

Квалиметрия как теоретическая база оценки качества образования

В.В. Красильников, В.С. Тоискин, А.В. Шумакова
г. Ставрополь

Осмысление проблемы измерения и оценки качества в сочетании с проблемой управления качеством продукции, процессов привело к разработке нового научного направления – квалиметрии, изучающей и разрабатывающей принципы и методы количественной оценки качества. Временем появления квалиметрии можно считать 1968 г., когда была опубликована программная статья по формированию этого направления [1]. Термин «квалиметрия» наиболее полно отражает суть дела, (*квали* по латыни – «какой, какого качества», а *метрео* на древнегреческом означает «мерить, измерять»).

Первоначально термин «квалиметрия» был использован для обозначения научной дисциплины, изучающей методологию и проблематику количественного оценивания качества объектов любой природы, главным образом продукции. К 1970 г. уже был накоплен опыт для достаточно всестороннего исследования квалиметрии, ее сущности и взаимосвязей с различными научными областями. Одновременно происходило осознание того, что сфера приложения методов квалиметрии должна быть расширена от качества только продукции до качества объектов любой природы, включая и социально-экономические объекты, такие, например, как качество жизни, качество образования, качество человека. Уместно упомянуть, что в таком древнейшем памятнике культуры, как Библия, говорится: «Но ты все расположил мерою, числом и весом».

Сегодня квалиметрия представляет собой относительно новую, но вполне сформировавшуюся науку, знания которой необходимы всем специалистам, занимающимся оценкой и последующим управлением качеством различных объектов.

Однако для понимания сущности квалиметрии необходимо рассматривать ее в рамках науки квалитологии как науки о качестве. В структуре квалитологии можно выделить следующие взаимосвязанные и взаимодействующие друг с другом составные части: это теория качества, теория управления качеством, квалиметрия, метрология.

Отсюда квалиметрия — это синтетическая система взаимосвязанных теорий, различающихся степенью общности, средствами и методами измерения и оценивания, к которым относятся общая, специальная и предметная квалиметрия, составляющие в совокупности синтетическую квалиметрию [2, 3].

Общая квалиметрия изучает такие общетеоретические проблемы, как система понятий, теория оценивания (законы и методы), аксиоматика квалиметрии (аксиомы и правила), теория квалиметрического шкалирования (в том числе ранжирование, весомость).

В *специальной квалиметрии* рассматриваются такие модели и алгоритмы оценки, как экспертная, вероятностно-статистическая, индексная, таксономическая, теория классификаций и систематизаций сложно-ориентированных объектов, имеющих иерархическое строение, а также исследуются точность и достоверность полученных оценок.

Предметная квалиметрия различается по предметам оценивания, к которым относятся продукция и техника, труд и деятельность, решения и проекты, процессы спроса, предложения, управления и информационного обеспечения.

Структурность, динамичность, определенность и целостность этих видов квалиметрии обеспечиваются механизмом их взаимодействия. Так, общая квалиметрия трансформирует специальную квалиметрию с учетом особенностей применяемых методов и моделей оценки, а последняя находит отражение в предметной квалиметрии. При этом взаимосвязи общей, специальной и предметной квалиметрии отражают суть философского подхода к изучению любых иерархических систем как отношения общего, особого (особенного) и единичного.

У квалиметрии как науки можно выделить следующие аспекты: общенаучный, системный, технико-экономический, экономический, педагогический, социологический, управленческий, территориальный.

Общенаучный аспект определяется философско-методологической и общенаучной функциями категории качества и подтверждается формированием большого числа предметных свойств квалиметрии (продукция, техника, труд и т. д.). Применение этого аспекта создает действенный механизм выбора лучших вариантов многокритериальных решений во всех сферах и на всех уровнях управления качеством.

Системный (систематический) аспект квалиметрии определяет ее как системную теорию, включающую все основные признаки системы (структурность, динамичность, определенность, упорядоченность) и позволяющую использовать системный подход к процессам оценки, анализа и управления.

