

# ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ЛОГИСТИКА В СТРОИТЕЛЬСТВЕ. ОРГАНИЗАЦИЯ ВЫСШЕГО ОБРАЗОВАНИЯ В ОБЛАСТИ СТРОИТЕЛЬСТВА И АРХИТЕКТУРЫ. ДОПОЛНИТЕЛЬНОЕ ОБРАЗОВАНИЕ И ПЕРЕПОДГОТОВКА КАДРОВ В СТРОИТЕЛЬНОЙ ОТРАСЛИ

УДК 372.862:004

DOI: 10.22227/2305-5502.2019.2.11

## Компьютерное тестирование как эффективная форма объективного контроля знаний при изучении технических дисциплин

**В.С. Кузнецов, Г.Е. Полехина, Ю.А. Шапошникова**

*Национальный исследовательский Московский государственный строительный университет  
(НИУ МГСУ), 129337, г. Москва, Ярославское шоссе, д. 26*

### АННОТАЦИЯ

**Введение.** Объективный и регулярный контроль знаний обучающихся в области технических дисциплин может выполняться с помощью специализированных тестов, обеспечивающих требуемый уровень изучения материала и надежность приобретенных знаний. Рассмотрены различные аспекты применения тестирования в учебном процессе.

**Материалы и методы.** Используемые в практике обучения тесты должны отвечать определенным требованиям: валидность, определенность, простота, однозначность, надежность. Выделение уровней освоения дает возможность диагностировать и совершенствовать учебный процесс и степень освоения компетенций обучающимися. В зависимости от оценки предстоящей деятельности по характеру изучения различают четыре уровня усвоения учебного материала. Тесты I уровня включают тесты на опознание, различение и классификацию. Тесты II уровня предназначены для контроля усвоения материала на уровне «репродукции», позволяющего воспроизводить по памяти и анализировать информацию, решать типовые задачи. Тесты III уровня включают специальные задачи, требующие от обучающегося поисковой деятельности, для которой не существует готовых алгоритмов, а решения ведут к получению субъективно новой информации. Тесты IV уровня выявляют умение учащихся принимать решения в новых проблемных ситуациях, решение которых результат творческой деятельности и сопровождается получением объективно новой информации.

**Результаты.** Для создания эффективной системы тестового контроля по определенной дисциплине требуется создать основные условия, такие как: база данных, выборка необходимого числа заданий, не менее 30 и не более 70, таймер в соответствии с трудоемкостью заданий, оценка заданий и критерии оценок, вывод результатов.

**Выводы.** При наличии необходимого числа компьютеров, в количестве минимум одного на двух обучающихся, правильно поставленное компьютерное тестирование значительно сокращает затраты времени на контрольное мероприятие, повышает ответственность и успеваемость студентов, гарантирует объективность контроля и исключает конфликтные ситуации.

**КЛЮЧЕВЫЕ СЛОВА:** валидность, дисциплина, знание, контроль, надежность, объективность, степень усвоения, тест

**ДЛЯ ЦИТИРОВАНИЯ:** Кузнецов В.С., Полехина Г.Е., Шапошникова Ю.А. Компьютерное тестирование как эффективная форма объективного контроля знаний при изучении технических дисциплин // Строительство: наука и образование. 2019. Т. 9. Вып. 2. Ст. 11. URL: <http://nso-journal.ru>. DOI: 10.22227/2305-5502.2019.2.11

## Computerized testing as efficient form of objective knowledge monitoring in studying of technical subjects

**Vitaly S. Kuznetsov, Galina E. Polekhina, Yulia A. Shaposhnikova**

*Moscow State University of Civil Engineering (National Research University) (MGSU),  
26 Yaroslavl'skoe shosse, Moscow, 129337, Russian Federation*

### ABSTRACT

**Introduction.** Objective and regular students' knowledge monitoring in technical subjects can be implemented by means of special tests allowing for the required mastering level of the matter and the reliable consolidation of the acquired knowledge. Various aspects of the application of tests in the academic activity were considered.

**Materials and methods.** Tests used in practical studies should meet specific requirements, such as: validity, definiteness, simplicity, unambiguity, reliability. The identification of mastering levels makes it possible to “troubleshoot” and to improve the academic activity and the mastering degree of the competences by the students. Based on the assessment of the studying pattern of the forthcoming activity, one could point out four mastering levels of the subject matter. Level I tests include recognition, discrimination and classification. Level II tests monitor the mastering of the subject in the level of “reproduction” allowing for retrieval of information from the memory and its analysis, for routine assignment solutions. Level III tests impose special assignments challenging a student with quests for which no ready algorithms are catered, whereas the solutions found lead to obtaining of subjectively new information. Level IV tests reveal students’ capability to take decisions in new problematic situations, the solutions found, being a result of creative activity, are followed by obtaining of objectively new information.

**Results.** To establish an efficient system of monitoring tests in a certain subject, a number of basic prerequisites is required, such as a data base, a sample group of with the required number of assignments, at least 30 and maximum 70, a time limit in accordance with the required labor intensity, assessment of the assignments and its criteria, the output of the results.

**Conclusions.** If there is a required number of computers of at least one PC per two students, correctly arranged computer testing considerably reduces time demand of a monitoring event, increases the responsibility and the progress of the students, guarantees the objectiveness of the knowledge monitoring and helps to avoid conflicts.

**KEYWORDS:** validity, subject, knowledge, monitoring, reliability, objectiveness, degree of mastering, test(ing)

**FOR CITATION:** Kuznetsov V.S., Polekhina G.E., Shaposhnikova Y.A. Computerized testing as efficient form of objective knowledge monitoring in studying of technical subjects. *Stroitel'stvo: nauka i obrazovanie* [Construction: Science and Education]. 2019; 9(2):11. URL: <http://nso-journal.ru>. DOI: 10.22227/2305-5502.2019.2.11 (rus.).

