

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ИЗМЕРЕНИЯ КОМПЕТЕНЦИЙ

*Маслак А.А.*

Филиал ФГБОУ ВО «Кубанский государственный университет», г. Славянск-на-Кубани  
e-mail: anatoliy\_maslak@mail.ru

**Аннотация.** В современном образовании акцент сделан на оценку результатов обучения. В каждой образовательной программе необходимо формировать десятки компетенций – универсальных, общепрофессиональных, профессиональных, общекультурных. Важным аспектом является измерение уровня сформированности этих компетенций. Требуемое в образовательных стандартах оценивание (пороговый, базовый, продвинутой) является недостаточным при проведении мониторинга и анализа факторов, влияющих на уровень сформированности компетенций. В статье проводится сравнение трех методов оценивания уровня сформированности компетенций: в рамках классической теории тестирования, на основе индексов и в рамках теории измерения латентных переменных.  
**Ключевые слова:** компетенции, латентные переменные, измерение, классическая теория тестирования, теория латентных переменных.

Компетенции выпускника вуза характеризуют его готовность выполнять свои профессиональные обязанности, что имеет важное значение для работодателя. Поэтому не случайно акцент в области образования сделан на измерение результатов обучения [5 - 8].

Существуют три метода измерения уровня сформированности компетенций: на основе классической теории тестирования, на основе индексов и на основе теории латентных переменных.

### **Классическая теория тестирования**

Это наиболее распространенный метод оценивания уровня сформированности компетенций. Так в работе [3] для измерения сформированности компетенций предлагается использовать оценки текущей успеваемости. Однако, как справедливо отмечает автор, эти оценки не являются метрическими, например, расстояние между «тройкой» и «четверкой» и расстояние между «четверкой» и «пятеркой» не являются одинаковыми [3]. Более того, разные преподаватели по-разному применяют оценки, например «тройка» одного преподавателя не эквивалентна «тройке» другого преподавателя.

Рассмотрим преимущества и недостатки классической теории тестирования. В рамках этой теории уровень сформированности компетенций определяется по числу набранных баллов или по доле правильных ответов.

Безусловным преимуществом классической теории тестирования являются доступность, наглядность и интерпретация получаемых оценок:

Однако этой теории присущи существенные недостатки.

- Оценка уровня сформированности компетенции испытуемого зависит от трудности теста. Если тест состоит из трудных тестовых заданий, то доля правильных ответов у испытуемого будет низкой. Если же тест состоит из легких тестовых заданий, то доля правильных ответов у того же самого испытуемого будет высокой.
- Оценка трудности тестового задания зависит от уровня сформированности компетенций испытуемых. Если выборка состоит из хорошо подготовленных испытуемых, то трудность тестовых заданий будет невысокой. Если же выборка состоит из плохо подготовленных испытуемых, то трудность тестовых заданий будет высокой.
- Шкала измерения уровня сформированности компетенции является нелинейной. Например, одна и та же разница в баллах (5 баллов) на краях и в середине шкалы

соответствует разному приращению в уровне сформированности компетенции (100 баллов – 95 баллов)  $\neq$  (50 баллов – 45 баллов).

- Дисперсия оценки уровня сформированности компетенции наибольшая в середине диапазона измерения и наибольшая по краям, что противоречит логике построения доверительных интервалов.
- Как было показано ранее, тестовый балл испытуемого нелинейно зависит от уровня его подготовленности.
- Близким к классической теории тестирования и более усложненным является индексный метод.

### **Индексный метод**

В работе [3] предлагается оценки успеваемости «взвешивать» в зависимости от трудоемкости заданий. Интегральная оценка получается путем суммирования этих взвешенных оценок. В результате интегральная оценка является средневзвешенной.

