theta-\beta_j)})^2},$$

- в трехпараметрической модели

$$I_j(\theta) = \frac{2,89 \cdot a_j^2 \cdot (1 - c_j)}{(c_j + e^{1,7a_j(\theta-\beta_j)})(1 + e^{-1,7a_j(\theta-\beta_j)})^2}.$$

Информационная функция теста складывается из информационных функций всех заданий

$$I(\theta) = \sum_{j=1}^N I_j(\theta).$$

### **Расчёт в рамках теории IRT**

Расчёты в рамках теории IRT не могут быть произведены в следующих случаях:

- 1) если среди испытуемых имеются те, кто выполнил все задания (в этом случае на листе **Характеристики заданий** будут трудности, равные 1);
- 2) если среди испытуемых имеются те, кто не выполнил ни одного задания (на листе **Характеристики заданий** будут трудности, равные 0);
- 3) если среди заданий имеются такие, которые выполнены всеми испытуемыми (на листе **Результаты тестирования** будут строки, для которых балл  $X_i$  равен сумме максимальных баллов за задание);
- 4) если среди заданий имеются такие, которые не выполнены ни одним испытуемым (на листе **Результаты тестирования** будут строки, для которых балл  $X_i = 0$ ).

Если среди данных имеются такие, при которых расчет невозможен, кнопка IRT будет недоступна. В этом случае рекомендуется сохранить в отдельный файл результаты обработки в рамках классической теории, после чего на листе **Данные тестирования** удалите строки и столбцы, которые делают невозможными вычисления, и произведите все расчеты заново. Если все данные допустимы, то на листе **Характеристики заданий** будет присутствовать кнопка **IRT**, после нажатия которой на листе **IRT** будут представлены результаты обработки данных в рамках современной теории создания тестов. На листе **ICC\_diagr** представлены характеристические кривые заданий в одно-, двух- или трех-параметрической модели (количество параметров модели можно выбрать с помощью соответствующего поля в правом верхнем углу диаграммы). На листе **IFC\_diagr** представлены информационные кривые заданий и теста в целом в одно-, двух- или трех-параметрической модели соответственно (количество параметров модели можно выбрать с помощью соответствующего поля в правом верхнем углу диаграммы).



Сдача государственного экзамена	1) Положительный, доброжелательный, располагающий 2) Нейтральный, спокойный, безэмоциональный 3) Отрицательный, недоброжелательный, депрессивный
Защита выпускной квалификационной работы	1) Положительный, доброжелательный, располагающий 2) Нейтральный, спокойный, безэмоциональный 3) Отрицательный, недоброжелательный, депрессивный

**9. Согласны Вы или нет с тем, что оценивание, осуществляемый в ходе итоговой аттестации, является формальным и не отражает Вашего действительного уровня знаний, умений и сформированности компетенций?**

Этап ГИА	Варианты ответа
Сдача государственного экзамена	1) Да, полностью согласен(-а) 2) Скорее согласен(-а), чем не согласен(-а) 3) Скорее не согласен(-а), чем согласен(-а) 4) Нет, полностью не согласен(-а)
Защита выпускной квалификационной работы	1) Да, полностью согласен 2) Скорее согласен, чем не согласен 3) Скорее не согласен, чем согласен 4) Нет, полностью не согласен

**10. Насколько на Ваш взгляд аттестационные испытания являются объективными?** (Пожалуйста, оцените по 5-ти балльной шкале, где 1-абсолютно необъективна, 5-полностью объективна).

Этап ГИА	Варианты ответа
Сдача государственного экзамена	1 2 3 4 5
Защита выпускной квалификационной работы	1 2 3 4 5

**11. Оцените по следующим критериям степень независимости оценки пройденных Вами аттестационных испытаний:**

Критерий	Государственный экзамен	Выпускная квалификационная работа
Все члены комиссии присутствовали на аттестационном испытании:	Да Нет Затрудняюсь ответить	Да Нет Затрудняюсь ответить
У каждого члена комиссии были оценочные листы:	Да Нет Затрудняюсь ответить	Да Нет Затрудняюсь ответить
Вы предварительно были ознакомлены с требованиями, критериями	Да Нет	Да Нет



и шкалой оценки Ваших образовательных результатов:		
--	--	--

**12. Насколько отметка, полученная в ходе аттестационных испытаний, отражает Ваш уровень знаний, навыков и компетенций?**

Этап ГИА	Варианты ответа
Государственный экзамен	1) Полученная отметка соответствует моим образовательным достижениям 2) Полученная отметка не соответствует моим образовательным достижениям – она завышена 3) Полученная отметка не соответствует моим образовательным достижениям – она занижена
ВКР	1) Полученная отметка соответствует моим образовательным достижениям 2) Полученная отметка не соответствует моим образовательным достижениям – она завышена 3) Полученная отметка не соответствует моим образовательным достижениям – она занижена

**13. Соответствуют ли результаты обучения в университете Вашим ожиданиям?**

- 1) Полностью соответствуют
- 2) Скорее соответствуют
- 3) Не знаю
- 4) Скорее не соответствуют
- 5) Не соответствуют

**14. Выразите свое отношение к адекватности существующих аттестационных испытаний, выбрав одно суждение из каждой пары.**

А	Результаты государственного экзамена отражают готовность выпускника к будущей профессиональной деятельности
А	В ходе государственного экзамена нельзя в полной мере оценить готовность выпускника к будущей профессиональной деятельности
Б	Успешность подготовки и защиты выпускной квалификационной работы свидетельствует о готовности выпускника к будущей профессиональной деятельности
Б	Подготовка и защита выпускной квалификационной работы не позволяет оценить готовность выпускника к будущей профессиональной деятельности

**15. Считаете ли Вы, что необходимо совершенствовать существующую систему оценивания в рамках итоговой аттестации?**

- 1) Безусловно необходимо
- 2) Скорее необходимо
- 3) Не вижу необходимости

4) Мне безразлично

**16. Как Вы считаете сколько этапов должна включать в себя государственная итоговая аттестация, чтобы объективно оценить готовность выпускника к будущей профессиональной деятельности?**

- 1) Одна стадия (например, только сдача государственного экзамена или иные испытания)
- 2) Две стадии (например, сдача гос. экзамена и защита ВКР или иные испытания)
- 3) Три стадии (например, тестирование, решение кейсов и защита ВКР или иные испытания)
- 4) Другое (напишите)

**17. Какой способ организации итоговой аттестации, на Ваш взгляд, является наиболее оптимальным для оценки профессиональных навыков и компетенций? Выберите несколько вариантов ответа.**

- 1) Государственный экзамен по билетам
- 2) Написание и защита ВКР
- 3) Решений кейсов
- 4) Стартап как диплом
- 5) Бланковое тестирование
- 6) Демонстрационный экзамен
- 7) Компьютерное тестирование
- 8) Решение профессиональных ситуационных задач
- 9) Другое

**18. Считаете ли Вы, что существующая система итоговой аттестации достаточно объективна и справедлива?**

- 1) Да – система итоговой аттестации объективна и справедлива.
- 2) Скорее да – в целом существующая система итоговой аттестации соответствует условиям и потребностям организации учебного процесса.
- 3) Скорее нет – систему итоговой аттестации необходимо совершенствовать и использовать более оптимальные способы оценивания.
- 4) Нет – необходимо оценивать не столько уровень образовательных результатов на аттестационных испытаниях, сколько отношение к учебе, индивидуальные способности, самостоятельную работу, активность на семинарах и ответ на экзамене, т. е. совокупность учебных достижений каждого студента.