## ВВЕДЕНИЕ

Цель данной работы — выявление особенностей составления тестов и их применения для текущего и рубежного контроля знаний обучающихся при изучении технических дисциплин.

Вопрос изучения возможностей применения персональных компьютеров для оценки качества знаний стал актуален достаточно давно — с появлением первых ПК и развитием компьютерной техники. Это направление постоянно совершенствуется и изучается российскими и зарубежными авторами [1–12]. Условно его можно разделить на три основные группы: совершенствование методов и алгоритмов тестового контроля [13–15], корректность их применения [14–17] и проблемы объективности контроля качества знаний обучающихся [3–4, 18–24].

Анализ учебного процесса с современных позиций показывает, что существенным недостатком в комплексе задач подготовки обучающихся различных уровней образования (бакалавров, специалистов, магистров) является, как правило, неэффективный контроль знаний как промежуточный, так и окончательный со стороны преподавателя и практически полное отсутствие системы самоконтроля для обучающихся [4, 18, 21].

Традиционный контроль в форме устного или письменного экзамена или зачета при существующих в настоящее время нормативах времени не позволяет всесторонне оценить знания обучающегося, а новые условия обуславливают необходимость получения оперативной, объективной, независимой

информации о результатах образовательной деятельности [1, 10–11, 18, 21, 23].

Применение тестов в качестве инструмента контроля дает возможность устранить названные недостатки и таким образом улучшить качество учебного процесса. Так, регулярное отслеживание качества знаний, умений и навыков в форме тестирования исключает влияние субъективных факторов на результат, не допускает принижения личности студента, сокращает время на контрольные мероприятия, позволяет вести самоконтроль. Правильно поставленные тесты несут не только диагностическую, обучающую, организующую, но и воспитательную функции.

Кроме того, по результатам тестирования преподаватель может оценить, насколько высока планка подачи учебного материала (степень абстракции), уровень требований при традиционном контроле и оперативно вносить соответствующие коррективы.

## МАТЕРИАЛЫ И МЕТОДЫ

Используемые в практике обучения тесты должны отвечать определенным требованиям, это — *валидность, определенность, простота, однозначность, надежность*:

- *валидность* теста — это соответствие теста смыслу и содержанию изучаемого материала и уровню тестирования;

- *определенность* теста предполагает конкретное понимание студентом требования теста, какие знания и в каком объеме он должен продемонстрировать;

- *простота* заключается в четкой и прямолинейной формулировке задания;
- *однозначность* теста предполагает, что задания не должны допускать двойного толкования;
- *надежность* теста состоит в обеспечении устойчивости результатов тестирования испытуемого.

В зависимости от оценки предстоящей деятельности по характеру изучения различают несколько уровней усвоения учебного материала [1]:

- уровень I — (знания–знакомства), узнавание объектов, свойств, процессов и др. в конкретной области знания на основе предшествующего обучения;
- уровень II — (знания–копии), воспроизведение информации, операций и др. путем самостоятельного применения типовых правил деятельности;
- уровень III — продуктивная (эвристическая) деятельность. Выполняется не по однозначным правилам, а основывается на догадке. В этом случае добывается субъективно новая информация путем трансформации ранее известного решения или создания нового;
- уровень IV — продуктивная творческая деятельность за счет самостоятельного создания нового решения или новой схемы деятельности.

По данным, изложенным в статье [18], соотношение требований от I до IV уровня составляет 1:4:8:16, что свидетельствует не только о степени абстракции и сложности учебного материала, но и о количественном соотношении обучающихся, способных анализировать и усвоить информацию того или иного уровня.

Выделение уровней освоения дает возможность диагностировать и совершенствовать учебный процесс и степень освоения компетенций обучающихся.

Степень освоения учебного материала  $K_a$  на конкретном уровне освоения количественно выражается формулой

$$K_a = A/P, \quad (1)$$

где  $A$  — число правильно выполненных задач или существенных операций;  $P$  — общее число задач или существенных операций.

Понятие «существенная операция» служит характеристикой трудоемкости отдельного задания и зависит от количества вариантов ответов, которые должен проанализировать обучающийся при ответе на поставленный вопрос.

Уровень сформированного навыка (степень усвоения)  $K_T$  характеризуется коэффициентом автоматизации

$$K_T = T_{\text{ст}}/T_{\text{ст}}, \quad (2)$$

где  $T_{\text{ст}}$  — время, необходимое группе специалистов (профессор, доцент, старший преподаватель, преподаватель) для выполнения того же задания;  $T_{\text{ст}}$  — время, необходимое студенту для выполнения задания.

Нормативное значение коэффициента  $K_T$  зависит от вида деятельности и задается в пределах 0,5...1,0 [18]. Для опасных видов деятельности, например, кораблестроение, автомобилестроение или авионавигация —  $K_T \rightarrow 1,0$  [18].

Для деятельности, связанной с проектированием и возведением объектов жилищного и гражданского назначения, в частности, для бакалавров, обучающихся по направлению 08.03.01 «Строительство», коэффициент  $K_T$  рекомендуется назначать в пределах 0,65...0,7.

Для оценки качества усвоения создается совокупность тестов одного уровня (пакет тестов), предъявляемая испытуемому для проверки. Количество заданий в пакете зависит от уровня теста, степени усвоения материала и заданной надежности теста.

**Тесты I уровня** включают тесты на опознание, различение и классификацию.

*Тесты на опознание* имеют всегда одну существенную операцию, как вывод из альтернативы «да–нет». Например, бетон хорошо сопротивляется сжатию и плохо растяжению. Эталонный ответ «да» содержит одну существенную операцию.

