

мерения и латентными параметрами испытуемых, выраженной в виде параметрических моделей педагогического измерения. Эта взаимосвязь обычно задается в виде специальной функции, содержащей то или иное число параметров. В зависимости от числа параметров в IRT выделяют одно-, двух- и трехпараметрические модели [6, 7]. Предполагается, что каждому испытуемому ставится в соответствие только одно значение латентного параметра, определяющего наблюдаемые результаты измерения. Требование одномерности не носит обязательного характера для тех случаев, когда измерения затрагивают проблемы оценивания компетенций.⁴

В условиях, когда оценки параметра подготовленности обучаемого и параметра трудности учебных заданий выражены в единой шкале логитов, принятой в IRT, однопараметрическая модель постулирует для обучаемого с уровнем подготовленности, равным трудности оценочного задания, вероятность правильного ответа в 50%. Для испытуемых с уровнем подготовленности намного большим, чем трудность заданий измерителей, эта вероятность стремится к единице. В случае же, когда трудность заданий намного превышает подготовленность обучаемого, вероятность правильного выполнения заданий будет стремиться к нулю.

Таким образом, в рамках основного предположения IRT устанавливается функциональная связь, в которую включены параметры двух латентных множеств. Элементы первого множества — оценки латентного параметра N испытуемых θ_i , где $i = 1, 2, \dots, N$. Второе множество образуют оценки латентного параметра β_j ($j = 1, 2, \dots, n$), равные трудности n заданий измерителя. Тогда однопараметрическая математическая модель IRT (логистическая модель Г. Раша), отображающая взаимосвязь между эмпирическими результатами измерения и оценками латентных переменных θ и β и задающая условную вероятность $P_j(\theta)$ правильного выполнения j -го задания трудностью β_j , когда θ считается независимой переменной, имеет следующий вид:

$$P_j(\theta) = \frac{e^{1,7(\theta - \beta_j)}}{1 + e^{1,7(\theta - \beta_j)}}$$

либо

$$P_j(\theta) = \{1 + \exp[1,7(\theta - \beta_j)]\}^{-1},$$

где коэффициент 1,7 является шкалирующим множителем, обеспечивающим возможность сопоставления результатов измерения, полученных на основе логистической модели с данными по другим моделям измерения.

Методика повышения качества

Наиболее сильный аргумент в пользу IRT связан с введением информационной функции на основе диф-

ференцированной ошибки измерения, которая используется для оценки эффективности измерителя и позволяет при проведении эвалюации значительно повысить качество результатов измерения.

Процесс повышения эффективности связан с подбором наиболее пригодных заданий, обеспечивающих минимальное отклонение оценки θ от параметра испытуемого. Степень пригодности принято характеризовать с помощью относительной величины, называемой информацией. По одному из определений, предложенных А. Бирнбаумом [7], объем информации, обеспеченный j -м заданием теста в данной точке θ , — это величина, обратно пропорциональная стандартной ошибке измерения данного значения θ , с помощью задания.

Достаточно важно для обеспечения объективности и эффективности управленческих воздействий правильно выбрать модель педагогического измерения в IRT.

Отдавая предпочтение классу однопараметрических моделей, можно выиграть в качестве измерений. Однако неизбежная в этой ситуации одномерность трактовки результатов измерений сузит круг управленческих решений. Выбирая двухпараметрические модели, можно получить более многогранную информацию, но проиграть в надежности измерений.

Несомненно, эти вопросы требуют систематических исследований. Они-то и проводятся автором статьи применительно к проблеме эвалюации программ и проектов, в частности по отношению к введению ФГОС в систему ВПО.

Внедрение ФГОС в эвалюации, очевидно, должно вестись в режиме мониторинга, поскольку необходимо систематически отслеживать результаты воздействия новых ФГОС на качество ВПО. Для проведения измерений должны быть выполнены определенные этапы, к числу которых можно отнести следующие:

- ♦ разработку описаний признаков проявления компетенций в операционализируемой (подающейся измерению) форме по различным направлениям профессиональной подготовки в вузах;
- ♦ разработку компетентностно-ориентированных измерителей, позволяющих надежно и валидно оценить компетенции;
- ♦ построение шкал результатов измерений интервального характера;
- ♦ проведение анализа шкалированных результатов измерений для выявления изменений (позитивных либо негативных) в качестве результатов образования;
- ♦ построение прогностических моделей эвалюации для выявления тенденций изменения качества результатов образования.

Таким образом, следует отметить в заключение, эвалюация является не только аппаратом для констатации воздействия нововведений, основанных на выполнении проектов и программ, но и эффективным прогнозирующим механизмом. Он позволяет предви-

⁴ В наши дни уже разработаны и используются, правда недостаточно широко, многомерные модели измерения, позволяющие отразить междисциплинарную природу компетенций.