
**Сенсорный анализ. Методология.
Руководящие указания по мониторингу
результативности работы группы по
количественному анализу**

*Sensory analysis — Methodology — Guidelines for monitoring the
performance of a quantitative sensory panel*

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 11132:2012

<https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012>

Ответственность за подготовку русской версии несёт GOST R
(Российская Федерация) в соответствии со статьёй 18.1 Устава ISO



Ссылочный номер
ISO 11132:2012(R)

Отказ от ответственности при работе в PDF

Настоящий файл PDF может содержать интегрированные шрифты. В соответствии с условиями лицензирования, принятыми фирмой Adobe, этот файл можно распечатать или вывести на экран, но его нельзя изменить, пока не будет получена лицензия на загрузку интегрированных шрифтов в компьютер, на котором ведется редактирование. В случае загрузки настоящего файла заинтересованные стороны принимают на себя ответственность за соблюдение лицензионных условий фирмы Adobe. Центральный секретариат ISO не несет никакой ответственности в этом отношении.

Adobe торговый знак фирмы Adobe Systems Incorporated.

Подробности, относящиеся к программным продуктам, использованным для создания настоящего файла PDF, можно найти в рубрике General info файла; параметры создания PDF были оптимизированы для печати. Были приняты во внимание все меры предосторожности с тем, чтобы обеспечить пригодность настоящего файла для использования комитетами-членами ISO. В редких случаях возникновения проблемы, связанной со сказанным выше, просьба проинформировать Центральный секретариат по адресу, приведенному ниже.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 11132:2012

<https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012>



ДОКУМЕНТ ЗАЩИЩЕН АВТОРСКИМ ПРАВОМ

© ISO 2012

Все права сохраняются. Если не указано иное, никакую часть настоящей публикации нельзя копировать или использовать в какой-либо форме или каким-либо электронным или механическим способом, включая фотокопии и микрофильмы, без предварительного письменного согласия ISO по соответствующему адресу, указанному ниже, или комитета-члена ISO в стране заявителя.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Опубликовано в Швейцарии

Содержание

Страница

Предисловие	iv
1 Область применения	1
2 Нормативные ссылки	1
3 Термины и определения	2
4 Принцип оценивания	3
5 Условия эксперимента	6
6 Квалификация испытателей	6
7 Процедура	6
7.1 Текущий контроль путём формальной валидации показателя результативности	6
7.2 Статистический анализ данных формальной валидации результативности (в одном сеансе)	7
7.3 Оценка общей результативности экспертной группы посредством формальной валидации показателя результативности	8
7.4 Оценка результативности испытателей методом формальной валидации	10
7.5 Обеспечение необходимой результативности	12
7.6 Текущий контроль в рамках регулярного профилирования продуктов	12
7.7 Планирование экспериментов для анализа зависимости показателей результативности от времени	12
7.8 Анализ временных рядов статистических данных	12
7.9 Воспроизводимость межгрупповых результатов	13
7.10 Статистический анализ сложных профилей	13
Приложение А (информативное) Пример практического применения	15
Приложение В (информативное) Пример использования критерия кумулятивных сумм	22
Приложение С (информативное) Пример использования карты Шухарта	25
Библиография	27

Предисловие

Международная организация по стандартизации (ISO) является всемирной федерацией национальных организаций по стандартизации (комитетов-членов ISO). Разработка международных стандартов обычно осуществляется техническими комитетами ISO. Каждый комитет-член, заинтересованный в деятельности, для которой был создан технический комитет, имеет право быть представленным в этом комитете. Международные правительственные и неправительственные организации, имеющие связь с ISO, также принимают участие в работе. ISO работает в тесном сотрудничестве с Международной электротехнической комиссией (IEC) по всем вопросам стандартизации в области электротехники.

Проекты международных стандартов разрабатываются согласно правилам, приведённым в Директивах ISO/IEC, часть 2.