*Тесты на различение* содержат всегда большее число существенных операций, так как обладают несколькими вариантами ответов, затрудняющими выбор правильного решения. Например, укажите правильное выражение для закона Гука

$$1. \sigma = E\varepsilon; 2. E = \varepsilon\sigma; 3. \varepsilon = \sigma E; 4. \sigma = E\varepsilon/2.$$

Эталон 1 — да, 2 — нет, 3 — нет, 4 — нет. Число существенных операций — 4.

Тесты на различение могут содержать и несколько правильных ответов из общего списка ответов.

*Тесты на классификацию* содержат многократный выбор ответов и число существенных операций равно числу соотношений. Например, указать правильный ответ (ответы) определения расчетной ширины свеса полки  $c_f$  для балки таврового сечения.

$$а) \text{ при } h'_f \geq 0,1h; б) \text{ при } 0,05h \leq h'_f < 0,1h.$$

Варианты ответов: 1.  $c_f = 6h'_f$ ; 2.  $c_f = 3h'_f$ ; 3.  $c_f = l_0/6$ .

Эталоны: а — 1,3; б — 2,3. Число существенных операций равно 6.

**Тесты II уровня** предназначены для контроля усвоения материала на уровне «репродукции», т.е. позволяющего воспроизводить по памяти и анализировать информацию, решать типовые задачи без подсказки. Основные типы этого уровня: тесты–подстановки, конструктивные тесты, типовые задачи.

*Тесты–подстановки* — наиболее простые из тестов II уровня, в которых намеренно пропущены или добавлены формула, слово или какой-либо другой существенный элемент. Например, перечислить классы по прочности на осевое растяжение для тяжелого бетона.  $B_1 0,8$ ;  $B_1 1,2$ ;  $B_1 1,6$ ;  $B_2 0$ ;  $B_2 4$ ;  $B_2 8$ ;  $B_3 2$ ;  $B_4 0$ . Число существенных операций — 8.

*Конструктивные тесты* требуют воспроизведения по памяти, без какой-либо помощи. На-

пример, для изгибаемого элемента прямоугольного сечения с одиночной арматурой воспроизвести формулу прочности по бетону. Эталон  $M \leq R_b \cdot b \cdot x \cdot (h_0 - 0,5 \cdot x)$ . Число существенных операций определяется суммой количества символов, индексов, математических знаков. Для приведенного примера число операций равно 18.

Типовые задачи позволяют по памяти изменять известную процедуру решения и получить необходимый ответ. Например, найти сумму координат точки пересечения прямой  $(x - 2)/1 = (y + 1)/3 = (z - 4)/2$  и плоскости  $x + 2y - 2z - 1 = 0$ . Варианты ответов 1) 13; 2) 15; 3) 20; 4) 20; 5) 23. Типовые задачи широко используются в практике обучения. При деятельности на первых двух уровнях студент не получает новой информации, а только демонстрирует знание изученного материала и способность применять приобретенные навыки в типовых условиях.

Тесты III уровня включают специальные задачи, требующие от студента поисковой деятельности, для которой не существует готовых алгоритмов, а решения ведут к получению субъективно новой информации. Оценка тестирования проводится путем сравнения с эталонным решением или экспертным путем.

Тесты IV уровня выявляют умение учащихся принимать решения в новых проблемных ситуациях, решение которых результат творческой деятельности и сопровождается получением объективно новой информации. Эталоном является созданная экспертами-специалистами методика решения проблемы.

В учебном процессе в вузах распространены тесты I и II уровней, а тесты III и IV уровней применяются для олимпиад, конкурсов городского, регионального или международного значения [4].

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Необходимое количество тестовых заданий можно определить в зависимости от заданной на-

дежности оценки  $P$  и отношения  $q = \varepsilon/s$ , где  $s$  — эмпирический стандарт ошибки;  $\varepsilon$  — доверительная оценка точности измерения, определяемая в соответствии с распределением Стьюдента [4, 25]. Так, для обеспечения надежности тестирования с достоверностью  $P = 70\%$  при одинаковой заданной точности измерений минимальное число заданий должно быть не менее 30. При увеличении количества заданий увеличивается надежность доверительных оценок и устойчивость результатов тестирования. На практике при назначении числа тестов I и II уровней можно пользоваться графиком, представленным на рис. 1, который позволяет просто определить требуемое число заданий  $n$  для обучающихся в тесте.

На рис. 1 показана зависимость надежности результатов тестирования от числа существенных операций (числа переборных), аппроксимированная логарифмической зависимостью типа  $P = 42,481 \ln(n) + 2,0438$ , полученной методом наименьших квадратов в соответствии с работой [25].

Из графика видно, что при количестве заданий  $n = 35$  надежность результатов составляет 71%, а при числе заданий  $n = 70$  надежность равна 92%.

Компьютерное тестирование применялось с 1995 по 2005 гг. на областном факультете ПГС, (ныне Мытищинский филиал МГСУ) по металлическим, железобетонным и деревянным конструкциям на компьютерах А3000 с операционной системой RICS OS, позволяющей использовать язык программирования BASIC BBC [19]. Использование тестов дало возможность повысить ответственность и успеваемость обучающихся, о чем свидетельствовали первые места филиала по успеваемости в этот период среди других подразделений университета [4]. С 2006 по 2015 гг. тестирование применялось на кафедре прикладной механики и математики по дисциплине «Сопrotивление материалов», где также отмечался положительный эффект [26].

В настоящее время тестирование применяется для контроля знаний обучающихся в дистанционных курсах в НИУ МГСУ.