В общем виде индексный метод измерения уровня сформированности компетенции состоит в следующем:

- на основе экспертных оценок каждому тестовому заданию (индикатору) приписывается вес: чем выше, по мнению эксперта, важность индикатора, тем больше его вес;
- абсолютные значения индикаторов переводятся в относительные. Это осуществляется следующим образом. Для каждого индикатора ищется его максимальное значение, которое и принимается за единицу; остальные значения индикатора выражаются в долях этого максимального значения;
- относительные значения индикатора умножаются на вес этого индикатора, интегральный показатель вычисляется суммированием взвешенных значений всех индикаторов;
- объекты ранжируются на основе этого интегрального показателя.

Однако и эта методика обладает многими существенными недостатками:

1. Экспертные оценки являются субъективными (варьируя весами, можно получить любое ранжирование объектов);
2. Используемый набор индикаторов может характеризовать не одну переменную, а несколько, что искажает получаемые оценки;
3. Получаемые оценки не измеряются на линейной шкале, что затрудняет мониторинг и сравнение объектов.

Необходимо подчеркнуть, что в рамках этих двух рассмотренных методов оценки уровня сформированности компетенций получаются на порядковой шкале. Порядковая шкала является нелинейной, поэтому эти оценки не являются измерениями.

Измерить уровень сформированности компетенций можно только в рамках теории латентных переменных [2, 9].

### **Теория измерения латентных переменных**

Измерение латентных переменных принципиально отличается от выше рассмотренных по многим аспектам:

- результаты измерений являются объективными в том смысле, что не зависят ни от того, кто измеряет, ни от того, какой измерительный инструмент используется;
- обязательным требованием является совместимость набора индикаторов, т. е. проверяется, действительно ли все индикаторы характеризуют одну и ту же латентную переменную;
- уровень подготовленности и трудность задания измеряются на одной и той же интервальной шкале в логитах. Поскольку мера измерения параметров модели Раша является линейной, то это позволяет использовать широкий спектр статистических процедур для анализа результатов измерений;
- модель измерения (модель Раша) является вероятностной;

- латентная переменная определяется операционально, через набор индикаторных переменных;
- чем больше индикаторов, тем выше точность измерения латентной переменной;
- модель Раша превращает измерения, сделанные в дихотомических и порядковых шкалах, в линейные измерения, в результате качественные данные анализируются с помощью количественных методов;
- оценка индикаторов не зависит от выборки испытуемых, на которых была получена;
- оценка уровня подготовленности испытуемых не зависит от используемого набора индикаторов;
- пропуск данных для некоторых комбинаций (испытуемый – индикатор) не является критическим;
- существуют удобные вычислительные процедуры для многоаспектной проверки адекватности модели: для всего набора тестовых результатов, для каждого испытуемого, для каждого индикатора и для каждой комбинации «испытуемый – индикатор».

Выделим наиболее важные аспекты измерений, получаемых на основе модели Раша: объективность измерений, совместимость индикаторов, линейная шкала измерений и вероятностный характер модели измерений [9].

### ***Объективность измерений компетенций***

Как измеряются компетенции на практике и каково же качество измерения компетенции?

Например, при использовании обычной 3-балльной системы оценивания «пороговый уровень», «базовый уровень» и «продвинутый уровень» за одни и те же достижения студентов разные преподаватели могут выбрать разные уровни. Необъективность такого рода измерений очевидна: результат измерений зависит от преподавателя (от того, кто измеряет).

При использовании классической системы тестирования этой необъективности уже нет. Уровень сформированности компетенции определяется по числу правильных ответов. Однако, как было отмечено выше, оценка уровня сформированности компетенции зависит от набора тестовых заданий.

Спрашивается: как можно использовать такого рода измерения для управления качеством образования?

Результаты же объективных измерений не должны зависеть ни от того, кто измеряет (преподавателя), ни от используемого измерительного инструмента (набора тестовых заданий). Именно эту объективность и обеспечивает теория измерения латентных переменных, краеугольным камнем которой является модель Раша.

### ***Совместимость индикаторных переменных***

Для наглядности рассмотрим очень показательную ситуацию, когда один студент правильно ответил на половину тестовых заданий, которые являются наиболее трудными, а на оставшиеся, легкие, задания ответил неправильно. Второй студент, наоборот, на трудную половину заданий ответил неправильно, а на оставшиеся, легкие, задания ответил правильно.