## Анкета для работодателя

Какую организацию Вы представляете?	<i>(Напишите название или тип организации)</i>
В реализации какой образовательной программы Вы принимаете участие?	Перечень
Ведете ли Вы учебные занятия в ИвГУ?	1) Да 2) Нет 3) Нет, но хотел(-а) бы 4) Не знал(-а), что это возможно
Принимаете ли Вы (организация, которую Вы представляете) обучающихся на практику?	1) Да 2) Нет 3) Затрудняюсь ответить
Принимаете ли Вы участие в работе государственной экзаменационной комиссии?	1) Да Нет
Принимаете ли Вы участие в проведении промежуточной аттестации по дисциплинам (практикам)?	1) Да 2) Нет 3) Нет, но хотел(-а) бы 4) Не знал(-а), что это возможно
Принимаете ли Вы участие в разработке и(или) в согласовании образовательной программы?	1) Да, в разработке образовательной программы в целом 2) Да, в разработке отдельных дисциплин (практик) 3) Нет, но хотел(-а) бы 4) Нет
Насколько Вы в целом удовлетворены сотрудничеством с ИвГУ?	1) Удовлетворен(-а) в полной мере 2) Скорее удовлетворен(-а) 3) Скорее не удовлетворен(-а) 4) Не удовлетворен(-а)
Какие, на Ваш взгляд, существуют основные проблемы взаимодействия организаций (предприятий) с вузами?	1) Отсутствие федеральных или региональных программ по вопросам взаимодействия бизнеса и образования 2) Отсутствие законодательной базы для участия работодателей в формировании образовательных стандартов 3) Отсутствие стимулов у работодателя вкладывать средства в образование 4) Отсутствие налоговых льгот для предприятий, инвестирующих в подготовку специалистов 5) Значительные риски при заключении договора о целевом обучении для

	адресной подготовки специалиста для организации (предприятия) б) Другое (напишите)
Как бы Вы в целом оценили условия обучения в ИвГУ (качество аудиторий, фондов читального зала и библиотеки, учебных лабораторий, оборудования и пр. материально-техническое обеспечение)?	1 2 3 4 5
Какие мероприятия по привлечению выпускников ИвГУ для Вас интересны?	1) «Ярмарки вакансий» (встречи с обучающимися с целью поиска сотрудников для организации) 2) Рекрутинг (размещение объявлений о вакансиях в Университете) 3) Использование интернет - ресурсов для информирования о вакансиях 4) Участие в реализации образовательной программы 5) Мастер-классы для обучающихся 6) Экскурсии обучающихся в Вашу организацию 7) «Дни организации» (презентации организации перед обучающимися) 8) Совместные конференции 9) Предоставление мест прохождения практики 10) Программы стажировок для обучающихся 11) Другое
Предоставляете ли Вы (организация, которую Вы представляете) возможность трудоустройства обучающихся по итогам прохождения практики в Вашей организации?	1) Да, часто 2) Иногда 3) Редко 4) Нет, никогда
Предоставляете ли Вы возможность трудоустройства выпускников программы в Вашу организацию?	1) Да, часто 2) Иногда 3) Редко 4) Нет, никогда
Работают ли в Вашей организации выпускники ИвГУ последних 3-х лет?	1) Да 2) Нет
Удовлетворены ли Вы (организация, которую Вы представляете) уровнем подготовки выпускников ИвГУ?	1) Полностью удовлетворена 2) Скорее удовлетворена 3) Отчасти удовлетворена, отчасти нет 4) Скорее не удовлетворена 5) Полностью не удовлетворена 6) Затрудняюсь ответить

<p>Укажите основные достоинства в подготовке выпускников образовательных программ ИвГУ:</p>	<ol style="list-style-type: none"> <li>1) Высокий уровень теоретических знаний</li> <li>2) Высокий уровень практической подготовки</li> <li>3) Владение современными методами и технологиями деятельности</li> <li>4) Профессионализм выпускников</li> <li>5) Готовность выпускников к быстрому реагированию в нестандартных ситуациях</li> <li>6) Высокий уровень производственной дисциплины</li> <li>7) Желание выпускников работать</li> <li>8) Желание выпускников к саморазвитию и самообразованию</li> <li>9) Другое</li> </ol>
<p>Укажите основные недостатки в подготовке выпускников ИвГУ:</p>	<ol style="list-style-type: none"> <li>1) Низкий уровень теоретических знаний</li> <li>2) Недостаточный уровень практической подготовки</li> <li>3) Отсутствие желания работать</li> <li>4) Низкая производственная дисциплина</li> <li>5) Отсутствие желания к саморазвитию и самообразованию</li> <li>6) Низкий уровень общей профессиональной подготовки</li> <li>7) Не отмечаю недостатков</li> <li>8) Другое</li> </ol>
<p>Требуется ли, на Ваш взгляд, молодым специалистам дополнительное обучение (профессиональная переподготовка, повышение квалификации, получение дополнительной квалификации) для работы в Вашей организации?</p>	<ol style="list-style-type: none"> <li>1) Да</li> <li>2) Скорее да, чем нет</li> <li>3) Скорее нет, чем да</li> <li>4) Нет</li> </ol>
<p>Как, на Ваш взгляд, изменилось за последние 3 года качество подготовки выпускников ИвГУ?</p>	<p>Значительно улучшилось  Незначительно улучшилось  Незначительно ухудшилось  Значительно ухудшилось  Затрудняюсь ответить  Не могу сравнить (в организации нет специалистов за указанный период выпуска)</p>

## Приложение 12

### 1. Основные характеристики компетенции и индикаторы достижения компетенции

**Профессиональная компетенция:** Способен осуществлять психолого-педагогическую диагностику метапредметных и личностных образовательных результатов обучающихся

Тип задач профессиональной деятельности: сопровождение.