Разработка международных стандартов является основной задачей технических комитетов. Проекты международных стандартов, принятые техническими комитетами, рассылаются комитетам-членам на голосование. Для публикации в качестве международного стандарта требуется одобрение не менее 75 % комитетов-членов, принявших участие в голосовании.

Принимается во внимание тот факт, что некоторые из элементов настоящего документа могут быть объектом патентных прав. ISO не принимает на себя обязательств по определению отдельных или всех таких патентных прав.

ISO 11132 был подготовлен Техническим комитетом ISO/TC 34, *Пищевые продукты*, Подкомитетом SC 12, *Органолептический анализ*.

[ISO 11132:2012](https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012)

<https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012>

Сенсорный анализ. Методология. Руководящие указания по мониторингу результативности работы группы по количественному анализу

1 Область применения

Настоящий международный стандарт содержит руководящие указания по контролю и оценке с помощью методов описательной статистики общего показателя эффективности экспертной группы количественного анализа и каждого члена этой группы в отдельности.

Такая группа экспертов может использоваться для количественного оценивания тех или иных качественных характеристик.

Показатель эффективности, о котором идёт речь, - это мера способности экспертной группы или отдельного её члена давать правильные оценки характеристик контролируемых изделий. Оценка характеристики может производиться в конкретный момент времени или отслеживаться в течение определённого периода. При этом показатель эффективности отражает способность группы экспертов обнаруживать, идентифицировать и измерять контролируемую характеристику, использовать оценочные характеристики аналогично другим экспертным группам или экспертам, различать входные сигналы, применять надлежащим образом оценочную шкалу, повторно получать собственные результаты и воспроизводить результаты других экспертных групп или отдельных экспертов.

Рассматриваемые методы обеспечивают согласованность, повторяемость и отсутствие смещения оценок, а также возможность выявления экспертных групп и отдельных экспертов, подлежащих контролю и оценке. Охватываются также задачи оценки согласованности решений между членами экспертной группы. Контроль и оценивание могут выполняться в одном сеансе или распределяться во времени.

Текущий контроль показателей результативности позволяет руководителю группы экспертов повышать эффективность работы каждого её члена и группы в целом, выявлять направления, по которым необходимо повышение квалификации специалистов, или определять, какие члены экспертной группы не выполняют свои функции должным образом и потому не могут продолжать оставаться участниками процедур экспертного оценивания.

Методы, определяемые данным международным стандартом, могут использоваться руководителем экспертной группы для организации непрерывного оценивания эффективности как экспертных групп, так и отдельных экспертов.

Настоящий международный стандарт применим как к специалистам или группам, проходящим обучение, так и к действующим экспертным группам.

2 Нормативные ссылки

Перечисленные ниже ссылочные документы обязательны для применения данного документа. В случае датированных ссылок действующим является только указанное издание. Применительно к недатированным ссылочным документам применяются их самые последние издания (включая все последующие изменения).

ISO 5492, *Органолептический анализ. Словарь*

ISO 8586, *Сенсорный анализ. Общее руководство по отбору, обучению и контролю за работой отобранных испытателей и экспертов*

ISO 8589, *Органолептический анализ. общее руководство по проектированию помещений для исследований.*

3 Термины и определения

В рамках настоящего документа используются термины и определения из стандарта ISO 5492 и термины, приведённые ниже.

3.1 согласованность agreement

способность различных экспертных групп или отдельных испытателей присваивать сравнимые оценки конкретному свойству выборочных экземпляров одного и того же изделия или продукта

3.2 единообразие, однородность homogeneity

мера согласованности балльных оценок, даваемых конкретными экспертами в сеансе тестирования: например, группой испытателей или отдельным испытателем в повторяющихся сеансах испытаний

3.3 смещение оценок испытателя assessor bias

характерная для испытателя тенденция формирования балльных оценок, постоянно оказывающихся выше или ниже истинной оценки, когда она известна, или средней групповой экспертной оценки в противном случае