Технико-экономический аспект отражает направленность квалиметрии на комплексные оценки экономических и технических свойств объектов и процессов и представляется в результатно-затратных мерах эффективности, технико-экономических и иных показателях и обусловлен необходимостью количественной оценки качества и классификации продукции по качественным категориям, а не только по отдельным свойствам.

Экономический аспект обусловлен политэкономическим содержанием категории качества в ее взаимодействии с потребительной стоимостью и стоимостью и включает в себя методы эконометрии как теоретического измерения экономических свойств создаваемых объектов и процессов.

Педагогическая квалиметрия определяется как «дисциплина, изучающая методологию и проблематику количественной оценки качества педагогических инноваций и педагогической деятельности в целом» [4, с. 159].

С точки зрения методологии качества образования остановимся на некоторых вопросах общенаучного аспекта квалиметрии.

Как любое научное направление квалиметрия использует специфическую терминологию. Приведем базовые термины, способствующие пониманию сущности квалиметрии.

Квалиметрия — научная дисциплина, изучающая методологию и проблематику количественного оценивания качества (и отдельных составляющих его свойств) объектов любой природы.

Квалиметрическая информация — количественная информация о качестве объекта, позволяющая сделать заключение о том выше или ниже (а также насколько или во сколько раз выше или ниже) качество данного объекта по сравнению с другим объектом.

Количественное оценивание качества или интегрального качества - процесс, на выходе которого получается в комплексной, количественной форме квалиметрическая информация о качестве (или интегральном качестве) объекта с учетом не отдельных, а одновременно всех его свойств.

Объект — любой предмет (процесс, явление) материальный или идеальный (мысленный), одушевленный или неодушевленный, естественный или искусственный, продукт труда или продукт природы.

Свойство — черта, особенность, характеристика объекта, проявляющаяся при его производстве (создании) или потреблении (применении, использовании, эксплуатации).

Простое свойство - свойство, которое не может быть подразделено на совокупность двух или более других, менее сложных свойств.

Сложное свойство - свойство, которое может быть подразделено (разбито, декомпозировано) на два или больше других, менее сложных свойств. Отметим, что совокупность свойств может представлять собой (хотя и не всегда) тоже свойство объекта, но свойство более сложное.

Для того чтобы улучшить качество, нужно прежде всего уметь его количественно определять. Оценка качества — это первый и основной этап системы управления качеством. Поэтому необходимо разработать объективные методы оценки качества. Качество и количество — это результат некоторого процесса. Этот результат в сфере потребления формирует эффект, представляющий собой меру качества соответствующей продукции.

Мерой качества (или числовым представлением) R и его элементов $r(t)$ называется их отображение на множество действительных чисел R_1 :

$$\mu: R \rightarrow R_1 \text{ или } r(t) \rightarrow R_1.$$

Аналогично определяется мера (числовое представление) факторов:

$$Z: V \rightarrow R_1 \text{ или } U(t) \rightarrow R_1.$$

Таким образом, меры — это функционалы качества или функционалы факторов. Отображение μ , как и Z , может быть композицией нескольких отображений:

$$\mu = \mu_n^0 \dots \mu_2^0 \mu_1 \quad \mu = \mu_n \circ \mu_{n-1} \circ \dots \mu_2 \circ \mu_1$$

или

$$\mu(r(t)) = \mu_n(\dots\mu_2(\mu_1(r(t)))\dots).$$

В частном случае $\mu \equiv E$, т. е. $\mu(r(t)) \equiv r(t)$, где E в общем случае единичная матрица [2, с. 9].

Синонимом меры качества является показатель качества: $\mu \sim p$.

В отображении вместо множества R_e реальных чисел можно использовать множество смысловых единиц S_e типа «отличный», «хороший», «пригодный», «непригодный» и т. д. Такую меру называют семантической мерой качества. В символической записи она будет выглядеть как

$$S : R \rightarrow S_e \text{ или } S : \{r_i\} \rightarrow S_e$$

Понятие семантической меры включает в содержание квалиметрии не только количественную, но и семантическую (качественную) оценку.