Задача профессиональной деятельности	Код и наименование ПК	Код и наименование индикатора достижения ПК	Основания включения в перечень планируемых результатов освоения ОП и установления индикаторов достижения ПК
<p>- организация и проведение психологической диагностики особенностей детей и обучающихся, в том числе с ограниченными возможностями здоровья, а также обучающихся, испытывающих трудности в освоении основных общеобразовательных программ, развитии и социальной адаптации;</p> <p>- осуществление мониторинга личностных и метапредметных результатов освоения</p>	<p>ПК-1. Способен осуществлять психолого-педагогическую диагностику метапредметных и личностных образовательных результатов обучающихся</p>	<p><b>Знать:</b> теорию психодиагностики, классификацию психодиагностических методов, их возможности и ограничения, предъявляемые к ним требования; методы сбора, обработки информации, результатов психологических наблюдений и диагностики; способы интерпретации и представления результатов психодиагностического обследования; содержание понятий: «универсальные учебные действия» (УУД), «личностные образовательные результаты», и «метапредметные образовательные результаты», психологические методы и средства освоения социокультурного опыта; закономерности личностного развития обучающихся.</p> <p><b>Уметь:</b></p>	<p><b>Код 01.002 Профессиональный стандарт «Педагог-психолог (психолог в сфере образования)»:</b>  <b>А/05.7 Психологическая диагностика детей и обучающихся</b>  <b>ТД:</b> Психологическая диагностика с использованием современных образовательных технологий, включая информационные образовательные ресурсы  <b>Скрининговые обследования (мониторинг)</b> с целью анализа динамики психического развития, определение лиц, нуждающихся в психологической помощи  <b>Составление психолого-педагогических заключений по результатам диагностического обследования с целью ориентации педагогов, преподавателей, администрации образовательных организаций и родителей (законных представителей) в проблемах личностного и социального развития обучающихся</b>  <b>НУ:</b> - Планировать и проводить диагностическое обследование с использованием стандартизированного инструментария, включая обработку результатов,                      - Проводить мониторинг личностных и метапредметных образовательных результатов обучающихся в соответствии с требованиями федеральных государственных образовательных стандартов общего образования соответствующего уровня,  <b>НЗ:</b> Федеральные государственные образовательные стандарты общего образования.</p>

<p>основной общеобразовательной программы с использованием современных средств информационно - коммуникационных технологий (ИКТ)</p>		<p>проводить диагностическое обследование с использованием стандартизированного инструментария, включая обработку результатов; проводить диагностическую работу по выявлению уровня готовности или адаптации детей и обучающихся к новым образовательным условиям; диагностировать интеллектуальные, личностные и эмоционально-волевые особенности, препятствующие нормальному протеканию процесса развития, обучения и воспитания;</p> <p>использовать психолого-педагогические средства для формирования, развития и оценки УУД, анализировать индивидуальные возможности обучающихся по достижению метапредметных и личностных образовательных результатов.</p> <p><b>Иметь практический опыт / навыки:</b> применения методов и методик психолого-педагогической диагностики; владеть способами сбора, анализа и интерпретации полученных результатов; приемами формирования и развития УУД, рефлексией личностных и метапредметных образовательных результатов обучающихся.</p>	<p><b>Код 01.001 Профессиональный стандарт "Педагог</b> (педагогическая деятельность в сфере дошкольного, начального общего, основного общего, среднего общего образования) (воспитатель, учитель)»: А/01.6 Общепедагогическая функция. Обучение</p> <p>НЗ: Пути достижения образовательных результатов и способы оценки результатов обучения</p> <p>А/03.6 Развивающая деятельность</p> <p>ТД: Применение инструментария и методов диагностики и оценки показателей уровня и динамики развития ребенка</p> <p>НУ: Составить (совместно с психологом и другими специалистами) психолого-педагогическую характеристику (портрет) личности обучающегося;</p> <p>Владеть стандартизированными методами психодиагностики личностных характеристик и возрастных особенностей обучающихся;</p> <p>Оценивать образовательные результаты: формируемые в преподаваемом предмете предметные и метапредметные компетенции, а также осуществлять (совместно с психологом) мониторинг личностных характеристик.</p>
--	--	--	--

## 1. Основные характеристики компетенции и индикаторы достижения компетенции

**Профессиональная компетенция:** Способен проводить консультации субъектов образовательного процесса по психологическим проблемам обучения и развития (ПК-3)

Тип задач профессиональной деятельности: сопровождение.

Задача профессиональной деятельности	Код и наименование ПК	Код и наименование индикатора достижения ПК	Основания включения в перечень планируемых результатов освоения ОП и установления индикаторов достижения ПК
<p>- Организация и проведение психолого-педагогического консультирования обучающихся, их родителей (законных представителей) и педагогических работников</p> <p>- Ознакомление педагогов, преподавателей, администрации образовательных организаций и организаций, осуществляющих образовательную деятельность, а также родителей (законных представителей) с основными условиями психического развития лиц с ограниченными возможностями здоровья, детей и</p>	<p>• ПК-3. Способен проводить консультации субъектов образовательного процесса по психологическим проблемам обучения и развития.</p>	<p><b>Знать:</b> этические нормы организации и проведения консультативной работы; современные теории и методы консультирования; приемов организации совместной и индивидуальной деятельности обучающихся в соответствии с возрастными нормами их развития.</p> <p><b>Уметь:</b> организовывать взаимодействие с участниками образовательного процесса; проводить индивидуальные и групповые консультации обучающихся по вопросам обучения, развития</p> <p><b>Иметь практический опыт / навыки:</b> консультирования педагогов, преподавателей, родителей (законных представителей) по психологическим проблемам обучения, воспитания и развития обучающихся.</p>	<p><b>Код 01.002 Профессиональный стандарт «Педагог-психолог (психолог в сфере образования)»:</b> <b>A/03.7</b></p> <p><b>ТД:</b> - Консультирование обучающихся по проблемам самопознания, профессионального самоопределения, личностным проблемам, вопросам взаимоотношений в коллективе и другим вопросам,</p> <p>- Консультирование администрации, педагогов, преподавателей и других работников образовательных организаций по проблемам взаимоотношений в трудовом коллективе и другим профессиональным вопросам,</p> <p>- Консультирование родителей (законных представителей) по проблемам взаимоотношений с обучающимися, их развития, профессионального самоопределения и другим вопросам,</p> <p>- Консультирование администрации образовательной организации, педагогов, преподавателей, родителей (законных представителей) по психологическим проблемам обучения, воспитания и развития обучающихся,</p> <p><b>НУ:</b> - Проводить индивидуальные и групповые консультации обучающихся по вопросам обучения, развития, проблемам осознанного и ответственного выбора дальнейшей профессиональной карьеры, самовоспитания, взаимоотношений со взрослыми и сверстниками,</p> <p><b>НЗ:</b> - Современные теории и методы консультирования,</p> <p>- Приемы организации совместной и индивидуальной деятельности обучающихся в соответствии с возрастными нормами их развития,</p> <p>- Этические нормы организации и проведения консультативной работы,</p>



<p>обучающихся, испытывающих трудности в освоении основных общеобразовательных программ, развитии и социальной адаптации (в рамках консультирования, педагогических советов)</p>			<p>- Содержание работы межведомственных организаций (ресурсных центров) для информирования субъектов образовательного процесса о способах получения отраслевой психолого-педагогической, медицинской и социальной помощи.</p>
--	--	--	---

## ОТЧЕТ

по результатам социологического опроса (анкетирования)  
**«Оценка удовлетворенности работодателей условиями, организацией  
и качеством образовательного процесса»**