Спрашивается, у кого из студентов, выше уровень сформированности компетенции? Естественно, возникает ответ, что у первого, но как тогда это согласовать с тем, что он не знает элементарных вещей? В классической системе тестирования этот вопрос решается просто: поскольку и тот, и другой студент ответили правильно на одно и то же число вопросов (50%), то оценки их уровня сформированности компетенции одинаковы.

В теории измерения латентных переменных этот вопрос решается принципиально по-другому. Прежде всего, проверяется, можно ли использовать полученные результаты тестирования для измерения уровня сформированности компетенции. Для данной ситуации наверняка будет получен вывод, что полученные результаты тестирования невозможно аппроксимировать моделью Раша и поэтому их нельзя использовать для измерения латентной переменной «уровень сформированности компетенции». Следовательно, вопрос о том, у кого выше уровень сформированности компетенции, останется открытым.

Единственное конструктивное решение здесь – это анализ причин, обусловивших такую ситуацию. Одна из возможных причин, приведших к таким результатам тестирования, может состоять в том, что набор тестовых заданий плохо отработан и сами тестовые задания являются несовместимыми. Как следствие, необходимо улучшать тест: корректировать, исключать или добавлять тестовые задания. Другая причина может заключаться в неадекватном тестировании: плохо организован сам процесс тестирования, в частности не исключены были списывания и т. д.

### ***Линейность шкалы измерения***

Еще раз подчеркнем важное свойство измерительной шкалы – ее линейность. Латентная переменная отображается на прямой, называемой линейным континуумом. Шкала является линейной, если равным приращениям измеряемой переменной соответствуют равные приращения на шкале.

Результаты же измерения при использовании классической системы тестирования являются существенно нелинейными. Так, одна и та же разность в числе правильных ответов по краям шкалы и в середине шкалы соответствует различной разнице в уровне сформированности компетенции. Например, при общем числе тестовых заданий, равном 100, разница между числом правильных ответов 100 и 95 соответствует большей разнице в уровне подготовленности, чем разность между 53 и 48.

### ***Модель измерения является вероятностной***

Существенная особенность модели Раша заключается в том, что она является вероятностной. Целесообразность использования вероятностной модели объясняется следующим. Дело в том, что невозможно описать и предсказать поведение отдельного индивида. Например, хорошо подготовленный студент может «провалиться» на легкой задаче, и, наоборот, плохо подготовленный студент может решить трудную задачу. Мы никогда не можем предсказать, решит или не решит студент задачу, но мы можем сказать, какие у него шансы (высокие или низкие) решить задачу. Поведение студента описывается вероятностной моделью, вероятностью того, что данный студент с известным уровнем подготовленности решит тестовое задание с заданной трудностью.

Вероятностная модель допускает, что даже отличник может ответить неправильно на очень легкое задание (правда, вероятность неправильного ответа очень мала). С другой стороны, если трудность задания значительно превосходит уровень знания студента, то есть вероятность (правда незначительная), что студент правильно ответит на трудное задание. Таким образом, модель Раша является достаточно гибкой и позволяет описать более широкий круг экспериментальных ситуаций по сравнению с детерминированной моделью (моделью Гутмана).

Важным этапом измерения компетенций является их формализация, то есть определение компетенции в виде набора индикаторов. Пример формализованного описания компетенции «Готов применять современные методики и технологии, методы диагностирования достижений обучающихся для обеспечения качества учебно-воспитательного процесса» приведен в работе [15, 16]. Формализованная структура этой компетенции представлена в таблице 1.

Таблица 1 – Формализованное описание компетенции «Готов применять современные методики и технологии, методы диагностирования достижений обучающихся для обеспечения качества учебно-воспитательного процесса»

<b>Аспект компетенции</b>	<b>Элемент компетенции</b>	<b>Характеристика элемента компетенции студента</b>
Когнитивный	Знания в области реальных объектов, по отношению к которым вводится компетенция	Знание: – современных методов, методик и технологий диагностирования достижений обучающихся; – основных показателей качества учебно-воспитательного процесса.

