ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р ИСО 10075-3— 2009

ЭРГОНОМИЧЕСКИЕ ПРИНЦИПЫ ОБЕСПЕЧЕНИЯ АДЕКВАТНОСТИ УМСТВЕННОЙ НАГРУЗКИ

Часть 3

Принципы и требования к методам измерений и оценке умственной нагрузки

ISO 10075-3:2004
Ergonomic principles related to mental workload — Part 3: Principles and requirements concerning methods for measuring and assessing mental workload (IDT)

Издание официальное

БЗ 10—2009/708

Москва
Стандартинформ
2010

gost.e-ecolog.ru

Предисловие

Цели и принципы стандартизации в Российской Федерации установлены Федеральным законом от 27 декабря 2002 г. № 184-ФЗ «О техническом регулировании», а правила применения национальных стандартов Российской Федерации — ГОСТ Р 1.0—2004 «Стандартизация в Российской Федерации. Основные положения»

Сведения о стандарте

1 ПОДГОТОВЛЕН Автономной некоммерческой организацией «Научно-исследовательский центр контроля и диагностики технических систем» (АНО «НИЦ КД») на основе собственного аутентичного перевода на русский язык стандарта, указанного в пункте 4

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 201 «Эргономика»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 7 декабря 2009 г. № 585-ст

4 Настоящий стандарт идентичен международному стандарту ИСО 10075-3:2004 «Эргономические принципы, относящиеся к умственной нагрузке. Часть 3. Принципы и требования к методам измерений и оценке умственной нагрузки» (ISO 10075-3:2004 «Ergonomic principles related to mental workload — Part 3. Principles and requirements concerning methods for measuring and assessing mental workload», IDT).

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ Р 1.5—2004 (подраздел 3.5).

При применении настоящего стандарта рекомендуется использовать вместо ссылочных международных стандартов соответствующие им национальные стандарты, сведения о которых приведены в дополнительном приложении ДА

5 ВВЕДЕН ВПЕРВЫЕ

Информация об изменениях к настоящему стандарту публикуется в ежегодно издаваемом информационном указателе «Национальные стандарты», а текст изменений и поправок — в ежемесячно издаваемых информационных указателях «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ежемесячно издаваемом информационном указателе «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет

© Стандартинформ, 2010

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 2
4 Измерение и оценка умственной нагрузки 3
Приложение А (справочное) Дополнительная информация, относящаяся к обобщаемости способа измерений 11
Приложение В (справочное) Контрольная таблица выбора оборудования для измерений 12
Приложение ДА (справочное) Сведения о соответствии ссылочных международных стандартов ссылочным национальным стандартам Российской Федерации (и действующим в этом качестве межгосударственным стандартам) 13
Библиография 14

Введение

Настоящий стандарт устанавливает требования к разработке и выбору измерительного оборудования для оценки умственной нагрузки, рассмотренной в ИСО 10075[1] и ИСО 10075-2[2]. Для понимания положений настоящего стандарта необходима осведомленность об основных понятиях, установленных в этих двух документах.

Умственная нагрузка является частью общей рабочей нагрузки, поэтому пользователи настоящего стандарта должны быть знакомы с идеями, понятиями и положениями, установленными в ИСО 6385[3].

Целью настоящего стандарта является установление требований к разработке измерительного оборудования и оценке пригодности использования процедур измерения умственной нагрузки.

Настоящий стандарт устанавливает требования к проведению измерений показателей умственной нагрузки, но не устанавливает используемое оборудование и методы. Выбор используемого метода возможен при наличии дополнительной информации.

Международный стандарт, на основе которого подготовлен настоящий стандарт, разработан техническим комитетом ИСО/ТС 159 «Эргономика».

---

[1] ИСО 10075:1991 «Эргономические принципы, относящиеся к умственной нагрузке. Термины и определения».

[2] ИСО 10075-2:1996 «Эргономические принципы, относящиеся к умственной нагрузке. Часть 2: Принципы проектирования».

[3] ИСО 6385:2004 «Применение эргономических принципов при проектировании производственных систем».