## Сводные результаты по всем образовательным программам

**Сроки проведения анкетирования:** 2022-23 учебный год

**Выборка:** 64 респондента.

**Перечень направлений подготовки в реализации которых участвуют респонденты:**

Направления подготовки:
01.03.01 Математика (Математика)
01.03.01 Математика (Математика, алгоритмы и анализ данных)
01.04.01 Математика (Фундаментальная математика)
01.04.01 Математика (Математические методы в экономике)
02.03.01 Математика и компьютерные науки (Математика и компьютерные науки)
02.04.01 Математика и компьютерные науки (Математические методы в компьютерных науках)
02.04.01 Математика и компьютерные науки (Цифровое моделирование экономических процессов)
02.03.02 Фундаментальная информатика и информационные технологии (Фундаментальная информатика и информационные технологии)
03.03.02 Физика (Физика)
03.03.02 Физика (Фундаментальная и прикладная физика)
04.03.01 Химия (Медицинская и фармацевтическая химия)
04.04.01 Химия (Инноватика в химии и химическом образовании)
04.05.01 Фундаментальная и прикладная химия (Фундаментальная и прикладная химия)
06.03.01 Биология (Биология)
06.03.01 Биология (Биоэкология и биоразнообразие)
06.03.01 Биология (Биохимия)
06.03.01 Биология (Биохимия и физиология)
06.04.01 Биология (Фундаментальная биология)
09.03.03 Прикладная информатика (Прикладная информатика в экономике)
09.03.03 Прикладная информатика (Прикладная информатика в цифровой экономике)
09.04.01 Прикладная информатика (Управление проектами цифровой трансформации)
10.03.01 Информационная безопасность (Безопасность компьютерных систем (по отрасли или в сфере профессиональной деятельности))
28.03.01 Нанотехнологии и микросистемная техника (Материалы микро- и наносистемной техники)
37.03.01 Психология (Психологическое консультирование и психодиагностика)
38.03.01 Экономика (Финансы и кредит)
38.03.01 Экономика (Бухгалтерский учет, анализ и аудит)
38.03.01 Экономика (Учет и анализ бизнес-процессов)
38.04.08 Финансы и кредит (Банки и банковская деятельность)
38.04.08 Финансы и кредит (Финансовый менеджмент)
38.03.01 Экономика (Экономика предприятий и организаций)
38.03.01 Экономика (Внешнеэкономическая деятельность)
38.04.01 Экономика (Экономика фирмы, отраслевых рынков и инновационной деятельности)
38.03.03 Управление персоналом (Управление личностно-профессиональным развитием персонала)

38.03.02 Менеджмент (Менеджмент организации)
38.04.02 Менеджмент (Менеджмент и технологии в ритейле)
38.04.02 Менеджмент (Общий и стратегический менеджмент)
39.03.01 Социология (Социология социальных процессов)
39.04.01 Социология (Социология управления современными социальными процессами)
39.03.02 Социальная работа (Социальная работа и социальное проектирование)
39.03.02 Социальная работа (Социальная работа с различными группами населения)
40.03.01 Юриспруденция (Юриспруденция)
40.04.01 Юриспруденция (Социальное государство: проблемы теории и юридической практики)
.04.01 Юриспруденция (Юрист в сфере правоприменения и правотворчества)
41.03.05 Международные отношения (Мировые политические процессы)
42.03.01 Реклама и связи с общественностью (Реклама и связи с общественностью)
42.03.02 Журналистика (Журналистика)
44.03.01 Педагогическое образование (Обучение иностранным языкам)
44.03.02 Психолого-педагогическое образование (Психология образования)
44.04.02 Психолого-педагогическое образование (
45.03.01 Филология (Зарубежная филология)
45.04.01 Филология (Зарубежная филология)
45.03.01 Филология (Отечественная филология (русский язык и литература))
45.04.01 Филология (Отечественная филология)
45.04.01 Филология (Русский язык и культура в современном мире)
46.03.01 История (История)
46.04.01 История (Человек и общественные институты в истории России и Европы)
46.04.01 История (Человек и общество в истории России)
46.03.01 История (Культурно-историческое проектирование)
49.03.01 Физическая культура (Физкультурно-оздоровительная деятельность)
49.03.01 Физическая культура (Лечебно-оздоровительная физическая культура)
49.03.01 Физическая культура (Спортивная тренировка в избранном виде спорта)

**Перечень организаций принявших участие в исследовании:**

<b>Организация</b>	<b>Проценты</b>
Банк России	1,6
Администрация города Иваново	4,7
Администрация Панинского сельского поселения Фурмановского района, тип организации - орган местного самоуправления Ивановской области	3,1
АО «Ивановский кабельный завод»	1,6
АО КИБ «ЕВРОАЛЬЯНС»	1,6
ООО «Априори»	1,6
БУ «Ивановские газеты»	4,7
ГАУДПО ИО «Университет непрерывного образования и инноваций»	1,6
ГБУ ИО «Государственный архив Ивановской области»	3,1
ГБУИО «Ивановский областной художественный музей»	1,6
Департамент социальной защиты населения Ивановской области	1,6
Департамент спорта Ивановской области	1,6
Ивановская городская Дума	1,6
Ивановская транспортная прокуратура	1,6

ИвГМА	1,6
ИКБ им. Куваевых	1,6
Иллюминатор	1,6
ИХР РАН	1,6
Коллегия адвокатов	1,6
МАУ ДО ЦТТ Новация	1,6
МБОУ «Лицей 67»	1,6
МБОУ «Подвязновская СШ»	1,6
МБУ «Методический центр в системе образования»	1,6
МБУ ДО ДЮСШ №1	1,6
МБУ ДО СШ №2	1,6
МБУ ДО ЦСК Притяжение	1,6
МБОУ "СШ № 42"	1,6
ОБСУСО «Ивановский дом-интернат»	1,6
ОБУСО Центр социальной помощи семье и детям "На Московской"	1,6
ООО «Технологии развития»	1,6
ОГБУ ДО СШОР № 8 «Спартак»	1,6
ООО «АВН»	1,6
ООО «Гарпикс»	1,6
ООО «ИМИДЖ-ФАКТОР»	1,6
ООО «ИНИЦИАТИВА»	1,6
ООО «Металлоконструкции»	1,6
ООО «СМАРТ ПРОЖЕКТС»	1,6
ООО «Кенгуру»	3,1
ООО «Профессионал»	3,1
Педагогический колледж	1,6
Правительство Ивановской области	1,6
ООО «Стандартпласт»	3,1
СШОР №1 «Текстильщик»	1,6
ЧОУ «Лицей Исток»	3,1
ЧОУ ДО «Альбион плюс»	1,6
ЧОУ «Школа Инглиш.ру»	1,6
<b>Иные</b> <i>в том числе</i> <i>Бюджетное учреждение</i> <i>Государственный орган</i> <i>Государственное учреждение</i> <i>Розничная и оптовая торговля</i> <i>Судебная система</i> <i>Сертифицированная ИТ компания</i> <i>Научно-исследовательский институт</i>	
<b>Итого</b>	

*Распределение ответов на вопросы анкеты*

1. Ведете ли Вы учебные занятия в ИвГУ?

<b>Причастность к ведению учебных занятий</b>	<b>Проценты</b>
Да	48,4
Нет	42,2
Нет, но хотел(-а) бы	9,4
Не знал(-а), что это возможно	0,0
<b>Итого</b>	

2. Принимаете ли Вы (организация, которую Вы представляете) обучающихся на практику?

<b>Участие в организации практики</b>	<b>Проценты</b>
Да	79,7
Нет	15,6
Затрудняюсь ответить	4,7
<b>Итого</b>	

3. Принимаете ли Вы участие в работе государственной экзаменационной комиссии?