Аспект компетенции	Элемент компетенции	Характеристика элемента компетенции студента
	Знания в области методов, способов и приемов деятельности в сфере данной компетенции	– знание основных целей и способов организации и проведения диагностики достижений учащихся
Праксиологический	Умения, навыки и способы деятельности в сфере компетенции	<p>Умение:</p> <ul style="list-style-type: none"> <li>– основываясь на учебной программе, определять цели диагностирования и выбирать адекватные им современные методы и технологии;</li> <li>– анализировать результаты диагностирования и делать выводы о степени достижения планируемых образовательных результатов</li> </ul>
	Минимально необходимый опыт деятельности студента в сфере компетенции	<p>Опыт:</p> <ul style="list-style-type: none"> <li>– диагностики достижений на основе использования современных методик и технологий;</li> <li>– анализа учебно-воспитательного процесса на основе результатов диагностики достижений учащихся и формулирования соответствующих выводов</li> </ul>
	Самоконтроль деятельности в сфере компетенции и ее результата	<p>Умение:</p> <ul style="list-style-type: none"> <li>– на основе рефлексии делать вывод о целесообразности выбора современных методов, методик и технологий диагностирования достижений обучающихся;</li> <li>– планировать самостоятельно деятельность по изучению современных методик диагностики достижений учащихся и опыта их использования;</li> <li>– следить за своевременностью выполнения плана и подводить итоги</li> </ul>
Аксиологический	Отношение к деятельности в сфере компетенции и ее результату (проявление интереса, активности, организованности и ориентированности на получение результата; понимание значения результата и его самооценка)	<ul style="list-style-type: none"> <li>– проявление интереса к современным методикам и опыту диагностики достижений учащихся;</li> <li>– понимание важности диагностики достижений учащихся и использования ее результатов для внесения коррективов в учебно-воспитательный процесс с целью повышения его качества;</li> <li>– проявление намерения в будущей профессиональной деятельности использовать современные методики и технологии диагностирования достижений учащихся;</li> <li>– понимание и оценивание значения результата использования современных методик диагностирования достижений учащихся</li> </ul>

На первый взгляд кажется, что это описание является трехуровневым. Однако для того чтобы оценить структурированные элементы компетенции (представленные в столбце 3 выше приведенной таблице 1), необходимы соответствующие тесты и/или опросники. Поэтому это описание компетенции является четырехуровневым.

Примеры описания и измерения латентных переменных приведены в работах [10, 13, 14]. В целом, описание компетенций является неформализованным этапом, обработка результатов тестирования и/или опроса хорошо формализована с использованием как отечественного, так и зарубежного программно-алгоритмического обеспечения [1, 11]. Примеры анализа качества тестов как измерительных инструментов для оценивания латентных переменных приведены в [4, 12]. Чем больше индикаторов (тестовых заданий и/или пунктов опросника), тем точнее описание компетенции и тем точнее можно измерить эту латентную переменную.