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ЭРГОНОМИЧЕСКИЕ ПРИНЦИПЫ ОБЕСПЕЧЕНИЯ АДЕКВАТНОСТИ УМСТВЕННОЙ НАГРУЗКИ

Часть 3

Принципы и требования к методам измерений и оценке умственной нагрузки

Ergonomic principles of assuring the adequacy of mental workload. Part 3. Principles and requirements concerning methods for measuring and assessing mental workload

Дата введения — 2010—12—01

1 Область применения

Настоящий стандарт устанавливает основные принципы и требования к измерению и оценке умственной нагрузки и определяет основные требования к методам измерений. Настоящий стандарт содержит информацию, необходимую для выбора соответствующих методов измерений и описывает основные аспекты оценки и измерения умственной нагрузки.

Настоящий стандарт предназначен, главным образом, для экспертов по эргономике, например, психологов, специалистов в области здравоохранения и/или физиологов, обладающих соответствующей теоретической подготовкой и практическим опытом в использовании методов измерений и оценке умственной нагрузки, а также в интерпретации полученных результатов. В стандарте приведена полезная информация для разработки и/или анализа методов оценки умственной нагрузки.

Работодатели, работники, менеджеры и разработчики систем в настоящем стандарте могут найти полезную информацию в области измерений и оценки умственной нагрузки, например, относительно выбора используемых методов, основных критериев для оценки и выбора измерительного оборудования.

Настоящий стандарт содержит основную информацию, необходимую для выбора методов измерений умственной нагрузки в различных ситуациях. Существует большое количество различных методов, предназначенных для различных целей, ситуаций и уровней прецизионности. Выбранный метод должен обеспечивать необходимую результативность и эффективность. Информация, представленная в настоящем стандарте, позволяет пользователям оценить различные методы измерений и выбрать наиболее подходящие для достижения поставленных целей.

Требования к документации должны быть согласованы с положениями настоящего стандарта.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ИСО 6385:2004 Применение эргономических принципов при проектировании производственных систем (ISO 6385:2004, Ergonomic principles in the design of work systems)

ИСО 10075:1991 Эргономические принципы, связанные с умственной нагрузкой. Термины и определения (ISO 10075:1991, Ergonomic principles related to mental work-load; general terms and definitions)

ИСО 10075-2:1996 Эргономические принципы, относящиеся к умственной нагрузке Часть 2: Принципы проектирования (ISO 10075-2:1996, Ergonomic principles related to mental workload — Part 2: Design principles)

Издание официальное

3 Термины и определения

В настоящем стандарте применены термины по ИСО 6385, ИСО 10075, а также следующие термины с соответствующими определениями.

3.1 **объективность** (objectivity): Качество, проявляющееся в том, что полученные результаты измерений не зависят от лиц, управляющих оборудованием и выполняющих измерения, анализ и обработку данных.

3.2 **достоверность** (reliability): Уровень прецизионности результатов измерений, полученных с использованием соответствующего метода или оборудования.

Примечание — Достоверность обычно характеризуют однородностью (гомогенностью), устойчивостью и стабильностью результатов измерений, а в случае, когда обработку результатов измерений выполняют два или более специалистов, групп или организаций, также и взаимной достоверностью, характеризующей их оценки. Достоверность обеспечивает обобщаемость полученных результатов.

3.2.1 **однородность (гомогенность)** (homogeneity): Свойство, характеризующее близость результатов частей или элементов процедуры измерений при измерении одной и той же характеристики.

3.2.2 **устойчивость** (consistency): Свойство метода измерений, характеризующее близость результатов, полученных с применением различных частей или видов аналогичного оборудования.

3.2.3 **стабильность** (stability): Свойство процедуры измерений при повторении (обычно с задержкой по времени) давать идентичные результаты.

3.2.4 **взаимная достоверность** (inter-rater-reliability): Свойство, характеризующее способность двух или более специалистов, выполняющих измерения одним и тем же методом, получать одинаковые результаты для одинаковых характеристик.

3.3 **валидность** (validity): Способность метода или средства измерений выполнять требуемые измерения.

Примечание — Оценку валидности проводят путем сопоставления результатов измерений, выполненных конкретным методом, с результатами применения известной процедуры измерений, проверки выполнения критериев (когда устанавливают связь с известным критерием) или факторной проверки (когда демонстрируют, что процедура измерений в состоянии оценить установленные свойства объекта).

3.4 **чувствительность** (sensitivity): Свойство, характеризующее способность метода или средства измерений распознавать различные состояния объекта измерений, например различные состояния умственного напряжения или усталости.