К основным типам мер качества относят отображение шкалирования и отображение свертывания (комплексирования, агрегирования, обобщения).

Шкалирование - метод моделирования реальных процессов с помощью числовых систем. В социальных науках - антропологии, педагогике, социологии, психологии и прочих - шкалирование является одним из важнейших средств математического анализа изучаемого явления, а также способом организации эмпирических данных, получаемых посредством наблюдения, изучения документов, анкетного опроса, экспериментов или тестирования. Большинство социальных и психологических объектов невозможно строго фиксировать относительно места и времени их существования, отчего они не поддаются прямому измерению. Поэтому возникает вопрос о специфике числовой системы, могущей соотноситься с эмпирическими данными такого рода. Различные методы шкалирования как раз служат особыми приемами трансформации качественных характеристик в некую числовую переменную.

В самом общем смысле шкалирование есть способ познания мира через моделирование реальности с помощью формальных (в первую очередь, числовых) систем. Применяется этот способ практически во всех сферах научного познания и имеет широкое прикладное значение.

Шкалированием называется мера качества, вводящая упорядочивающие отношения на измеряемом множестве свойств $\{r_i\}$ или их мер. Или, другими словами, шкалирование - это процесс отображения по заданным правилам эмпирических множеств в формальные.

Под эмпирическим множеством понимается любая совокупность реальных объектов (людей, явлений, свойств, процессов, событий), находящихся в определенных отношениях друг с другом. Эти отношения могут быть представлены четырьмя типами (эмпирическими операциями): равенство (равно — не равно); ранговый порядок (больше — меньше); равенство интервалов; равенство отношений.

Понятие «меры качества» и «свертывания», «отображения шкалирования» («квалиметрического шкалирования») позволили расширить понятие «предмета квалиметрии как науки» и выдвинуть к концу 70-х годов

новую парадигму квалиметрии – синтетическую, охватывающую все виды измерений качества в разных предметных областях.

Под формальным множеством понимается произвольная совокупность символов (знаков, чисел), связанных между собой определенными отношениями, которые соответственно эмпирическим отношениям описываются четырьмя видами формальных (математических) операций: «равно — не равно» ($= \neq$); «больше — меньше» ($> <$); «сложение — вычитание» ($+ -$); «умножение — деление» ($\times \div$).

При шкалировании обязательным условием является взаимнооднозначное соответствие между элементами эмпирического и формального множеств. Это означает, что каждому элементу первого множества должен соответствовать только один элемент второго, и наоборот. При этом взаимнооднозначное соответствие типов отношений между элементами обоих множеств (изоморфизм структур) не обязательно. В случае изоморфности этих структур производится так называемое прямое (субъективное) шкалирование, при отсутствии изоморфизма производится косвенное (объективное) шкалирование.

Итогом шкалирования является построение шкал (лат. *scala* — «лестница»), т. е. некоторых знаковых (числовых) моделей исследуемой реальности, с помощью которых можно эту реальность измерить. Таким образом, шкалы являются измерительными инструментами. Общее представление обо всем многообразии шкал можно получить, например, из работы [5], где приведена их классификационная система и даны краткие описания каждого вида шкал.

Шкалирование может быть одномерным и многомерным. Одномерное шкалирование — это процесс отображения эмпирического множества в формальное по одному критерию. Получаемые одномерные шкалы отображают либо отношения между одномерными эмпирическими объектами (или одними и теми же свойствами многомерных объектов), либо изменения одного свойства многомерного объекта. Реализуется одномерное шкалирование с помощью методов и прямого (субъективного), и косвенного (объективного) шкалирования.

Прямое, или субъективное, шкалирование представляет собой процесс отображения эмпирического множества в формальное при взаимнооднозначном соответствии (изоморфизм) структур этих множеств.