3.4 постороннее значение, выброс outlier

оценка, которая не согласуется с общей совокупностью данных или резко отличается от других оценок того же самого продукта или аналогичных продуктов

3.5 дрейф экспертных оценок panel drift

явление, заключающееся в том, что показатели чувствительности группы испытателей с течением времени изменяются или начинают дрейфовать, приводя в конечном итоге к смещению диапазона оценок на шкале, по которой ранжируются значения некоторой характеристики постоянно контролируемого продукта

3.6 результативность performance

способность экспертной группы или отдельного испытателя давать правильные и надёжные оценки конкретных воспринимаемых раздражителей (стимулов) и характеристик каждого конкретного стимула

3.7 повторяемость repeatability

согласованность оценок эквивалентных выборок продуктов одной и той же экспертной группой или одним и тем же испытателем при одних и тех же условиях

3.8**воспроизводимость
reproducibility**

согласованность оценок эквивалентных выборок продуктов при разных условиях тестирования, в рамках разных задач или разными экспертными группами либо другим испытателем

ПРИМЕЧАНИЕ Воспроизводимость может оцениваться одним из следующих способов:

- путём проверки краткосрочной воспроизводимости результатов экспертной группы в двух или более сеансах, разделённых интервалом в несколько дней;
- путём проверки среднесрочной или долгосрочной воспроизводимости результатов экспертной группы в сеансах, разделённых интервалом в несколько месяцев;
- путём проверки воспроизводимости результатов разными экспертными группами в одной и той же лаборатории или в разных лабораторных условиях;
- путём проверки воспроизводимости оценок разных характеристик продукта одним испытателем.

3.9**валидация
validation**

процесс установления факта корреляционной связи конкретных сенсорных данных с другими данными по образцам того же самого продукта (например, с данными лабораторных измерений, с потребительской реакцией, с результатами других экспертных групп, с жалобами потребителей) или факта соответствия экспертной группы либо отдельного испытателя установленным критериям результативности

3.10**сеанс
session**

конкретный случай оценивания характеристик продукта

ПРИМЕЧАНИЕ В отдельном сеансе могут оцениваться один или несколько продуктов одним или несколькими испытателями. Применительно к каждому испытателю, работающему отдельно или в составе экспертной группы, сеансы разделяются во времени.

3.11**повторные сеансы
replicate sessions**

сеансы, в которых испытатели, продукты, условия тестирования и целевые задачи не меняются

4 Принцип оценивания

Настоящий международный стандарт касается экспертных групп сенсорного анализа, которые используются для проведения оценки значений одной или нескольких сенсорных характеристик продуктов с целью подготовки их количественных описаний, или профилей. Существует множество адекватных методов оценки и текущего контроля результативности экспертных групп, используемых для проведения такого тестирования.

Результативность экспертной группы сенсорного анализа может определяться с помощью уже существующих оценочных показателей или в рамках сессий экспертного оценивания, проводимых специально для получения данных, касающихся результативности.

Настоящий международный стандарт может использоваться для организации как периодического контроля, так и пересмотра текущих профильных данных.

В стандарте представлена специализированная процедура, подходящая для проведения периодической аттестации, сертификации и других аналогичных процессов. Блок-схема этой процедуры показана на Рисунке 1.

Для анализа текущих профильных данных, порождаемых экспертной группой, целесообразно использовать данные, полученные в ходе совершенно разных экспериментов по определению профилей с использованием разных типов продуктов, разного числа изделий и т.п. Процедура анализа отображена на Рисунке 1. Однако в связи с тем, что никакие возможные значения отклонений заранее не определены, рекомендуется, чтобы характеристики данного профиля, которые согласованно выделяются всей экспертной группой в целом, использовались как ключевые критерии для проверки результативности отдельных участников экспертной группы. Характеристики, по которым не наблюдается значимых отклонений, не могут служить надёжной основой для проверки согласованности оценок экспертов, поскольку наличие разногласий по ним в рамках экспертной группы и между её членами может означать, что эти продукты слишком похожи друг на друга, чтобы служить критериями для оценки результативности экспертов.