<b>Участие в работе ГЭК</b>	<b>Проценты</b>
Да	75,0
Нет	25,0
<b>Итого</b>	

4. Принимаете ли Вы участие в проведении промежуточной аттестации по дисциплинам (практикам)?

<b>Участие в проведении промежуточных аттестаций</b>	<b>Проценты</b>
Да	42,2
Нет	50,0
Нет, но хотел(-а) бы	6,3
Не знал(-а), что это возможно	1,6
<b>Итого</b>	

5. Принимаете ли Вы участие в разработке и(или) в согласовании образовательной программы?

<b>Участие в разработке и(или) в согласовании ОП</b>	<b>Проценты</b>
Да, в разработке образовательной программы в целом	7,8
Да, в разработке отдельных дисциплин (практик)	21,9
Нет, но хотел(-а) бы	12,5
Нет	57,8
<b>Итого</b>	

6. Насколько Вы в целом удовлетворены сотрудничеством с ИвГУ?

<b>Уровень удовлетворенности сотрудничеством</b>	<b>Проценты</b>
Удовлетворен(-а) в полной мере	70,3
Скорее удовлетворен(-а)	26,6
Скорее не удовлетворен(-а)	3,1
Не удовлетворен(-а)	

<b>Итого</b>	
--------------	--

7. Какие, на Ваш взгляд, существуют основные проблемы взаимодействия организаций (предприятий) с вузами?

<b>Основные проблемы взаимодействия вуза и организаций</b>	<b>Проценты</b>
Отсутствие федеральных или региональных программ по вопросам взаимодействия бизнеса и образования	26,6
Отсутствие законодательной базы для участия работодателей в формировании образовательных стандартов	23,4
Отсутствие стимулов у работодателя вкладывать средства в образование	31,3
Отсутствие налоговых льгот для предприятий, инвестирующих в подготовку специалистов	4,7
Значительные риски при заключении договора о целевом обучении для адресной подготовки специалиста для организации (предприятия)	6,3
Другое	7,8
<b>Итого</b>	

8. Как бы Вы в целом оценили условия обучения в ИвГУ (качество аудиторий, фондов читального зала и библиотеки, учебных лабораторий, оборудования и пр. материально-техническое обеспечение)?

<b>Оценка по 5-ти балльной шкале</b>	<b>Проценты</b>
	0,0
	3,1
	6,3
	56,3
	34,4
<b>Итого</b>	

<b>Среднее значение:</b>	
--------------------------	--

9. Какие мероприятия по привлечению выпускников ИвГУ для Вас интересны?

<b>Мероприятия по привлечению выпускников</b>	<b>Проценты</b>
«Ярмарки вакансий» (встречи с обучающимися с целью поиска сотрудников для организации)	53,1
Рекрутинг (размещение объявлений о вакансиях в Университете)	39,1
Использование интернет - ресурсов для информирования о вакансиях	37,5
Участие в реализации образовательной программы	34,4
Мастер-классы для обучающихся	31,3
Экскурсии обучающихся в Вашу организацию	51,6
«Дни организации» (презентации организации перед обучающимися)	25,0
Совместные конференции	42,2
Предоставление мест прохождения практики	56,3
Программы стажировок для обучающихся	37,5

<b>Итого</b>	<i>Предоставлена возможность выбора нескольких вариантов ответа</i>
--------------	---

10. Предоставляете ли Вы (организация, которую Вы представляете) возможность трудоустройства обучающихся по итогам прохождения практики в Вашей организации?

<b>Возможность трудоустройства обучающихся</b>	<b>Проценты</b>
Да, часто	37,5
Иногда	34,4
Редко	21,9
Нет, никогда	6,3
<b>Итого</b>	

11. Предоставляете ли Вы возможность трудоустройства выпускников программы в Вашу организацию?

<b>Возможность трудоустройства выпускников</b>	<b>Проценты</b>
Да, часто	37,5
Иногда	35,9
Редко	20,3
Нет, никогда	6,3
<b>Итого</b>	

12. Работают ли в Вашей организации выпускники ИвГУ последних 3-х лет?

<b>Наличие сотрудников выпускников ИвГУ</b>	<b>Проценты</b>
Да, часто	71,9
Нет	28,1
<b>Итого</b>	

13. Удовлетворены ли Вы (организация, которую Вы представляете) уровнем подготовки выпускников ИвГУ?

<b>Уровень удовлетворенности подготовкой выпускников</b>	<b>Проценты</b>
Полностью удовлетворен(-а)	18,8
Скорее удовлетворен(-а)	53,1
Отчасти удовлетворен(-а), отчасти нет	25,0
Скорее не удовлетворен(-а)	3,1
Полностью не удовлетворен(-а)	
Затрудняюсь ответить	
<b>Итого</b>	

14. Укажите основные достоинства в подготовке выпускников образовательных программ ИвГУ:

<b>Достоинства в подготовке выпускников</b>	<b>Проценты</b>
Высокий уровень теоретических знаний	33,9
Высокий уровень практической подготовки	16,1
Владение современными методами и технологиями деятельности	37,1
Профессионализм выпускников	14,5

Готовность выпускников к быстрому реагированию в нестандартных ситуациях	33,9
Высокий уровень производственной дисциплины	12,9
Желание выпускников работать	43,5
Желание выпускников к саморазвитию и самообразованию	41,9
<b>Итого</b>	<i>Предоставлена возможность выбора нескольких вариантов ответа</i>

15. Укажите основные недостатки в подготовке выпускников ИвГУ:

<b>Недостатки в подготовке выпускников</b>	<b>Проценты</b>
Низкий уровень теоретических знаний	7,9
Недостаточный уровень практической подготовки	44,4
Отсутствие желания работать	22,2
Низкая производственная дисциплина	15,9
Отсутствие желания к саморазвитию и самообразованию	14,3
Низкий уровень общей профессиональной подготовки	4,8
Не отмечаю недостатков	27,0
<b>Итого</b>	<i>Предоставлена возможность выбора нескольких вариантов ответа</i>

16. Требуется ли, на Ваш взгляд, молодым специалистам дополнительное обучение (профессиональная переподготовка, повышение квалификации, получение дополнительной квалификации) для работы в Вашей организации?

<b>Необходимость дополнительного обучения</b>	<b>Проценты</b>
Да	
Скорее да, чем нет	
Скорее нет, чем да	
Нет	
<b>Итого</b>	

17. Как, на Ваш взгляд, изменилось за последние 3 года качество подготовки выпускников ИвГУ?

