

О качестве систем педагогических измерений

Корсак Константин Витальевич — творческий директор Киевского института образовательной политики, кандидат физико-математических наук

Подобно определённым фасонам одежды тесты американского типа периодически становятся объектом повышенного внимания в той или иной европейской стране. Однажды от широкомасштабного эксперимента не удержалась и Франция, отличающаяся особым консерватизмом в области образования, до сих пор сохраняющего немало инноваций наполеоновских времён. Применяв разработанные в США варианты тестов по математике, французы получили в итоге ребус, на решение которого их лучшие учёные потратили несколько месяцев: при тестировании лучшие ученики регулярно показывали пониженные результаты, худшие и средние — заметно или слегка повышенные. После дополнительных экспериментов и исследований учёные пришли к выводу, что тестовые измерения с их множеством лёгких заданий и выбором ответа из малого числа вариантов чрезмерно отличались от привычных для французских школьников и подобных нашим контрольных работ с небольшим числом заданий различной сложности и без ответов. Во Франции американские тесты не столько измеряли знания учеников по математике, сколько определяли подвижность их нервной системы, уровень способностей приспособляться к необычному и предрасположенность к риску. Итогом этих событий стал отказ руководителей системы образования Франции от массового применения на важных экзаменах стандартизованных тестов американского образца.

Если ознакомиться с применяемыми в различных странах методами слежения за прогрессом учеников и студентов путём измерения накопленных знаний и приобретённых умений и навыков (так называемыми системами *педагогических измерений*), то поражает, что фактически каждая страна делает это по-своему. Часто система педагогических измерений является самой консервативной частью системы образования, формируясь вместе с ней сотни лет. Подобно отсутствию не критикуемых систем образования, нет и идеальных систем педагогических измерений. Более того, можно даже утверждать, что в среднем в мире количество претензий к ним в течение последних десятилетий постоянно возрастает. Объяснений последнему обстоятельству столь много, что укажем лишь часть главных, не ранжируя их по глубине влияния на педагогические измерения.

Самый заметный фактор — **переход от элитарных видов образования к всеобщим и обязательным, на уровне высшей школы — к массовым, охватывающим половину и более соответствующей возрастной группы**. В рамках элитарного образования требовалось всего лишь отсечь наименее подготовленных, не заботясь об их судьбе — государственную ценность представляли лишь несколько процентов всех учащихся. С такой простой задачей справлялись привычные нам формы устных и письменных экзаменов, хотя их объективность часто была очень низкой. В современных условиях «отсекать» нельзя по той простой причине, что выбросить на улицу или направить в почти ничему не обучающие профтехучилища уже невозможно — и по законам, и «по жизни» необходимо учить практически всех и как можно дольше. Не только у нас, но и во многих других странах конфликт старых приоритетов во время всех педагогических измерений и новых высших общественно-экономических требований становится слишком острым и всё более опасным.

Почти столь же очевидно то, что удвоение с периодом 10–12 лет объёма знаний о природе и обществе в сочетании с ошибочным убеждением, что школа данного времени способна и обязана впихнуть в головы учеников основы всех накопленных на этот момент знаний, привели к весьма печальным последствиям. Учёные-педагоги, отказавшись от отбора действительно нужного, с азартом вводят новые предметы и обновляют (т.е. расширяют) программы

старых. Они преуспели — лишь отдельные школьники-уникумы удовлетворяют «педагогическим стандартам», остальным же от класса к классу учение становится всё менее интересным, а экзамены всё более ненавистными.

Ещё один фактор — изменение характера труда. Ещё столетие тому школа и вуз не имели информационных конкурентов, обеспечивая выпускников почти пожизненной информацией. Работа не требовала новых больших знаний, росли лишь производственные навыки. Необходимость изменений специальности была уникальным явлением. **Наше время — эпоха непрерывного образования, а наша школа и вуз готовят к чему угодно, но не к этой жизненно важной деятельности.** Рынок труда в лице работодателей (преимущественно из частного сектора) уже давно требует новых акцентов в обучении. Отметим мимоходом, что **по важности для успешной современной карьеры абсолютный объём знаний выпускника вуза откатился на третье или четвёртое место среди тех его характеристик, которых жаждут нынешние работодатели. Лидируют — умение быстро и эффективно учиться, желание совершенствоваться, способность «держат удар» и адаптироваться к новому.** Но именно эти-то качества и умения современные методы педагогических измерений определяют исключительно плохо, так как всегда совершенствовались с целью точного измерения объёма знаний. Ещё хуже то, что вся наша система обучения не ориентирована на формирование новых качеств у выпускников школ и вузов.