a) Текущий контроль посредством валидации результативности

Используется небольшое число образцов (возможно, три или четыре), о которых заведомо известно, что их характеристики различаются. Эти характеристики могут в дальнейшем служить ключевыми показателями для измерения результативности.



b) Общая результативность экспертной группы

- 1) Сколько ожидаемых ключевых характеристик существенно отличается друг от друга?
- 2) Сколько из этих ключевых характеристик имеет признаки зависимости от образца и испытателя? Ответ на этот вопрос первоначально указывает точку наименьшей согласованности оценок в рамках экспертной группы (7.3.2)
- 3) Какова повторяемость оценок ключевых характеристик в повторных сеансах (7.3.3)?

ISO 11132:2012
308079351781 6-11132-2012



c) Результативность отдельного испытателя

- 1) Различающая способность: сколько ожидаемых ключевых характеристик были существенно различными?
- 2) Повторяемость: есть ли согласие в различающей способности по данной характеристике и по продукту (7.4.2)?
- 3) Вклад во взаимодействие: для каких характеристик появляется взаимодействие?
 - i) Взаимодействие перекрёстных эффектов (7.4.4)
 - ii) Взаимодействие из-за различий в способе использования оценочной шкалы (7.4.5)



d) В тех случаях, когда выявляются проблемы с результативностью экспертной группы в целом или отдельных испытателей, требуется проведение соответствующих тренировочных сеансов.

Рисунок 1 — Блок-схема процедуры контроля результативности

В рамках отдельного сеанса подлежат определению следующие показатели:

- *смещение оценок, даваемых испытателем*, которое вычисляется как разность среднего значения оценки испытателя и известного “истинного значения” или усреднённой оценки экспертной группы как приближения для “истинного значения”;
- *повторяемость оценок испытателя*, представляемая обратной величиной стандартного отклонения (SD) повторных оценок испытателем того же самого образца или разных реплик одного и того же продукта;
- *воспроизводимость оценок испытателя*, представляемая стандартным отклонением (SD) смещения оценок испытателя на множестве отдельных продуктов;
- *различительная способность испытателя*, представляемая как его способность присваивать существенно отличающиеся балльные оценки разным продуктам.

Смещение оценок испытателя может указывать на то, что его сенсорная чувствительность отличается от сенсорной чувствительности других испытателей и/или его способ использования оценочной шкалы отличен от того, который свойствен другим испытателям.

Если возникает ситуация, когда какой-то испытатель даёт оценки, существенно отличающиеся от оценок других испытателей, то пересматриваются все результаты с целью определения:

- a) устойчивы или изменчивы оценки повторно отобранных образцов того же продукта;
- b) одинаковы или различны оценки разных образцов разных продуктов;
- c) сохраняется ли смещение оценок при использовании любых или только определённых оценочных шкал.

Для получения ответов на эти вопросы может использоваться метод дисперсионного анализа [Analysis of variance (ANOVA)].

В некоторых случаях смещение оценок может указывать на то, что соответствующий испытатель обладает более высокой сенсорной чувствительностью, что делает его результаты особенно ценными. Однако бывают и такие случаи, когда выявленное смещение оценок испытателя может потребовать его переобучения и удаления из состава экспертной группы.

Ниже описывается единый согласованный подход к использованию статистического анализа результатов экспертного оценивания. Однако ряд характеристик результативности экспертной группы может оцениваться с помощью нескольких мер описательной статистики. Например, как средний квадрат, так и стандартное отклонение ошибок SD (среднее значение квадратного корня) выражают изменчивость оценки продукта. Используемые на практике способы измерения должны быть привычными для конкретной области применения.