Косвенное, или объективное, шкалирование — это процесс отображения эмпирического множества в формальное при взаимном несоответствии (отсутствии изоморфизма) между структурами этих множеств. К примеру, в психологии в основе такого несоответствия лежит первый постулат Фехнера о невозможности прямой субъективной оценки величины своих ощущений. Для количественного выражения ощущений используются внешние по отношению к ним (косвенные) единицы измерения, базирующиеся на различных оценках испытуемых: едва заметные различия, время реакции (ВР), дисперсия различения, разброс категориальных оценок.

Косвенные шкалы по способам их построения, исходным допущениям и единицам измерения образуют несколько групп, главные из которых следующие: 1) шкалы накопления или логарифмические шкалы; 2) шкалы, основанные на измерении ВР; 3) шкалы суждений (сравнительных и категориальных).

Под многомерным шкалированием понимается процесс отображения эмпирического множества в формальное одновременно по нескольким критериям. Многомерные шкалы отражают либо отношения между многомерными объектами, либо одновременные изменения нескольких признаков одного объекта. Процесс многомерного шкалирования в отличие от одномерного характеризуется большей трудоемкостью формализации данных. В связи с этим привлекается мощный статистико-математический аппарат, например, кластерный или факторный анализы, входящие неотъемлемой частью в методы многомерного шкалирования.

Шкалы измерений принято классифицировать по типам измеряемых данных, которые определяют допустимые для данной шкалы математические преобразования, а также типы отношений, отображаемых соответствующей шкалой. Современная классификация шкал была предложена в 1946 году Стэнли Смитом Стивенсом.

Шкала наименований (номинальная шкала). Это самая простая из всех шкал. В ней числа выполняют роль ярлыков и служат для обнаружения и различения изучаемых объектов. Числа, составляющие шкалу наименований, разрешается менять местами. В этой шкале нет отношений типа «больше - меньше», поэтому некоторые полагают, что применение шкалы наименований не стоит считать измерением. При использовании шкалы наименований могут проводиться только некоторые математические операции. Например, ее числа нельзя складывать и вычитать, но можно подсчитывать, сколько раз (как часто) встречается то или иное число.

Шкала порядка. Места, занимаемые величинами в шкале порядка, называются рангами, а сама шкала называется *ранговой*. В такой шкале составляющие ее числа упорядочены по рангам (т.е. занимаемым местам), но интервалы между ними точно измерить нельзя. В отличие от шкалы наименований шкала порядка позволяет не только установить факт равенства или неравенства измеряемых объектов, но и определить характер неравенства в виде суждений: «больше—меньше», «лучше—хуже» и т.п.

С помощью шкал порядка можно измерять качественные, не имеющие строгой количественной меры, показатели. Особенно широко эти шкалы используются в гуманитарных науках: педагогике, психологии, социологии. К рангам шкалы порядка можно применять большее число математических операций, чем к числам шкалы наименований.

Шкала интервалов. Это такая шкала, в которой числа не только упорядочены по рангам, но и разделены определенными интервалами. Особенность, отличающая ее от описываемой дальше шкалы отношений, состоит в том, что нулевая точка выбирается произвольно. Примерами могут быть календарное время (начало летоисчисления в разных календарях

устанавливалось по случайным причинам, температура, потенциальная энергия поднятого груза, потенциал электрического поля и др.).

Результаты измерений по шкале интервалов можно обрабатывать всеми математическими методами, кроме вычисления отношений. Данные шкалы интервалов дают ответ на вопрос «на сколько больше?», но не позволяют утверждать, что одно значение измеренной величины во столько-то раз больше или меньше другого.

Шкала отношений. Эта шкала отличается от шкалы интервалов только тем, что в ней строго определено положение нулевой точки. Благодаря этому шкала отношений не накладывает никаких ограничений на математический аппарат, используемый для обработки результатов наблюдений.