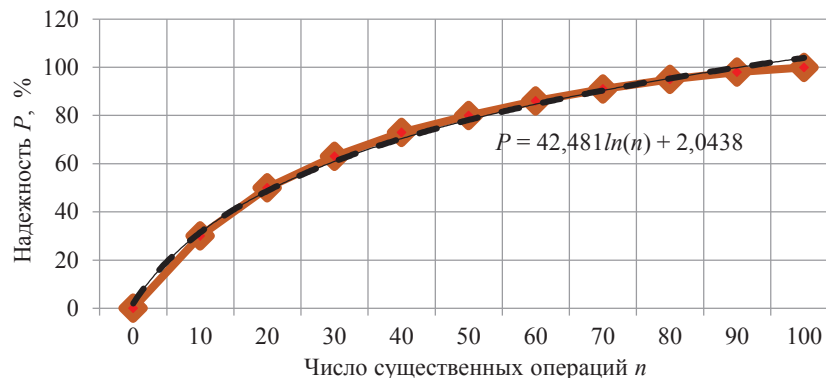


Рис. 1. Зависимость надежности тестов от числа существенных операций



## ЗАКЛЮЧЕНИЕ И ОБСУЖДЕНИЕ

Для создания эффективной системы тестового контроля по определенной дисциплине необходимо:

- создать базу данных, содержащую такое количество заданий, которое позволяло бы полностью покрыть изучаемый материал и исключало бы повторение заданий для одного испытуемого;
- с помощью генератора случайных чисел установить выборку необходимого числа заданий (для обеспечения требуемой надежности теста количество заданий должно быть не менее 30, а исходя из лимита времени на проведение тестирования не более 70);
- настроить таймер в соответствии с трудоемкостью заданий. Время выполнения устанавливается экспертным путем;
- установить оценку отдельного задания как отношение числа правильно выполненных существенных операций к нормативному числу существенных операций;
- вычислить процент правильных ответов в пакете;
- установить критерии оценок (например, 70–79 % — удовлетворительно, 80–89 % — хорошо, 90–100 % — отлично);

- в случае превышения лимита времени тест считается невыполненным — оценка неудовлетворительная;

- вывести результаты на экран или лучше распечатать.

Таким образом, при наличии необходимого числа компьютеров (в количестве минимум одного на двух обучающихся), правильно поставленное компьютерное тестирование значительно сокращает затраты времени на контрольное мероприятие, повышает ответственность и успеваемость обучающихся, гарантирует объективность контроля и практически исключает конфликтные ситуации.

По результатам тестирования преподаватель может судить о слабых и сильных сторонах учебного процесса, а также имеет возможность оперативно влиять на учебный процесс.

Применение и расширение тестирования при изучении различных дисциплин в НИУ МГСУ способствовало бы более глубоким и устойчивым знаниям обучающихся.

## ЛИТЕРАТУРА

1. Смирнова Е.Э. Оценка качества образования: подходы и практика // Оценка качества образования в российских вузах : сб. тр. СПб., 2014.

2. Янченко С.И. Независимый тестовый контроль и его место в системе контроля уровня учебных достижений студентов // Новые образовательные технологии в вузе : сб. тез. докл. участников конференции. Екатеринбург : Уральский федеральный университет им. первого Президента России Б.Н. Ельцина, 2014. С. 1564–1569.

3. Шейн А.А., Привалов Н.И. Методика оценки качества знаний студентов при тестовом контроле усвоения материала // Известия Волгоградского государственного технического университета. Сер.: Новые образовательные системы и технологии обучения в вузе. 2010. Т. 7. № 8 (68). С. 215–217.

4. Кузнецов В.С., Посвятенко Ю.В. Объективный контроль качества знаний студентов в условиях компетентностного подхода к образованию // Гуманитарное образование в строительном вузе. Компетентностный подход и современные педагогические технологии : сб. науч.-метод. мат. факультета гуманитарного образования. М. : МГСУ, 2007. Вып. 4. С. 255–279.

5. Осипов А.В. Эффективная оценка доли знаний при тестовом контроле // В мире научных открытий. 2010. № 4–11 (10). С. 127–128.

6. Bolysbayev D., Almetov N. The technology of applying the didactic potential of ethnoculture at

school // The Scientific Journal of the Modern Education & Research Institute. 2017. Pp. 19–29.

7. Honour E., Workinger S. Principles of test & evaluation // A Workshop in Two Days. 2013. Vol. 2.1.

8. L'Horty Y., Petit P. Évaluation aléatoire et expérimentations sociales // Revue Française d'Économie. 2011. Vol. XXVI. Issue 1. Pp. 13–48. DOI: 10.3917/rfe.111.0013

9. Nizonkiza D., Van De Poel K. Testing controlled productive knowledge of adverb-verb collocations in junior researchers using English as a foreign language // Stellenbosch Papers in Linguistics. 2016. Vol. 46. Pp. 99–119. DOI: 10.5774/46-0-235

10. Kahlert J., Thomsen R., Pedersen L., Madsen M., Moller, N., Beck-Nielsen H., Sorensen H. Testing the robustness of background knowledge confounder selection in a case-control study // Pharmacoepidemiology and drug safety. 2014. Vol. 23. Pp. 96–97.

11. Čaušević A., Sundmark D., Punnekkat S. Impact of test design technique knowledge on test driven development: a controlled experiment // Lecture Notes in Business Information Processing. 2012. Vol. 111. Pp. 138–152. DOI: 10.1007/978-3-642-30350-0\_10

12. Брюханова Т.Н., Кирюшин А.В. О тестовом контроле знаний по физике студентов заочного факультета ТОГУ // Ученые заметки ТОГУ. 2013. Т. 4. № 4. С. 1298–1300.

13. Дубан Р.Н. Построение профилей IRT в тестовом контроле знаний // Актуальные проблемы

гуманитарных и естественных наук. 2014. № 3–1. С. 52–57.

14. *Атоев Э.Х.* Некоторые аспекты применения компьютерной техники при тестовом контроле знаний // Молодой ученый. 2016. № 21 (125). С. 849–850.