Другие релевантные методы измерения согласованности способов использования испытателями оценочной шкалы для той или иной характеристики продукта – это взаимодействие испытателя и продукта и коэффициент корреляции между оценками испытателя и оценками экспертной группы. Оценки испытателя могут быть несмещёнными, но сам он может использовать оценочную шкалу иным способом, чем его коллеги. Согласованность испытателя с остальными членами экспертной группы считается хорошей при коэффициенте корреляции близком к единице, наклоне линии регрессии близком к единице и близкой к нулю величине отрезка, отсекаемого линией регрессии на координатной оси.

При небольшом числе оценок (меньше шести) коэффициент корреляции должен интерпретироваться с особой осторожностью, поскольку в таких случаях он может быть высоким (достигать величины 0,7) лишь чисто случайно.

5 Условия эксперимента

Условия эксперимента должны соответствовать требованиям ISO 8589.

6 Квалификация испытателей

Группа должна иметь уровень квалификации и практического опыта, установленный применительно к экспертам в ISO 8586, или более высокий.

7 Процедура

7.1 Текущий контроль путём формальной валидации показателя результативности

В каждом сеансе группа экспертов снабжается множеством образцов, подобных тем, что подлежат экспертной оценке в процессе контроля продукции, для которых хотя бы применительно к одной паре может гарантироваться наличие статистически значимых отличий от других образцов не менее чем по восьми характеристикам.

Именно такое число образцов, рекомендуется для того, чтобы содействовать лидерам групп сенсорного анализа или организаторам сенсорного контроля в идентификации и отборе аттестационных образцов, которые позволят определить реальный и статистический показатели результативности экспертной группы.

Значения этих ключевых характеристик используются как эталоны, относительно которых будет оцениваться результативность группы экспертов. Набор образцов должен включать в себя реплики оцениваемых образцов в количестве, одинаковом для каждого из них. Число испытателей, оцениваемых образцов и реплик зависит от конкретных продуктов, контролируемых сенсорных характеристик и от целей процедуры экспертных оценок. Например, могут использоваться две или три реплики трёх или четырёх образцов. Следует ограничивать число оценок в сеансе во избежание притупления сенсорной чувствительности испытателей. Значения характеристик образцов должны лежать в диапазоне, который используется экспертной группой при оценке контролируемых продуктов.

Для проведения оценки строится рандомизированный блочный план, в котором испытатели выступают в качестве "блоков".

Если ожидается эффект переноса обобщений от текущего образца к следующему, то подходящим блочным планом будет латинский квадрат Вильямса. В этом случае в базовом плане используются четыре испытателя и четыре образца.

Таблица 1 — Латинский квадрат Вильямса

Испытатель	Порядок следования			
	1	2	3	4
1	A	B	C	D
2	B	D	A	C
3	C	A	D	B
4	D	C	B	A

В этом блочном плане каждый испытатель отбирает образцы четырёх продуктов в разной последовательности, и каждый отобранный продукт сопровождается другим продуктом для каждого испытателя: например, за А следует В для испытателя 1, С для испытателя 2, D для испытателя 3 и не следует ничего для испытателя 4.

При наличии множественных четвёрок испытателей для каждой четвёрки может повторяться тот же самый план.

7.2 Статистический анализ данных формальной валидации результативности (в одном сеансе)

Таблица 2 иллюстрирует один из способов табулирования и суммирования результатов. Некоторые компьютерные программы могут требовать иной организации данных: например с расположением образцов в столбцах, а испытателей – в строках.