По шкале отношений измеряют и те величины, которые образуются как разности чисел, отсчитанных по шкале интервалов.

При использовании шкалы отношений (и только в этом случае!) измерение какой-либо величины сводится к экспериментальному определению отношения этой величины к другой подобной, принятой за единицу.

Шкала абсолютных величин (абсолютная шкала). Во многих случаях напрямую измеряется величина чего-либо. При таких измерениях на измерительной шкале отмечаются абсолютные количественные значения измеряемого. Такая шкала абсолютных значений обладает и теми же свойствами, что и шкала отношений, с той лишь разницей, что величины, обозначенные на этой шкале, имеют абсолютные, а не относительные значения. Результаты измерений по шкале абсолютных величин имеют наибольшую достоверность, информативность и чувствительность к неточностям измерений.

Шкалы интервалов, отношений и абсолютных величин называются метрическими, так как при их построении используются некоторые меры, т.е. размеры, принятые в качестве единиц измерений.

Рассмотренные типы шкал и допустимые преобразования (указаны в скобках) приведены на рис. 1. Прямоугольник, соответствующий более высокому типу шкал, расположен выше прямоугольника, соответствующего более низкому типу шкал.

В квалиметрии употребляются два термина: *измерение* и *оценка*. Под количественной оценкой в квалиметрии понимается некоторая функция отношения (выраженная чаще всего в %) показателя качества рассматриваемой продукции (процесса) к показателю качества продукции (процесса), принятой (принятого) за эталон.

Качество в квалиметрии рассматривается как некоторая иерархическая совокупность свойств, которые представляют интерес для некоторой рассматриваемой проблемы (задачи). Для удобства можно принять, что качество как некоторое наиболее обобщенное, комплексное свойство объекта рассматривается на самом низком, нулевом уровне иерархической совокупности свойств, а составляющие его менее обобщенные свойства на

более высоком уровне. Каждое из свойств первого уровня может состоять из некоторого числа менее общих свойств второго уровня и т. д. (рис. 2).