3.5 **диагностичность** (diagnosticity): Способность метода или средства измерений распознавать различные виды или источники умственной нагрузки, например усталость, монотонность деятельности, умственное пресыщение или пониженную бдительность.

3.6 **обобщаемость** (generalizability): Возможность распространения наблюдаемого явления на установленные ситуации (условия умственного напряжения) и/или всю популяцию работников.

3.6.1 **относительная обобщаемость** (relative generalizability): Обобщаемость, относящаяся лишь к некоторым рабочим условиям и работникам, подвергающимся рабочей нагрузке (от низкой до высокой).

3.6.2 **абсолютная обобщаемость** (absolute generalizability): Обобщаемость, относящаяся к абсолютным уровням стресса/напряжения в установленных рабочих условиях для групп работников без учета стресса/напряжения, связанного с другими условиями/работниками.

Примечание — Дополнительная информация о обобщаемости приведена в приложении А.

3.7 **пригодность использования** (usability): Свойство продукции при ее предназначенном использовании определенным пользователем обеспечивать эффективность и результативность достижения установленных целей и удовлетворенность пользователя.

[ИСО 9241-11:1998, определение 3.1].

Примечание — В настоящем стандарте под продукцией следует понимать измерительное оборудование.

3.8 **критические значения** (critical values): Значения, используемые для оценки результатов измерений или принятия решений.

Примечание — Установление критических значений предполагает, что измерения с применением данного оборудования позволяют принимать решения с высокой точностью. Кроме того, должны быть установлены условия применения критических значений (например при измерениях после эксперимента и т.п.).

4 Измерение и оценка умственной нагрузки

4.1 Общие принципы

Умственная нагрузка является сложным и многомерным понятием, поэтому для оценки измерений умственной нагрузки не может быть применена единственная процедура. Не существует наилучшего способа измерений умственной нагрузки, так как наиболее подходящий способ оценки или измерений умственной нагрузки зависит от цели оценки, которая предполагает оценку различных аспектов умственной нагрузки, применения различных методов измерений и различной точности результатов измерений.

Модель, используемая в настоящем стандарте, позволяет учесть три компонента оценки умственной нагрузки:

- различные аспекты умственной нагрузки, такие как стресс, напряжение, усталость и т.п.;
- различные методы измерений, например анализ рабочего задания, оценка производительности, субъективные предпочтения или психологические измерения;
- различные уровни прецизионности, например при назначении контроля или оценке точности результатов измерений.

В соответствии со стандартами серии ИСО 10075 различные аспекты умственной нагрузки следует рассматривать по отдельности: стресс, умственное напряжение и его последствия. Таким образом, оценка и измерения должны учитывать этапы процесса стресс—напряжение—последствие, воздействующие на здоровье, безопасность, благополучие и производительность труда, т.е:

- оценку рабочих условий, приводящих к умственному напряжению при разработке и оценке проекта рабочей системы;
- оценку умственной нагрузки и напряжения, например, для определения максимальной нагрузки;
- оценку влияния напряжения на сотрудников, например влияния усталости, монотонии, умственного пресыщения или понижения внимания.

Различным целям необходимы разные уровни прецизионности, например, анализ риска опасных производственных процессов требует более высокой прецизионности, чем обработка анкет с общим обзором работы системы. Настоящий стандарт не устанавливает необходимый уровень прецизионности, т.к. требования к прецизионности зависят от цели и условий выполнения измерений, например от метрологических и законодательных требований и соотношений затрат и прибыли.

Метод измерений и оценки умственной нагрузки, независимо от используемых для этого технологий, классифицирован в настоящем стандарте по трем (низкий, средний и высокий) уровням прецизионности:

- уровень 1 применяют для целей точных измерений. Цель оценки умственной нагрузки на этом уровне заключается в получении достоверной, точной и правомерной информации о природе источника недостаточной или избыточной нагрузки для оптимизации рабочих условий. Методы, соответствующие этому уровню прецизионности, обычно предназначены для применения только специалистами, например психологами, экспертами в области эргономики, охраны труда и/или физиологии, имеющими соответствующую теоретическую подготовку и практический опыт применения (в плане использования таких методов и интерпретации полученных результатов);
- уровень 2 применяют для целей контроля, который требует менее точных измерений, чем уровень 1, например в случае, когда необходимо предотвратить последствия умственной нагрузки или идентифицировать причины возникновения неблагоприятной умственной нагрузки. Процедуры, соответствующие этому уровню, должны обладать менее высокой достоверностью, валидностью по сравнению с уровнем 1 и должны быть способны выявлять необходимость корректирующих измерений;
- уровень 3 применяют для целей получения грубой ориентировочной оценки. Методы, соответствующие этому уровню, позволяют получать информацию об умственной нагрузке при низком уровне точности. Они дают общую информацию о рабочих условиях, субъективных и психофизиологических состояниях сотрудника в отношении умственной нагрузки и, как правило, не являются ресурсоемкими. Информация на этом уровне должна допускать предотвращение негативных последствий умственной нагрузки путем принятия организационных решений, таких как изменение рабочего задания и/или методов и условий работы. Измерения на этом уровне обычно сводятся к приближенным оценкам, грубым измерениям (с умеренной достоверностью и валидностью), анализу рабочего задания, субъективным предпочтениям рабочих условий и субъективных состояний по отношению ко всем аспектам умственной нагрузки.

Для установления необходимого уровня прецизионности оборудования должен быть указан объект измерений. Необходимо установить, является ли объект измерений человеком или характеристикой ситуации. Если предполагаемый объект измерений является человеком, то должны быть оценены и установлены психометрические свойства для оценки людей. Если предполагаемый объект измерений является характеристикой ситуации, то психометрические свойства необходимо относить к оцениваемым условиям. Это более детально рассмотрено при анализе обобщаемости результатов испытаний (см. 4.2.7).

При установлении уровня прецизионности оборудования необходимо учитывать, на чем основаны результаты измерений: на единственной оценке/наблюдении или среднеарифметическом оценок/наблюдений.

Каждой процедуре оценки соответствует погрешность измерений, которая может быть уменьшена путем усреднения по некоторому количеству оценок/наблюдений. Оборудование, соответствующее требованиям уровня 3 для единичной оценки/наблюдения, может при достаточном количестве оценок достичь требований уровня 2 для среднеарифметического оценок/наблюдений. Таким образом, для более высокого уровня прецизионности должно быть определено необходимое количество оценок/наблюдений.

Оборудование для оценки умственной нагрузки должно продемонстрировать валидность в отношении заявленных аспектов умственной нагрузки и иметь четко установленную область применения (например умственная усталость или монотонность деятельности). Если область применения оборудования охватывает более одного аспекта умственной нагрузки, то документация по его валидности должна содержать подтверждение достоверности измерений для каждого аспекта.

Для оценки умственной нагрузки могут быть использованы различные методы, но для конкретных ситуаций некоторые из них более пригодны, чем другие. В частности, могут быть использованы следующие методы:

- физиологические измерения (методы предоставляют информацию о физиологическом состоянии сотрудника в заданных рабочих условиях);
- субъективное шкалирование (метод предоставляет информацию о субъективных оценках сотрудниками различных аспектов умственной нагрузки на рабочем месте, например, при использовании психометрических шкал и оценок сотрудниками своих рабочих условий);
- оценка производительности труда (методы дают возможность оценить умственную и психомоторную производительность труда сотрудника в определенных рабочих условиях, например, оценить уменьшение или изменение производительности труда под влиянием умственной нагрузки);
- анализ работы и производственного задания (методы оценивают элементы рабочего задания, физические и психологические условия работы, условия окружающей среды и организацию работы как источников умственной нагрузки).

Для достижения различных уровней достоверности необходимы различные методы, например, короткого опросного листа может быть достаточно для измерений на уровне 3, тогда как для верификации проекта системы на отсутствие однообразия в работе требуется использование шкал с обоснованной достоверностью. Для защиты особо ответственных или опасных систем от отрицательного воздействия снижения работоспособности вследствие неадекватной умственной нагрузки необходима самая высокая достоверность и валидность оборудования. Если такое оборудование недоступно, оно должно быть разработано. При отсутствии такого оборудования необходимо использовать методы с самыми высокими психометрическими критериями. Прецизионность зависит не только от методов измерений, а также от конструкции оборудования, психометрических свойств и правильного применения метода и оборудования.

При использовании методов, не удовлетворяющих задачам измерений, должны быть проведены специальные исследования в области умственной нагрузки и ее измерений:

- для оценки риска, связанного с использованием неоптимального оборудования;
- для обоснования полученных оценок.