Таблица 2 — Результаты испытателей

Образец	Испытатель								Среднее
	1		2		j		n_q		
	Оценки	Среднее	Оценки	Среднее	Оценки	Среднее	Оценки	Среднее	
1	Y_{111} Y_{112} Y_{11n_r}	$\bar{Y}_{11..}$			Y_{1j1} Y_{1j2} Y_{1jn_r}	$\bar{Y}_{1j.}$			$\bar{Y}_{1..}$
2									
i	Y_{i11} Y_{i12} Y_{in_r}	$\bar{Y}_{i1.}$			Y_{ij1} Y_{ij2} Y_{ijn_r}	$\bar{Y}_{ij.}$			$\bar{Y}_{i..}$
n_p									
Среднее					$\bar{Y}_{.j.}$				$\bar{Y}_{...}$

В этой таблице предполагается, что:

- $n_p \equiv$ число образцов ($i = 1, 2 \dots n_p$);
- $n_q \equiv$ число испытателей ($j = 1, 2 \dots n_q$);
- $n_r \equiv$ число реплик на образец ($k = 1, 2 \dots n_r$).

Для измерения результативности экспертной группы в целом и отдельных испытателей без признаков смещения оценок необходимо проведение дисперсионного анализа данных.

Подробности, касающиеся выполняемых для этого базовых вычислений, в настоящем международном стандарте не рассматриваются, поскольку операции дисперсионного анализа обычно реализуются пакетом прикладных программ.

Данные по каждому испытателю исследуются с применением одностороннего дисперсионного анализа (Таблица 3).

Таблица 3 — Дисперсионный анализ данных отдельного испытателя по одной характеристике

Источник дисперсии	Число степеней свободы	Сумма квадратов	Средний квадрат	Коэффициент F
Между образцами	$\nu_1 = n_p - 1$	S_1	$MS_1 = s_1/\nu_1$	$F = MS_1/MS_2$
Ошибка	$\nu_2 = n_p(n_r - 1)$	S_2	$MS_2 = s_2/\nu_2$	
Сумма	$\nu_3 = n_p n_r - 1$	S_3		
$n_p \equiv$ число образцов				
$n_r \equiv$ число реплик на образец				

Данные по всему сеансу исследуются с применением дисперсионного анализа рандомизированного блочного плана (Таблица 4).

Таблица 4 — Дисперсионный анализ данных всего сеанса по одной характеристике

Источник дисперсии	Число степеней свободы	Сумма квадратов	Средний квадрат	Коэффициент F
Между образцами	$\nu_4 = n_p - 1$	S_4	$MS_4 = s_4/\nu_4$	$F = MS_5/MS_7^a$
Между испытателями	$\nu_5 = n_q - 1$	S_5	$MS_5 = s_5/\nu_5$	
Взаимодействие	$\nu_6 = (n_p - 1)(n_q - 1)$	S_6	$MS_6 = s_6/\nu_6$	$F = MS_6/MS_7$
Ошибка	$\nu_7 = n_p n_q (n_r - 1)$	S_7	$MS_7 = s_7/\nu_7$	
Сумма	$\nu_8 = n_p n_q n_r - 1$	S_8		
$n_p \equiv$ число образцов				
$n_q \equiv$ число испытателей				
$n_r \equiv$ число реплик на образец				
<p>^a Если взаимодействие значимо, то коэффициент F для дисперсии между испытателями вычисляется по формуле $F = MS_5/MS_6$ со среднеквадратическим значением показателя взаимодействия в знаменателе дроби.</p>				

7.3 Оценка общей результативности экспертной группы посредством формальной валидации показателя результативности

7.3.1 Выделение ключевых характеристик

Подлежат определению ожидаемые ключевые характеристики, которые существенно отличаются друг от друга. Каждая такая характеристика идентифицируется по значимому изменению от образца к образцу при уровне значимости 0,05 в таблице дисперсионного анализа (Таблица 4). Чем больше доля различающихся ключевых характеристик, тем эффективней работа экспертной группы. Экспертная группа, не продемонстрировавшая ожидаемой различительной способности ключевых характеристик, должна пройти процедуру повышения квалификации.

7.3.2 Обеспечение однородности экспертной группы

Экспертная группа считается неоднородной, если какие-либо входящие в её состав испытатели дают оценки, не согласующиеся с оценками остальных коллег.