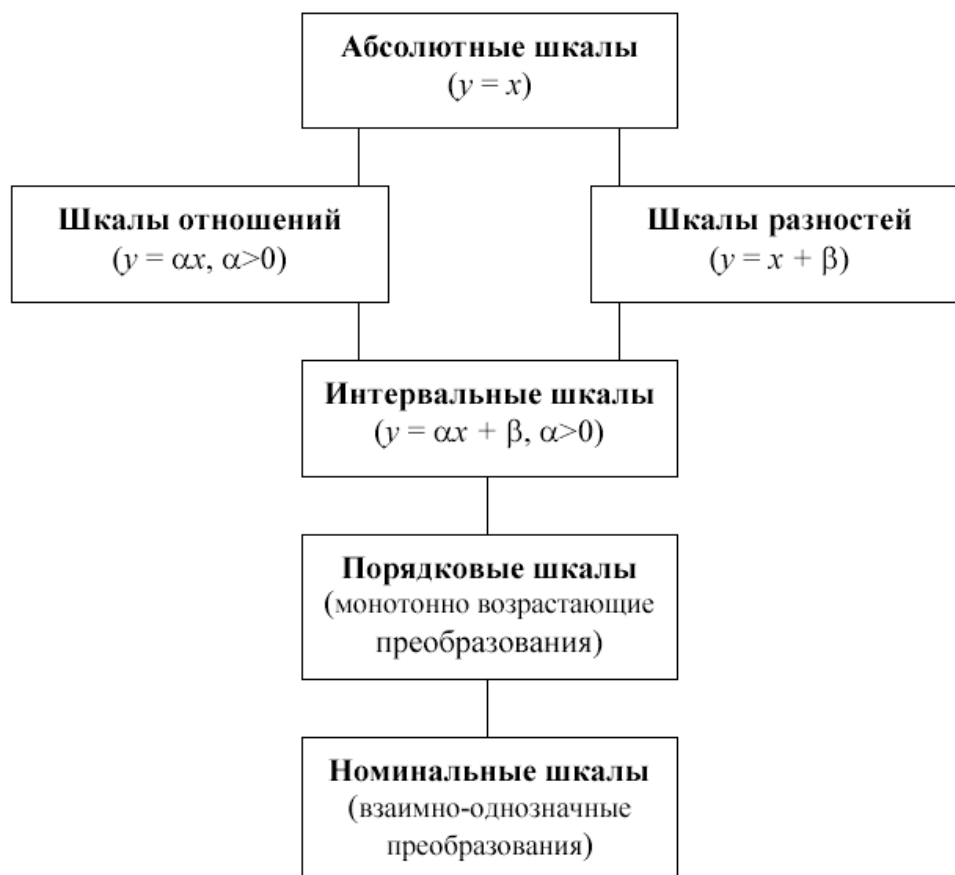


Рис. 1. Соотношение типов шкал (в скобках указаны преобразования)

Таким образом строится так называемое иерархическое дерево свойств, число уровней рассмотрения которого по мере познания может неограниченно возрастать. Желательно подняться до такого высокого уровня m рассмотрения, на котором находятся неразлагаемые, простые свойства объекта.

В большинстве случаев простые свойства могут быть измерены, а также могут быть получены их численные характеристики - абсолютный показатель свойств P_{ij} , где j – число свойств, лежащих на i -м уровне).

Абсолютные показатели ничего не говорят о свойстве с точки зрения «много-мало», «хорошо-плохо» и т. д. Чтобы определить уровень свойства необходимо определить относительный показатель

$$K_{ij} = f(P_{ij}, P_{ij}^{\text{баз}}),$$

где $P_{ij}^{\text{баз}}$ - абсолютный базовый показатель, а $f()$ – некоторая функциональная (с точки зрения математики) зависимость абсолютного и абсолютного базового показателей. Наиболее часто эта функциональная зависимость представляет собой отношение абсолютного и абсолютного базового показателей:

$$K_{ij} = f(P_{ij}, P_{ij}^{\text{баз}}),$$



Рис. 2. Дерево свойств качества

Если показатель характеризует одно свойство – это единичный показатель, все или несколько – это комплексный.

Оценка любого свойства K_{ij} зависит от выбранного базового показателя, вне выбранного эталона нельзя говорить об оценке качества.

Качество любого продукта (процесса) может оцениваться с точки зрения отдельного индивидуума (множество различных оценок качества), а также с точки зрения группы потребителей (общественной потребности). В квалиметрии оценка качества K определяется с точки зрения общественной потребности.

Показатель любого свойства, в том числе и показатель качества в целом, зависит от абсолютных показателей простых свойств K_{ij} , т. е. чтобы вычислить показатель качества, нужно свести воедино показатели простых свойств K_{ij} . Все они могут иметь разную размерность, поэтому необходимо с помощью специальных приемов на единой методологической основе перевести все простые свойства из шкал с разными размерностями в шкалу, имеющую единую размерность (или безразмерную). Эта операция носит название «трансформация шкал».

В квалиметрии считается, что любое свойство качества может быть определено двумя числовыми параметрами: весомостью (важностью q) и оценкой качества (K). Во многих методиках принимается, что весомость всех свойств, находящихся на одном и том же уровне удовлетворяет условию

нормирования $\forall i \sum_{j=1}^n q_{ij} = 1$, т.е. сумма важностей всех показателей на каждом уровне равна 1.

Другие методики предполагают разную суммарную весомость всех свойств на различных уровнях. Однако все методики подчиняются одному правилу: весомости всех свойств, находящихся на одном уровне, связаны друг с другом так, что сумма весомостей всегда остается постоянным, заранее заданным числом: $\sum_{j=1}^n q_{ij} = const$.

При вычислении комплексной оценки K_0 в подавляющем числе случаев применяют одну из средневзвешенных формул (арифметическую, геометрическую, гармоническую), а также применяют принципы теории машинного «распознавания образов».