15. *Shyuka S.* Problem of optimization of computer test control of educational achievements of students in ukrainian language (for professional purposes) // Information Technologies and Learning Tools. 2017. Vol. 60. Issue 4. Pp. 288–299. DOI: 10.33407/itlt.v60i4.1783

16. *Aleksenko O., Baranova I., Vashchenko S., Sobol A.* Information system of students' knowledge test control // Communications in Computer and Information Science. 2015. Vol. 538. Pp. 104–114. DOI: 10.1007/978-3-319-24770-0\_10

17. *Khidirova Ch.M.* Methods and algorithms of determination of complexity of test questions for formation a database system of the adaptive test-control of knowledge // International conference on information science and communications technologies (ICISCT). 2017. DOI: 10.1109/icisct.2017.8188572

18. *Зенкина С.В.* Компьютерное тестирование при оценке уровня учебных достижений студентов // Стандарты и мониторинг в образовании. 2017. № 1. С. 30–32.

19. *Кузнецов В.С., Малахова А.Н.* Опыт применения персональных компьютеров в преподавании

дисциплины железобетонные и каменные конструкции // Основные результаты и перспективы высшего строительного образования : сб. науч.-метод. разработок. М. : МГСУ, 2001. С. 270–280.

20. *Киреев Б.Н.* Использование мультимедиа технологий при изучении технических дисциплин // Альманах современной науки и образования. 2013. № 10 (77). С. 69–73.

21. *Александрова Л.В.* Тестовый контроль на уроках обществознания как средство обучения и контроля (из опыта работы) // Симбирский научный вестник. 2015. № 1 (19). С. 50–52.

22. *Bon A., Cros F.* Les evaluations en education au niveau international: impacts, contradictions, incertitudes. CIEP. 2006. 172 p.

23. *Wiggins G.P.* Assessing student performance: exploring the purpose and limits of testing. 1st ed. San Francisco : Jossey-Bass Publishers, 1993. 336 с.

24. *Aiken L.S., West S.G.* Multiple regression: Testing and Interpreting interactions. SAGE Publications, Inc. 1991. DOI: 10.1037/0021-9010.84.6.897

25. *Румицкий Л.З.* Математическая обработка результатов эксперимента. М. : Наука, 1971. 192 с.

26. *Агапов В.П., Кузнецов В.С., Голованов Р.О.* Компьютерное тестирование по дисциплине сопротивление материалов. М. : МГСУ, 2007. С. 15–20.

Поступила в редакцию 8 ноября 2018 г.

Принята в доработанном виде 27 ноября 2018 г.

Одобрена для публикации 27 марта 2019 г.

**О Б АВТОРАХ:** **Кузнецов Виталий Сергеевич** — кандидат технических наук, доцент, доцент кафедры архитектурно-строительного проектирования, **Национальный исследовательский Московский государственный строительный университет (НИУ МГСУ)**, 129337, г. Москва, Ярославское шоссе, д. 26, umzmf@mgsu.ru;

**Полехина Галина Евгеньевна** — кандидат педагогических наук, доцент, доцент кафедры прикладной механики и математики, **Национальный исследовательский Московский государственный строительный университет (НИУ МГСУ)**, 129337, г. Москва, Ярославское шоссе, д. 26, PolehinaGE@mgsu.ru;

**Шапошникова Юлия Александровна** — кандидат технических наук, доцент кафедры железобетонных и каменных конструкций, **Национальный исследовательский Московский государственный строительный университет (НИУ МГСУ)**, 129337, г. Москва, Ярославское шоссе, д. 26, gbk@mgsu.ru.

## INTRODUCTION

It is the purpose of this paper to reveal specific features of tests compilation and use for current and milestone knowledge monitoring of students studying technical subjects.

The problem of the feasibility study of the use of personal computers for knowledge quality assessment has been vital since long ago, that is since the very first PCs appeared, and the computer technology started to progress. This sphere is constantly experiencing im-

provements and being studied by Russian and foreign authors [1–12]. As a convention, it can be subdivided into three basic groups: improvement of methods and algorithms of testing [13–15], correctness of their application [14–17] and problems of the objectiveness of students' knowledge quality monitoring [3–4, 18–24].

Analyses of the academic activity from contemporary points of view demonstrate that the essential drawback in the comprehensive mission of the education of students of different education levels (Bachelors, Specialists, Masters) is, as a rule, inefficient knowledge

monitoring, both intermediary and final, by the tutor and almost complete absence of a self-check system for the students [4, 18, 21].

A traditional monitoring event, such as oral or written exam or credit, scarcely allows for a comprehensive evaluation of a student's knowledge due to the current time limits, whereas the new realities stipulate a necessity for prompt, objective, independent information on the results of the educational activity [1, 10–11, 18, 21, 23].

Using of tests as a monitoring tool provides for the elimination of the above-stated drawbacks, thus contributing to the quality improvement of the academic activity. For example, regular monitoring of the quality of the knowledge, the capacities and the skills in the form of testing provides for avoidance of the influence of subjective factors on the result, of any detraction of a student's personality, contributes to the reduction of the time demand of the monitoring events, allows for self-checking. Correctly arranged testing bears not only "troubleshooting", teaching, organizing functions but also the pedagogical one.

Besides, based on the testing results, the tutor can evaluate the degree of refinement the subject is presented with (degree of abstraction), the level of requirements in the traditional monitoring in order to respond promptly with corresponding corrections.