<b>Насколько изменились выпускники</b>	<b>Проценты</b>
Значительно улучшилось	
Незначительно улучшилось	
Незначительно ухудшилось	
Значительно ухудшилось	
Затрудняюсь ответить	
Не могу сравнить (в организации нет специалистов за указанный период выпуска)	
<b>Итого</b>	

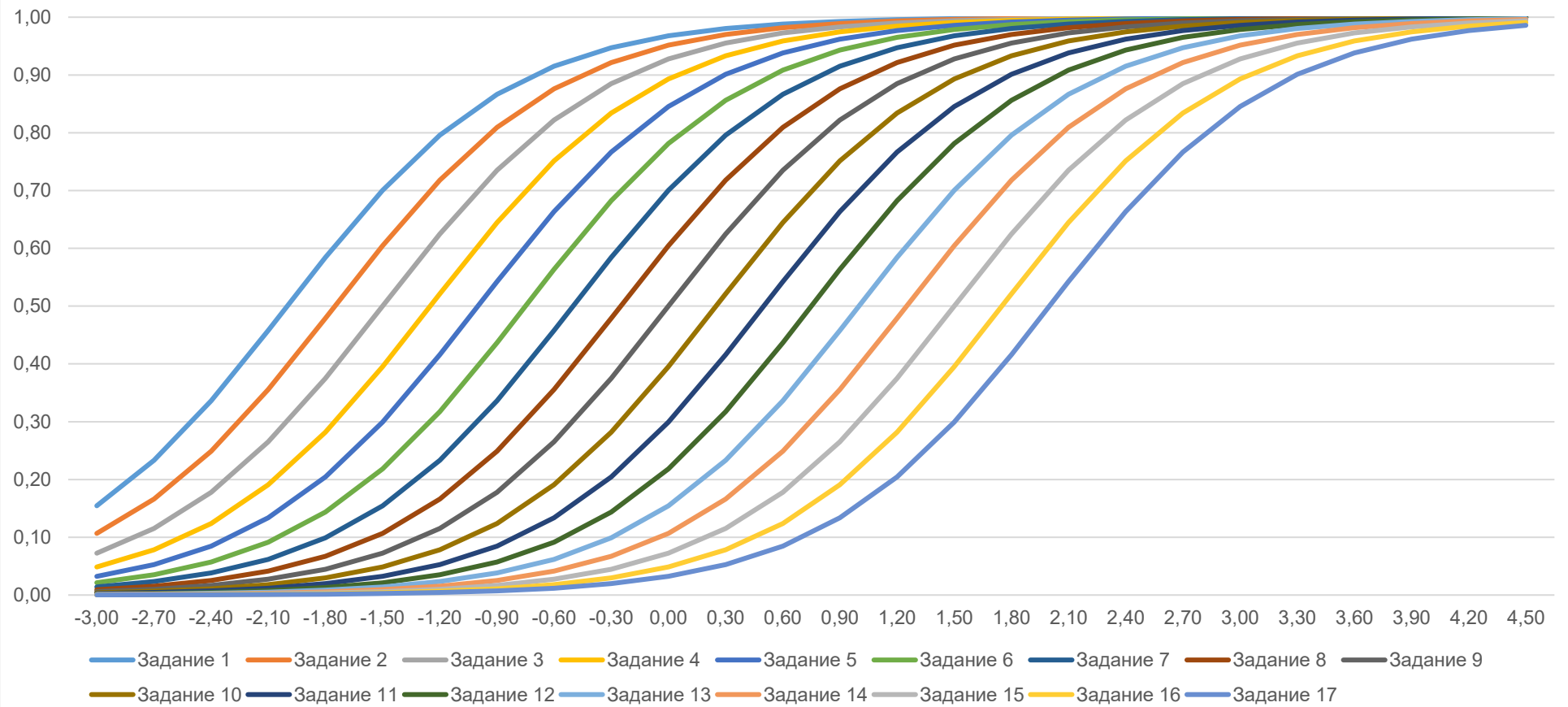


## Приложение 14

### Расчеты вероятностей правильных ответов для 17 заданий по однопараметрической модели IRT при различных значениях уровня подготовленности $\theta$

$\theta$	Задание 1	Задание 2	Задание 3	Задание 4	Задание 5	Задание 6	Задание 7	Задание 8	Задание 9	Задание 10	Задание 11	Задание 12	Задание 13	Задание 14	Задание 15	Задание 16	Задание 17
-2,00	-1,75	-1,50	-1,25	-1,00	-0,75	-0,50	-0,25	0,00	0,25	0,50	0,75	1,00	1,25	1,50	1,75	2,00	
-3,00	0,15	0,11	0,07	0,05	0,03	0,02	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
-2,70	0,23	0,17	0,12	0,08	0,05	0,04	0,02	0,02	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
-2,40	0,34	0,25	0,18	0,12	0,08	0,06	0,04	0,03	0,02	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00
-2,10	0,46	0,36	0,27	0,19	0,13	0,09	0,06	0,04	0,03	0,02	0,01	0,01	0,01	0,00	0,00	0,00	0,00
-1,80	0,58	0,48	0,38	0,28	0,20	0,14	0,10	0,07	0,04	0,03	0,02	0,01	0,01	0,01	0,00	0,00	0,00
-1,50	0,70	0,60	0,50	0,40	0,30	0,22	0,15	0,11	0,07	0,05	0,03	0,02	0,01	0,01	0,01	0,00	0,00
-1,20	0,80	0,72	0,62	0,52	0,42	0,32	0,23	0,17	0,12	0,08	0,05	0,04	0,02	0,02	0,01	0,01	0,00
-0,90	0,87	0,81	0,73	0,64	0,54	0,44	0,34	0,25	0,18	0,12	0,08	0,06	0,04	0,03	0,02	0,01	0,01
-0,60	0,92	0,88	0,82	0,75	0,66	0,56	0,46	0,36	0,27	0,19	0,13	0,09	0,06	0,04	0,03	0,02	0,01
-0,30	0,95	0,92	0,88	0,83	0,77	0,68	0,58	0,48	0,38	0,28	0,20	0,14	0,10	0,07	0,04	0,03	0,02
0,00	0,97	0,95	0,93	0,89	0,85	0,78	0,70	0,60	0,50	0,40	0,30	0,22	0,15	0,11	0,07	0,05	0,03
0,30	0,98	0,97	0,96	0,93	0,90	0,86	0,80	0,72	0,62	0,52	0,42	0,32	0,23	0,17	0,12	0,08	0,05
0,60	0,99	0,98	0,97	0,96	0,94	0,91	0,87	0,81	0,73	0,64	0,54	0,44	0,34	0,25	0,18	0,12	0,08
0,90	0,99	0,99	0,98	0,97	0,96	0,94	0,92	0,88	0,82	0,75	0,66	0,56	0,46	0,36	0,27	0,19	0,13
1,20	1,00	0,99	0,99	0,98	0,98	0,96	0,95	0,92	0,88	0,83	0,77	0,68	0,58	0,48	0,38	0,28	0,20
1,50	1,00	1,00	0,99	0,99	0,99	0,98	0,97	0,95	0,93	0,89	0,85	0,78	0,70	0,60	0,50	0,40	0,30
1,80	1,00	1,00	1,00	0,99	0,99	0,99	0,98	0,97	0,96	0,93	0,90	0,86	0,80	0,72	0,62	0,52	0,42
2,10	1,00	1,00	1,00	1,00	0,99	0,99	0,99	0,98	0,97	0,96	0,94	0,91	0,87	0,81	0,73	0,64	0,54
2,40	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,98	0,97	0,96	0,94	0,92	0,88	0,82	0,75	0,66
2,70	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,98	0,98	0,96	0,95	0,92	0,88	0,83	0,77
3,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,99	0,98	0,97	0,95	0,93	0,89	0,85
3,30	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,99	0,98	0,97	0,96	0,93	0,90
3,60	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,99	0,98	0,97	0,96	0,94
3,90	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,98	0,97	0,96
4,20	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,98	0,98
4,50	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,99