Средневзвешенная арифметическая без учета коэффициента весомости:

$$K_0 = \frac{\sum_{i=1}^m K_i}{m};$$

с учетом коэффициента весомости

$$K_0 = \sum_{i=1}^m q_i K_i.$$

Среднегеометрическая:

$$K_0 = \sum_{i=1}^m q_i \sqrt[m]{\prod_{i=1}^m K_i^{q_i}}.$$

Обосновывая правильность применения средней геометрической, приводят следующие доводы. Средняя геометрическая обладает свойством обращать комплексную оценку качества в нуль, если оценка одного из показателей равна нулю.

Среднегармоническая:

$$K_0 = \frac{\sum_{i=1}^m q_i}{\sum_{i=1}^m \frac{q_i}{K_i}}.$$

Эта функция занимает некоторое промежуточное положение между средней арифметической и геометрической. Она, как и среднегеометрическая, учитывает разброс показателей вокруг среднего значения и, как среднегеометрическая в большей степени учитывает важность показателей.

Среднеарифметическая, как и другие средние, предполагает отсутствие корреляции между отдельными свойствами. Однако на самом деле многие свойства взаимосвязаны. Следовательно, модели, основанные на использовании средних, являются достаточно грубыми и требуют разработки более точных и достоверных моделей.

Следует отметить, что комплексная оценка качества посредством интегрального показателя не всегда уместна. Причины этого следующие.

1. Качество представляет собой совокупность различных полезных свойств, каждое из которых имеет свою единицу измерения. Это различие в единицах измерения и не позволяет объединить многочисленные показатели свойств объектов в единый комплексный показатель качества. Однако этот довод теряет силу, если сопоставление свойств ведется не непосредственно через их абсолютные показатели, а опосредованно, с помощью некоторых безразмерных функций этих показателей.

2. Единичные показатели качества, т. е. показатели только отдельных свойств, имеют определенную размерность, а, следовательно, и определенный смысл. Комплексная же оценка, выраженная в безразмерных единицах (баллах или процентах), конкретного физического смысла не имеет.

3. Любая методика комплексной оценки качества строится на предположении о неодинаковой весомости (важности) сопоставляемых между собой отдельных полезных свойств. Эти весомости, вводимые в расчет, находят так называемым экспертным методом как среднеарифметические значения, полученные в результате обработки мнений небольшого числа специалистов (носят субъективный характер).

4. Использование комплексных показателей качества создает возможность для «прикрытия» низкого уровня одних свойств более высоким других. Такое прикрытие одних свойств другими возможно далеко не всегда и не во всех пределах.

5. Применение комплексного показателя качества связано с учетом не всех, а только небольшого числа из общей совокупности свойств, характеризующих качества. Таким образом, в комплексной модели происходит упрощение, огрубление реально существующего качества.

6. Порой комплексная оценка качества не нужна, поскольку главная задача оценки – выявить, какие показатели свойств объекта (процесса) и насколько отстают от наилучших образцов данного типа и назначения.

Литература.

1. Азгальдов Г.Г., Гличев А.А. Квалиметрия – наука об измерении качества продукции // Стандарты и качество. № 1. 1968.

2. Субетто А.И. Квалиметрия человека и образования: генезис, становление, развитие, проблемы и перспективы // Материалы XI симпозиума «Квалиметрия в образовании: методология, методика и практика» – М.: Исследовательский центр проблем качества подготовки специалистов, 2006.

3. Бабкина Л.Н., Скотаренко О.В. Применение квалиметрического подхода в управлении региональной экономикой // Научно-технические ведомости СПбГПУ. Экономические науки № 4 (175). 2013. С. 45-52.

4. Педагогический словарь. – М.: Академия, 2006.

5. Гусев Е.К. Никандров В. В. Психофизика. Часть П. Психологическое шкалирование. - Л.: Наука, 1985.