## MATERIALS AND METHODS

Tests used in the education practice shall comply with certain requirements, these are *the validity, the definiteness, the simplicity, the unambiguity, the reliability*:

- *the validity* of a test is the compliance of the test to the sense and the contents of the subject being studied and to the testing level;
- *the definiteness* of a test supposes student's precise understanding of the requirements imposed by the test, that is which knowledge and in which quantity he (she) should demonstrate;
- *the simplicity* supposes a clear and plain assignment statement;
- *the unambiguity* of a test supposes that assignments admit no multiple interpretation;
- *the reliability* of a test is the feature providing for the stability of the results demonstrated by the testee.

Based on the assessment of the forthcoming activity in terms of the studying pattern, there are several mastering levels of the subject matter [1].

- level I (acquaintance knowledge) comprises recognition of objects, properties, processes, etc. in a specific sphere of knowledge on the background of previous studies;
- level II (replica knowledge) comprises reproducing of information, operations, etc. by means of autonomous application of the framework activity rules;
- level III (productive (heuristic) activity). It is not carried out in accordance with unambiguous rules but

is based on speculation. In this case, subjectively new information is obtained, by means of transforming of a previously known solution or a creation of a new one;

- level IV is productive creative activity due to independent development of a new solution or a new activity pattern.

According to the data of the paper [18], the requirements ratio of the levels from I to IV looks like 1:4:8:16, being not only the evidence of the abstraction degree and the complex nature of the subject matter but also that of the proportion of students capable of analyzing and mastering of the information of the specific level.

The identification of the mastering levels makes it possible to "troubleshoot" and to improve the academic activity and the competence mastering degrees of the students.

The *mastering degree of the subject matter*  $K_a$  is quantitatively expressed on a specific mastering level by means of the following formula

$$K_a = A/P, \quad (1)$$

thereby  $A$  is the number of correctly fulfilled assignments or significant transactions;  $P$  is the total number of the assignments or significant transactions.

The notion "significant transaction" serves to characterize the labor intensity of a specific assignment dependent upon the multitude of the solution alternatives to be analyzed by the student while answering the posed question.

The *level of the acquired skill* (mastering degree)  $K_i$  is characterized by the automation factor

$$K_i = T_{sp}/T_{st} \quad (2)$$

thereby  $T_{sp}$  is the time demand of a group of specialists (professor, assistant professor, senior professor, tutor) to process the same assignment;  $T_{st}$  is the time demand of a student to fulfill the assignment.

The rated value of  $K_i$  factor is dependent on the type of activity and is set within the limits of 0.5...1.0 [18]. For highly safety relevant activities, such as shipbuilding, motor vehicle manufacturing or aircraft navigation,  $K_i \rightarrow 1.0$  [18].

For activities associated with the design and the erection of residential and civil construction projects, in particular, for bachelors with subject 08.03.01 "Construction", the value of  $K_i$  is recommended to be assigned within 0.65...0.7.

*For assessment of the mastering quality* a multitude of tests of one and the same level is compiled (a test package) and offered to the testee for a check. The number of assignments in the test package depends upon the test level, the mastering degree of the subject matter and the assigned test reliability.

**Level I tests** include recognition, discrimination, and classification tests.

*Recognition tests* always contain one significant transaction like the choice between "yes" or "no". E.g., concrete features good resistance to compression and

poor resistance to extension. The reference answer “yes” contains one significant transaction.

*Discrimination tests* always contain a higher number of significant transactions since they offer several answer alternatives making correct choice a difficult issue. E.g., indicate the correct expression for Hooke’s law

1.  $\sigma = E\varepsilon$ ; 2.  $E = \varepsilon\sigma$ ; 3.  $\varepsilon = \sigma E$ ; 4.  $\sigma = E\varepsilon/2$ .

Reference 1 — yes, 2 — no, 3 — no, 4 — no. The number of significant transactions is 4.

Discrimination tests may also contain several correct answers from the full answers list.

*Classification tests* contain a multiple selection of answers with the number of significant transactions equal to that of the ratios. E.g., indicate correct answer(s) for the design flange overhung width  $c_f$  of an T-beam.

- a) at  $h'_f \geq 0.1h$ ; b) at  $0.05h \leq h'_f < 0.1h$ .

Alternatives of the answer: 1.  $c_f = 6h'_f$ ; 2.  $c_f = 3h'_f$ ; 3.  $c_f = l_0/6$ .

References: a — 1.3; b — 2.3. The number of significant transactions is 6.

**Level II tests** are suitable for monitoring of the mastered subject matter in the “reproduction” level, that is, making it possible to retrieve from the memory and to analyze information, to solve routine assignments without help. The basic test types of this level are substitution tests, construing tests, routine assignments.

The *substitution tests* are the simplest of Level II tests with deliberately omitted or added formula, word or any other essential element. E.g., list strength classes of axial extension for heavy concrete.  $B_10.8$ ;  $B_11.2$ ;  $B_11.6$ ;  $B_12.0$ ;  $B_12.4$ ;  $B_12.8$ ;  $B_13.2$ ;  $B_14.0$ . The number of significant transactions is 8.

The *construing tests* require information to be reproduced from the memory without any help. E.g., reproduce the concrete strength formula for a bent element of rectangular cross-section with a single reinforcement rod. The reference is  $M \leq R_b \cdot b \cdot x \cdot (h_0 - 0.5 \cdot x)$ . The number of significant transactions is determined by the sum of the number of symbols, indexes, math signs. For the example above, the number of operations is 18.

The *routine assignments* presume retrieval of a known solution procedure formula from the memory to obtain the correct answer selection. E.g., find the sum of the coordinates of the crossing point of the line  $(x - 2)/1 = (y + 1)/3 = (z - 4)/2$  with the plane  $x + 2y - 2z - 1 = 0$ . The answer variants are 1) 13; 2) 15; 3) 20; 4) 20; 5) 23. Routine assignments are widely used in the practical studies. The activities in the first two levels do not offer any new information to the student, being pure demonstration of the studied matter and of the capacity to use the acquired skills in regular situations.