Характеристические кривые для 17 заданий (однопараметрическая модель)



## Протоколы государственной экзаменационной комиссии

Ивановский государственный университет  
Государственная экзаменационная комиссия

### ПРОТОКОЛ № 1/1

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 9 час.00 мин. до 9 час. 30 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
Ф.И.О ЧЛЕНА ГЭК, профессор, канд. пед. наук  
Ф.И.О РАБОТОДАТЕЛЯ, представитель работодателя  
Ф.И.О ЧЛЕНА ГЭК, доцент, канд.психол.наук  
Ф.И.О работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**Б. В. С.**

Вопросы: Билет № 12

**1.** Ситуационная задача № 4.

**2. Охарактеризуйте** контрольно-оценочную деятельность учителя и ученика.

**Проанализируйте** основные недостатки современной системы оценивания результатов обучения.

**Покажите** пути совершенствования этой системы.

**3. Раскройте** содержание понятия «речь».

**Назовите** виды и функции речи.

**Проанализируйте** связь мышления и речи.

**4.** Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Назовите критерии контрольно-оценочной деятельности учителя и ученика? (Ф.И.О ЧЛЕНА ГЭК)

2. Какую роль играет для педагога-психолога знания об особенностях речевого развития? (Ф.И.О работодателя-члена ГЭК)

3. Целесообразно ли изучить особенности психологического климата 8 «А» при определении причин поведения ученика этого класса? (Ф.И.О работодателя-члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/2**

заседания государственной экзаменационной комиссии (ГЭК)

по приему государственного экзамена

Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 9 час. 30 мин. до 10 час. 00 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд.психол.наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**Л. Ю. А.**

Вопросы: Билет № 17.

1. Ситуационная задача № 2.

2. **Охарактеризуйте** понятие «процесс обучения» как педагогическую систему.

**Раскройте** сущность компонентов и характер их взаимосвязи в зависимости от дидактической (образовательной) модели.

**Сравните** подробнее содержание понятия «результаты обучения» в различных дидактических моделях.

3. **Дайте** определение понятиям «мотив» и «мотивация».

**Приведите** основные классификации мотивов.

**Раскройте** роль мотивации в различных видах деятельности.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Назовите особенности деятельности учителя и ученика как участников образовательного процесса? Приведите примеры. (ФИО ЧЛЕНА ГЭК)

2. Какие основные мероприятия развития мотивации в учебной деятельности учеников Вы можете выделить? (ФИО ЧЛЕНА ГЭК)

3. Допускаете ли Вы, что у Ивана может и не быть потребности в повышении самооценки? (ФИО работодателя-члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/3**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 10 час. 00 мин. до 10 час. 30 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд.психол.наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**В. О. А.**

Вопросы: Билет № 15.

1. Ситуационная задача № 1.

2. **Охарактеризуйте** на основе ПС «Педагог (педагогическая деятельность в дошкольном общем, начальном общем, основном общем и среднем общем образовании (учитель, воспитатель))» основные направления деятельности современного педагога по воспитанию и развитию обучающихся.

**Определите** роль педагога-психолога в обеспечении продуктивной деятельности современного учителя (воспитателя) – по выбору.

**Опишите** необходимые для этого знания и умения.

3. **Дайте** определение понятию «способности».

**Назовите** виды способностей. Приведите примеры.

**Проанализируйте** связь способностей с интересами и склонностями.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Можно ли отнести знания, умения и навыки к способностям? (ФИО члена ГЭК)

2. У ребенка рано проявляется интерес к музыке, но педагоги убеждены, что задатков нет. Как можно помочь ребенку в данной ситуации? (ФИО ЧЛЕНА ГЭК)

3. Определите статус ПС «Педагог (педагогическая деятельность в дошкольном общем, начальном общем, основном общем и среднем общем образовании (учитель, воспитатель))»? (ФИО ЧЛЕНА ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент при ответе на все вопросы продемонстрировал исчерпывающее, последовательное и логически стройное изложение.

Признать, что студент сдал государственный экзамен с оценкой «отлично».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/4**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 10 час. 30 мин. до 11 час. 00 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд.психол.наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**К. Ю. С.**

Вопросы: Билет № 18.

1. Ситуационная задача № 3.

2. **Раскройте** взаимосвязь обучения и развития.

**Охарактеризуйте** идеи Л.С. Выготского о развивающем обучении.

**Охарактеризуйте** одну из систем РО в начальной школе (на выбор).

3. **Дайте** психологическую характеристику младенческого возраста.

**Опишите** социальную ситуацию и ведущую деятельность данного возрастного периода.

**Охарактеризуйте** кризис первого года жизни.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Какова точка зрения Ж. Пиаже на обучение и развитие? (ФИО работодателя - члена ГЭК)

2. Какие важные этапы проходит ребенок в младенческом возрасте? Перечислите их. (ФИО члена ГЭК)

3. Если мама такого ребенка перекладывает всю ответственность на свою дочь, то, как вести себя педагогу-психологу в такой ситуации? (ФИО работодателя-члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент при ответе на все вопросы продемонстрировал исчерпывающее, последовательное и логически стройное изложение.

Признать, что студент сдал государственный экзамен с оценкой «отлично».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/5**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 11 час. 00 мин. до 11 час. 30 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
Малыгин А.А., зав. кафедрой НППО, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд. психол. наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**Н. Д. А.**

Вопросы: Билет № 11.

1. Ситуационная задача № 5.

2. **Охарактеризуйте** понятие «урок» как основную форму классно-урочной системы в современной школе.

**Дайте** типологию урока в традиционном обучении и в рамках деятельностного подхода и **охарактеризуйте** структуру урока (на выбор).

**Раскройте** педагогические требования к современному уроку.

3. **Сформулируйте** определение понятия «мышление».

**Проанализируйте** связь мышления с другими познавательными процессами.

**Охарактеризуйте** виды мышления и основные мыслительные операции.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Можно ли назвать урок современным, если учитель не использует информационные технологии? (ФИО ЧЛЕНА ГЭК)

2. Охарактеризуйте урок, который проводится вне стен школы? (Малыгин А. А.)

3. Какими свойствами должно обладать понятие? Приведите пример понятия. (ФИО члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/6**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 11 час. 30 мин. до 12 час. 00 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд.психол.наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**Г. И. С.**

Вопросы: Билет № 5.

1. Ситуационная задача № 11.

2. **Дайте** определения основным категориям педагогики: образование, обучение, воспитание на основе научного знания и 273–ФЗ «Об образовании в РФ».

**Охарактеризуйте** систему образования в РФ.

**Покажите** тенденции ее развития.

3. **Дайте** психологическую характеристику раннего возраста.

**Опишите** социальную ситуацию и ведущую деятельность данного возрастного периода.

**Охарактеризуйте** кризис трех лет.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Перечислите основные тенденции развития современного образования? (ФИО ЧЛЕНА ГЭК)

2. Чем отличается предметно-манипулятивная деятельность от предметно-орудийной? (ФИО члена ГЭК)

3. Какой приоритетный вариант решения в сложившейся ситуации матери с дочерью десятиклассницей? (ФИО ЧЛЕНА ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря



Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/7**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 12 час. 00 мин. до 12 час. 30 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
Круглдова Е.А., доцент, канд. психол. наук  
ФИО работодателя, представитель работодателя, г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**Л. А. Е.**

Вопросы: Билет № 13.