### Список литературы

1. Andrich, D., Sheridan, & Luo B. G. (2005). RUMM2020: RaschUnidimensional Measurement Models software and manual. Perth, Australia, RUMM Laboratory. – 98 p.
2. Rasch G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen, Denmark: Danish Institute for Educational Research. (Expanded edition, 1980. Chicago: University of Chicago Press.). -199 p.
3. Братищенко В.В. Измерение сформированности компетенций по текущей успеваемости студентов // Наука. Информатизация. Технологии. Образование. Материалы XII международной научно-практической конференции «Новые информационные технологии в образовании и науке НИТО 2019». 2019. С. 451-4
4. Дроздов В.И., Маслак А.А., Новиков Ю.М. Использование современной теории тестологии при оценке качества АПИМ // Известия Курского государственного технического университета. 2008. № 4 (25). С. 87-95.
5. Ефремова Н. Ф. Компетенции в образовании: формирование и оценивание. Методическое пособие. Федеральный институт педагогических измерений. - Сер. Контроль и оценивание в современной системе образования. Методический портфель учителя. Донскойгосударственныйтехническийуниверситет. М. 2012. – 416 с.
6. Замолодская О.М., Корчинский А.В., Троицкий Ю.Л., Шкаренков П.П. Инструменты измерения студенческих компетенций // Вестник РГГУ. Серия: Психология. Педагогика. Образование. 2015. № 1 (144). С. 83-91.
7. Кутькина О.П., Куран М.С. Проблема разработки тестов как средств оценки уровня сформированности компетенций // Мир науки, культуры, образования. № 6 (61) 2016. С. 216-217.
8. Лопаткин Е. В. Компетентностный подход в педагогическом образовании будущих учителей //Вестник Таганрогского государственного педагогического института Изд-во: Таганрогский государственный педагогический институт им. А.П. Чехова. г. Таганрог. 2012, Номер 5. С. 44-49.
9. Маслак А.А. Теория и практика измерения латентных переменных в образовании. Монография. М., Изд-во «Юрайт», 2016. – 255 с.
10. Маслак А.А., Леус О.В., Данилов А.А. Разработка методики измерения качества образования в школе. Методические рекомендации / Государственное образовательное учреждение высшего профессионального образования Славянский-на-Кубани государственный педагогический институт; Лаборатория объективных измерений. Славянск-на-Кубани, 2009. – 67 с.
11. Маслак А.А., Моисеев С.И., Осипов С.А. Сравнительный анализ оценок параметров модели Раша, полученных методами максимального правдоподобия и наименьших квадратов // Проблемы управления. 2015. № 5. С. 58-66.
12. Маслак А.А., Поздняков С.А. Анализ качества тестовых заданий с выбором одного правильного ответа. Методические рекомендации / Государственное образовательное

- учреждение высшего профессионального образования Славянский-на-Кубани государственный педагогический институт, Лаборатория объективных измерений. Славянск-на-Кубани, 2009. – 50 с.
13. Маслак А.А., Поздняков С.А. Измерение качества выпускных квалификационных работ. Методические рекомендации / Государственное образовательное учреждение высшего профессионального образования Славянский-на-Кубани государственный педагогический институт, Лаборатория объективных измерений. Славянск-на-Кубани, 2009. – 47 с.
  14. Маслак А.А., Рыбкин А.Д. Анализ качества опросника для измерения на линейной шкале любознательности школьников // Педагогическое образование в России . 2014. № 12. С. 158-165.
  15. Шкерина Л. В., Юшипицина Е. Н. Мониторинг компетенций студентов: диагностические карты, портфолио // Высшее образование сегодня / Красноярский государственный педагогический университет им. В.П. Астафьева. М.Ю. Изд-во Российский новый университет, 2012, номер 7. С. 19-27.
  16. Шкерина Л.В. Измерение и оценивание уровня сформированности профессиональных компетенций студентов – будущих учителей математики: учебное пособие; Краснояр. гос. пед. ун-т им. В.П. Астафьева. Красноярск, 2014. – 136 с.

## COMPARATIVE ANALYSIS OF METHODS FOR MEASURING COMPETENCIES

*Maslak A.A.*

Branch of the Kuban State University at Slavyansk-on-Kuban, Slavyansk-on-Kuban  
e-mail: anatoliy\_maslak@mail.ru

**Abstract.** *In modern education, the emphasis is on evaluating learning outcomes. In each educational program, it is necessary to form dozens of competencies – universal, general professional, professional, and general cultural. An important aspect is to measure the level of development of these competencies. The assessment required in educational standards (basic, advanced, excellent) is insufficient when monitoring and analyzing factors that affect the level of competence formation. The article compares three methods for assessing the level of competence formation: in the framework of the classic theory of testing, based on indices, and in the framework of the theory of measurement of latent variables.*

**Keywords:** *competencies, latent variables, measurement, classic testing theory, theory of latent variables.*