**Level III tests** include special assignments requiring queries from the student for which no ready algorithms are available, whereas the solutions lead to obtaining of subjectively new information. The evaluation of the testing is performed by comparison with a reference solution or by an expert judgement.

**Level IV tests** reveal students’ capacities to take decisions in new problematic situations, the solutions being a result of creative activities and followed by acquisition of objectively new information. The reference is a method of problem solving developed by experts.

Popular in the academic activity in institutions of higher education are Level I and Level II tests, whereas Level III and Level IV tests are used for olympiads, municipal, regional and international contests [4].

## RESULTS OF STUDIES

The required number of test assignments can be determined based on the preset evaluation reliability  $P$  and the ratio  $q = \varepsilon/s$ , thereby  $s$  is the empiric error standard;  $\varepsilon$  is the estimation of the measurement accuracy by means of confidence regions, determined in accordance with Student’s distribution [4, 25]. Thus, the minimum number of assignments shall be at least 30 to ensure testing reliability with  $P = 70\%$  trustworthiness at equal preset measurement accuracy. More assignments mean higher fidelity of estimations by means of confidence regions and higher stability of test results. In practice, the number of tests of Level I and Level II can be assigned based on the diagram in Fig. 1 which

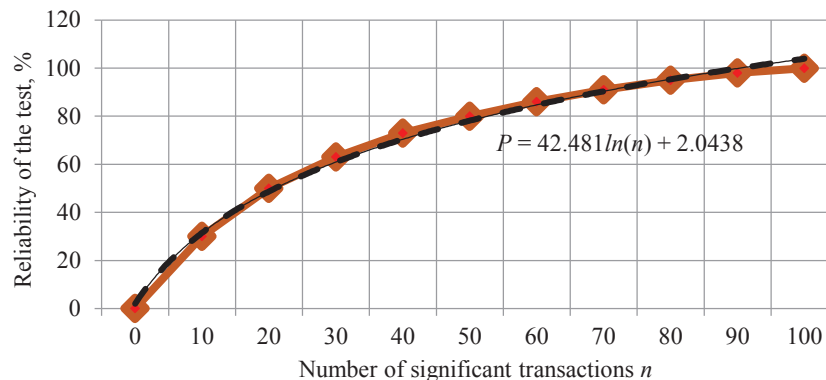


Fig. 1. Dependence of reliability of tests on the number of significant transactions



makes it simple to determine a required number of assignments  $n$  for students in their test.

Fig. 1 shows the dependence of the reliability of test results on the number of significant transactions (number of sorting steps) approximated by means of a logarithmic relationship like  $P = 42.481 \ln(n) + 2.0438$ , obtained by the method of least squares, in accordance with the paper [25].

The diagram demonstrates that at a number of assignments  $n = 35$  the reliability of the results is 71 %, whereas at the number of assignments  $n = 70$  the reliability makes up 92 %.

Computerized testing was used from 1995 to 2005 in the Regional Faculty of ICC (now Mytishy Branch of MSCU) concerning steel, reinforced concrete and wooden structures in A3000 hardware with *RICS OS* allowing for use of *BASIC BBC* programming language [19]. The use of the tests contributed to the increase the responsibility and the progress of the students the evidence whereof were the leading positions of the Branch in students' progress within the aforesaid period among other subdivisions of the University [4]. From 2006 to 2015, the testing was used by the Chair of Applied Mechanics and Mathematics in the subject "Strength of materials" also resulting in a positive effect [26].

For the time being, testing is used for monitoring of knowledge of the students of extramural courses of NRU MSCU.

## SUMMARY AND DISCUSSION

In order to build-up an efficient system of monitoring tests in a certain subject, the following is required:

- compilation of a database containing a number of assignments allowing for complete coverage of the subject matter avoiding repeated assignments for one testee;

- determination of the selection of the required number of assignments by means of the random number generator (to ensure the required test reliability, the number of the assignments shall not be less than 30, whereas the maximum number of assignments should not exceed 70 due to the time limit of the testing);

- time limit setup in accordance with the labor intensity of the assignments. The fulfillment time limit should be determined by expertise;

- definition of the evaluation of a single assignment as the ratio of the number of correctly performed significant transactions vs. the standard number of the significant transactions;

- calculation of the percentage of the right answers in the package;

- definition of the evaluation criteria (e.g., 70–79 % — satisfactory, 80–89 % — good, 90–100 % — excellent);

- if the time limit is exceeded, the test is deemed failed with "unsatisfactory" note;

- displaying the results in the monitor or better printing them out.

Thus, if there is a required number of computers of at least one PC per two students, correctly arranged computer testing considerably reduces the time demand of a monitoring event, increases the responsibility and the progress of the students, guarantees the objectiveness of the knowledge monitoring and helps to avoid conflicts.

Based on the testing results, a tutor can judge the strengths and the weaknesses of the academic activity and take prompt corrective measures.

Using and extending the testing to different subjects studied in NRU MSCU could contribute to deeper and more stable knowledge of the students.