1. Ситуационная задача № 3.

2. **Перечислите** основные направления психолого-педагогической деятельности.

**Охарактеризуйте** более глубоко содержание 2-х видов психолого-педагогической деятельности (по выбору).

**Покажите** их отражение в профессиональном стандарте «Педагог-психолог», **соотнесите** с обобщенными трудовыми функциями, трудовыми действиями, умениями и знаниями (на примере одной трудовой функции).

3. **Дайте** определение понятию «эмоции».

**Перечислите и охарактеризуйте** виды эмоций.

**Назовите** способы регуляции эмоциональных состояний.

4. **Дополнительный ситуационный вопрос.**

Дополнительные вопросы:

1. Можно ли сказать, что профессиональный стандарт педагога-психолога выделяет трудовые функции в соответствии с научными работами, которые посвящены психологической деятельности как специфического вида деятельности? (ФИО ЧЛЕНА ГЭК)

2. Соотнесите понятие психодиагностика и психометрика? (Мальгин А. А.)

3. Назовите основоположника учения о стрессе? (ФИО члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент при ответе на все вопросы продемонстрировал исчерпывающее, последовательное и логически стройное изложение.

Признать, что студент сдал государственный экзамен с оценкой «отлично».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/8**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 12 час. 30 мин. до 13 час. 00 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
Малыгин А.А., зав. кафедрой НППО, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд. психол. наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**М. П. А.**

Вопросы: Билет № 16.

1. Ситуационная задача № 11.

2. **Раскройте** содержание основных понятий: человек, индивид, личность, индивидуальность.

**Охарактеризуйте** биологизаторскую, социологизаторскую теории развития личности, теорию факторов.

**Сформулируйте** современный взгляд на эту проблему и педагогические следствия, вытекающие из этой теории.

3. **Дайте** психологическую характеристику подросткового возраста.

**Опишите** социальную ситуацию и ведущую деятельность данного возрастного периода.

**Проанализируйте** психологические и педагогические подходы к воспитанию подростка

4. Дополнительный ситуационный вопрос

Дополнительные вопросы:

1. Что входит в содержание понятия «личность»? (ФИО работодателя-члена ГЭК)

2. Назовите представителей биологизаторских и социологизаторских концепций? (ФИО члена ГЭК)

3. Назовите возрастные рамки подросткового возраста? (ФИО члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/9**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 13 час. 00 мин. до 13 час. 30 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
Малыгин А.А., зав. кафедрой НППО, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд. психол. наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**П. Е. Н.**

Вопросы: Билет № 6.

1. Ситуационная задача № 10.

2. **Раскройте** содержание основных образовательных моделей на основе работ

Г. Батищева: «педагогика формирования», «педагогика развития ради развития», «педагогика сотрудничества» и на основе других научных источников.

**Выделите** преимущества и недостатки каждой из них.

**Обоснуйте** ответ на вопрос: «На какую модель ориентируется современная образовательная практика?».

3. **Дайте** психологическую характеристику дошкольного возраста.

**Опишите** социальную ситуацию и ведущую деятельность данного возрастного периода.

**Охарактеризуйте** кризис семи лет.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Почему Г. Батищев разделяет педагогику на «педагогику формирования», «педагогику развития ради развития», «педагогику сотрудничества»? (ФИО ЧЛЕНА ГЭК)

2. Почему «игра по правилам» так важна в дошкольном детстве? (ФИО работодателя-члена ГЭК)

3. Может ли сформироваться самооценка в три года? (ФИО работодателя-члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/10**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 13 час. 30 мин. до 14 час. 00 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
Малыгин А.А., зав. кафедрой НППО, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд. психол. наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**Т. Д. А.**

Вопросы: Билет № 20.

**1.** Ситуационная задача № 5.

**2. Раскройте** сущность понятий «активные методы обучения», «интерактивные методы обучения».

**Приведите** конкретные примеры этих методов.

**Покажите** механизм их реализации на практике в условия продуктивного использования.

**2. Дайте** общую характеристику самостоятельной работы учащихся.

**Сформулируйте** основные требования к самостоятельной работе.

**Покажите** пути психолого-педагогического сопровождения самостоятельной работы обучающихся.

**4.** Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. В чем психологические основы самостоятельной работы учащихся школьного возраста?  
(ФИО ЧЛЕНА ГЭК)

2. Приведите примеры «активных методов обучения» и «интерактивных методов обучения»?  
(ФИО ЧЛЕНА ГЭК)

3. Какие методы работы можно использовать при работе с таким ребенком? (ФИО работодателя-члена ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря

Ивановский государственный университет  
Государственная экзаменационная комиссия

**ПРОТОКОЛ № 1/11**

заседания государственной экзаменационной комиссии (ГЭК)  
по приему государственного экзамена  
Направление подготовки 44.03.02 Психолого-педагогическое образование  
(направленность «Психология образования»)

«04» июня 2019 г.

с 14 час. 00 мин. до 14 час. 30 мин.

Присутствовали из состава ГЭК:

Г.Г.В., председатель ГЭК, директор МБУ, г. Иваново  
ФИО ЧЛЕНА ГЭК, профессор, канд. пед. наук  
ФИО РАБОТОДАТЕЛЯ, представитель работодателя  
ФИО ЧЛЕНА ГЭК, доцент, канд.психол.наук  
ФИО работодателя, представитель работодателя,  
г. Иваново

О сдаче государственного экзамена.

Экзаменуется студент социолого-психологического факультета очной формы обучения

**Фамилия Имя Отчество аттестуемого**

**В. А. Р.**

Вопросы: Билет № 24.

1. Ситуационная задача № 9.

2. **Сформулируйте** основные принципы диагностики в воспитании.

**Охарактеризуйте** сущность контроля воспитательной деятельности.

**Покажите** механизмы организации диагностики и контроля воспитательной деятельности в практике на примере опыта современных педагогов.

3. **Рассмотрите** педагога как субъект педагогической деятельности.

**Раскройте** психофизиологические предпосылки деятельности педагога.

**Опишите** профессиональные деформации личности педагога и методы их предупреждения.

4. Дополнительный ситуационный вопрос.

Дополнительные вопросы:

1. Кто осуществляет воспитательную деятельность в школе? (ФИО ЧЛЕНА ГЭК)

2. Почему педагоги-психологи не любят работать с проблемой профессиональной деформации личности педагога? (ФИО ЧЛЕНА ГЭК)

3. Что является предметом диагностики в воспитательной деятельности? (ФИО ЧЛЕНА ГЭК)

Общая характеристика ответа студента на заданные ему вопросы: студент правильно применял теоретические положения, подтвержденные примерами; сделал вывод; один вопрос освещен полностью, а другой был доведен до логического завершения при наводящих вопросах членов комиссии.

Признать, что студент сдал государственный экзамен с оценкой «хорошо».

Председатель ГЭК

\_\_\_\_\_

И.О. Фамилия председателя

Секретарь ГЭК

\_\_\_\_\_

И.О. Фамилия секретаря