Известно, что обострение потребности — главный побудительный мотив начала серьёзной деятельности человека разумного. Вот и тесты американцы изобрели не как рождественский подарок. После раздумий и опытов они решили, что стандартизированные тесты являются их последней надеждой определить уровень знаний школьников и студентов, так как каждое учебное заведение в США имеет почти неограниченные возможности применять собственный набор предметов, не говоря уж о содержании их программ. Нам, учившимся по немыслимо стандартизированным учебникам, просто невозможно вообразить масштаб американской свободы и все её положительные и отрицательные последствия для образования и экономики.

Но вернёмся к потребности усовершенствовать педагогические измерения и приблизить их возможности к приоритетам современного общества. Возникла специальная наука, фигурирующая пока под разными названиями (квалиметрия, доцимология и др.). Главные её достижения ещё впереди, укажем лишь основное, касающееся критериев качества педагогических измерений.

Среди главных выделяются четыре — **объективность, надёжность, валидность и точность.**

Объективность экзамена или других измерений — независимость результатов от состояния, личностных качеств и прочих характеристик тех, кто их готовит и проводит. Лучший способ обеспечения объективности — стандартизация экзаменов. Истинная объективность касается всех стадий и процедур, включая использование и распространение результатов экзаменов.

Надёжность педагогического измерения характеризует устойчивость его результатов, их высокую повторяемость на дополнительных экзаменах. В педагогике обеспечить абсолютную надёжность невозможно хотя бы из-за большого влияния на результаты измерений состояния ученика или студента, хотя стремиться к максимальной надёжности можно и нужно.

Валидность измерения — его комплексная характеристика, являющаяся мерой получения в результате экзамена данных именно о том, что планировалось определить и измерить. Следует обеспечить надлежащее (валидное) содержание вопросов и задач, пригодность методов и средств, правильное планирование и прогноз результатов.

Степень точности измерения определяет систематическую ошибку, зависящую в первую очередь от особенностей применяемого метода. Значительную помощь педагогам может оказать математика (теория ошибок и статистика).

Читателям могут быть интересны выжимки из таблиц, составленных автором для всех

применяемых в наше время форм педагогических измерений, — наблюдения, устного опроса, письменного экзамена, интервью, стандартного и нестандартного тестирования. Ограничимся лишь двумя из них — письменным экзаменом и стандартизированным тестированием, приводя положительные и отрицательные аспекты.

ПИСЬМЕННЫЙ ЭКЗАМЕН

Положительные аспекты

1. Высокий уровень стандартизации условий проведения экзамена
2. Приемлемый уровень влияния субъективных факторов в процессе выполнения работы
3. Отличная сохраняемость итогов
4. Возможность охватить большие группы учащихся
5. Получение данных об общем развитии учащихся
6. Возможность измерения проблемного мышления, общей грамотности

Отрицательные аспекты

1. Большие затраты времени на проверку работ учеников
2. Чрезмерная субъективность во время интерпретации работ, низкая итоговая объективность измерения
3. Низкий уровень надёжности
4. Недостаточная общая валидность измерения
5. Низкая точность измерения при допустимой длительности экзамена. Повысить её можно только проводя слишком длительные экзамены

СТАНДАРТНОЕ ТЕСТИРОВАНИЕ

Положительные аспекты

1. Достаточно высокая объективность процесса педагогического измерения и интерпретации его итогов
2. Возможность стандартизации педагогического измерения
3. Достаточный уровень надёжности экзаменов
4. Хороший уровень валидности измерения
5. Приемлемая точность при введении новой (расширенной) шкалы оценок
6. Малые затраты времени на массовые измерения
7. Хорошая сохраняемость итогов и лёгкость их компьютерной обработки.
8. Содействие необходимой интеграции в образовательное пространство Земли

Отрицательные аспекты

1. Необходимость больших изменений в нашей системе образования, включая психологию и личностные приоритеты
2. Необходимость новых учебников и другого учебного материала
3. Большие затраты денег и времени на подготовку тестового материала
4. Сложность преодоления инертности старой системы измерений
5. Нехватка специалистов по тестированию
6. Затраты на частое обновление тестового материала
7. Негативные воспоминания о прежних попытках внедрения тестов

В заключение приведём одно сравнение — абсолютизация тестов как единственной формы педагогических измерений подобна попытке принудить нас всех питаться одним и тем же блюдом.