## REFERENCES

1. Smirnova E.Je. Evaluation of the quality of education: approaches and practices. *Assessment of the quality of education in Russian universities: collected Works*. Saint Petersburg, 2014. (rus.).
2. Yanchenko S.I. Independent test control and its place in the system of monitoring the level of students' academic achievements. *New educational technologies in the university. collection of abstracts of the conference participants*. Ekaterinburg, Ural Federal University named after the first President of Russia B.N. Yeltsin Publ., 2014; 1564-1569. (rus.).
3. Shein A.A., Privalov N.I. Methodology for assessing the quality of students' knowledge in the test control of the assimilation of material. *New Volgograd State Technical University. Series: New educational systems and technologies of training in high school*. 2010; 7:8(68):215-217. (rus.).
4. Kuznecov V.S., Posvjatenko Ju.V. Objective control of the quality of students' knowledge in the context of a competence approach to education. *Humanitarian education in a construction institution. Competence approach and modern pedagogical technologies: collection of scientific and methodological materials of the Faculty of Humanitarian Education*. Moscow, MGSU Publ., 2007; 4:255-279. (rus.).
5. Osipov A.V. Effective assessment of the share of knowledge in the test control. *In the world of scientific discoveries*. 2010; 4-11(10):127-128. (rus.).
6. Bolysbayev D., Almetov N. The technology of applying the didactic potential of ethnoculture at school. *The Scientific Journal of the Modern Education & Research Institute*. 2017; 19-29.
7. Honour E. Workinger S. Principles of test & evaluation. *A Workshop in Two Days*. 2013; 2.1.

8. L'Horty Y., Petit P. Évaluation aléatoire et expérimentations sociales. *Revue Française d'Économie*. 2011; XXVI(1):13-48. DOI: 10.3917/rfe.111.0013
9. Nizonkiza D., Van De Poel K. Testing controlled productive knowledge of adverb-verb collocations in junior researchers using English as a foreign language. *Stellenbosch Papers in Linguistics*. 2016; 46(0):99-119. DOI: 10.5774/46-0-235
10. Kahlert J., Thomsen R., Pedersen L., Madsen M., Moller, N., Beck-Nielsen H., Sorensen H. Testing the robustness of background knowledge confounder selection in a case-control study. *Pharmacoepidemiology and drug safety*. 2014; 23:96-97.
11. Čaušević A., Sundmark D., Punnekkat S. Impact of test design technique knowledge on test driven development: a controlled experiment. *Lecture Notes in Business Information Processing*. 2012; 111:138-152. DOI: 10.1007/978-3-642-30350-0\_10
12. Bryukhanova T.N., Kiryushin A.V. About the test control of knowledge on the physics of students of the correspondence faculty of the TOGU. *Uchenye zapiski TOGU*. 2013; 4(4):1298-1300. (rus.).
13. Duban R.N. Construction of IRT profiles in the test knowledge control. *Actual problems of the humanities and natural sciences*. 2014; 3-1:52-57. (rus.).
14. Atoyev E.Kh. Some aspects of the use of computer technology in the test control of knowledge. *A young scientist*. 2016; 21(125):849-850. (rus.).
15. Shyyka S. Problem of optimization of computer test control of educational achievements of students in ukrainian language (for professional purposes). *Information Technologies and Learning Tools*. 2017; 60(4):288-299. DOI: 10.33407/itlt.v60i4.1783
16. Aleksenko O., Baranova I., Vashchenko S., Sobol A. Information system of students' knowledge test control. *Communications in Computer and Information Science*. 2015; 538:104-114. DOI: 10.1007/978-3-319-24770-0\_10
17. Khidirova Ch.M. Methods and algorithms of determination of complexity of test questions for formation a database system of the adaptive test-control of knowledge. *International conference on information science and communications technologies (ICISCT)*. 2017. DOI: 10.1109/icisct.2017.8188572
18. Zenkina S.V. Computer testing in assessing the level of students' achievements. *Standards and monitoring in education*. 2017; 1:30-32. (rus.).
19. Kuznecov V.S., Malahova A.N. Experience in the use of personal computers in the teaching of discipline reinforced concrete and stone structures. *Main results and prospects of higher construction education: collection of scientific and methodological developments*. Moscow, MGSU Publ., 2001; 270-280. (rus.).
20. Kireev B.N. The use of multimedia technologies in the study of technical disciplines. *Almanac of Modern Science and Education*. 2013; 10:69-73. (rus.).
21. Alexandrova L.V. Test control at the lessons of social studies as a means of learning and control (from work experience). *Simbirsk Scientific Bulletin*. 2015; 1(19):50-52. (rus.).
22. Bon A., Cros F. *Les évaluations en éducation au niveau international: impacts, contradictions, incertitudes*. CIEP. 2016; 172.
23. Wiggins G.P. *Assessing student performance: exploring the purpose and limits of testing*. 1st ed. San Francisco, Jossey-Bass Publishers, 1993; 336.
24. Aiken L.S., West S.G. *Multiple regression: testing and Interpreting interactions*. SAGE Publications, Inc. 1991. DOI: 10.1037/0021-9010.84.6.897
25. Rumshiskij L.Z. *Mathematical processing of experimental results*. Moscow, Nauka Publ., 1971; 192. (rus.).
26. Agapov V.P., Kuznecov V.S., Golovanov R.O. *Computer testing on the discipline of resistance materials*. Moscow, MGSU Publ., 2007; 15-20. (rus.).

Received November 8, 2018

Adopted in a modified form on November 27, 2018

Approved for publication March 27, 2019

**BIONOTES:** **Vitaly S. Kuznetsov** — Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Architectural and Construction Design, **Moscow State University of Civil Engineering (National Research University) (MGSU)**, 26 Yaroslavskoe shosse, Moscow, 129337, Russian Federation, umzmf@mgsu.ru;

**Galina E. Polekhina** — Candidate of Pedagogical Sciences, Associate Professor, Associate Professor of the Department of Applied Mechanics and Mathematics, **Moscow State University of Civil Engineering (National Research University) (MGSU)**, 26 Yaroslavskoe shosse, Moscow, 129337, Russian Federation, PolekhinaGE@mgsu.ru;

**Yulia A. Shaposhnikova** — Candidate of Technical Sciences, Associate Professor of the Department of Reinforced Concrete and Stone Structures, **Moscow State University of Civil Engineering (National Research University) (MGSU)**, 26 Yaroslavskoe shosse, Moscow, 129337, Russian Federation, gbk@mgsu.ru.