На выбор подходящего оборудования влияют также законодательные и регулирующие требования, а также соображения экономической целесообразности.

4.2 Требования к процедуре

4.2.1 Общие положения

Оценка объективности, достоверности, валидности, чувствительности и диагностичности может быть выполнена любым подходящим и научно обоснованным способом. Тем не менее, должна быть разработана четкая стратегия оценки психометрических свойств способа измерений, которая должна

быть использована для правильного выбора процедуры и анализа полученных результатов. Фактически, валидность процедуры зависит от принятой модели измерений, например, вероятностная модель измерений требует применения вероятностной психометрии, в то время как подход обобщаемости потребует оценки компонентов дисперсии и вычисления коэффициентов обобщаемости (см. 4.2.7).

Не существует единственного наилучшего способа оценки психометрических свойств, так как он зависит, в частности, от принципов, лежащих в основе конструкции оборудования. Однако подход обобщаемости позволяет применять эффективные и результативные процедуры оценки параметров объективности, достоверности, валидности, чувствительности и диагностичности с затратой разумных усилий. В любом случае процедуры должны быть описаны способом, не требующим повторной оценки психометрических свойств.

4.2.2 Объективность

Для документирования объективности метода измерений должны быть представлены объективные свидетельства того, что сотрудник, выполняющий оценку или измерения, не влияет на результат оценки/наблюдений. Лучше всего это может быть достигнуто при участии более двух сотрудников, выполняющих измерения одним методом и обрабатывающих полученные данные с использованием дисперсионного анализа (ANOVA[4]) для определения их среднего или оценки взаимного влияния. Результаты должны показать, что такого влияния не существует, а соответствующие оценки компонентов дисперсии значимо не отличаются от 0. В исследованиях по проверке взаимного влияния необходимо обеспечить, чтобы были учтены воздействия сотрудника, выполняющего оценку/измерения. Поэтому вероятность ошибки β[5] должна быть равна 0,05, а вероятность ошибки α[6] должна быть не менее 0,20. Если невозможно устранить воздействия сотрудника, выполняющего оценку/измерение, метод измерений не может быть признан объективным.

4.2.3 Достоверность

Достоверность результатов измерений можно оценить с помощью корреляционных методов, если объектами измерений являются два сотрудника и/или две ситуации, относительное положение которых в совокупности должно быть достоверно воспроизводимо.

Коэффициенты корреляции не дают информации о различиях между наборами данных на абсолютном уровне. Такая информация может быть получена при использовании методов ANOVA, которые также позволяют вычислить коэффициенты устойчивости, например через коэффициенты корреляции внутри класса. Вычисление, например, корреляции между экспертами для контрольной таблицы может привести к высокой корреляции их относительных суждений и, следовательно, к положительной оценке достоверности, хотя оценки этих экспертов могут значительно различаться по абсолютной величине. Использование методов ANOVA позволяет получить такую корреляцию, но указывает на взаимосвязь оценок.

Однородность и/или устойчивость измерительного оборудования обычно оценивают на основе межэлементной корреляции или корреляции между частями в целом или корреляции между эквивалентными методами измерений. Строгие модели метода измерений требуют более строгих исследований, например с помощью анализа шкалограммы. Факторный анализ для одномерного измерения может быть использован для демонстрации однородности оборудования.

4.2.4 Валидность

Валидность процедуры измерений должна быть продемонстрирована путем совместной проверки с различными валидированными процедурами.

Если такие процедуры недоступны, валидность должна быть продемонстрирована экспериментально на основе изменения условий выполнения рабочих заданий, в процессе которых производят измерения, например, рабочего напряжения, приводящего к различным уровням усталости, если усталость является объектом измерений. В этом случае необходимо изменять интенсивность рабочего напряжения (например сложность выполнения рабочего задания) и длительность его воздействия. Изменения только интенсивности или только продолжительности воздействия недостаточно для демонстрации валидности. Должен быть использован подходящий набор рабочих заданий (в эксплуатационных или лабораторных условиях) вместе с выделением достаточного времени для выполнения задания, чтобы получить дополнительные воздействия или результат их взаимодействия.

---

[4] ANOVA — Analysis of variance.

[5] β — Ошибка второго рода.

[6] α — Ошибка первого рода.

Примечание — Если объектом измерений являются эмоциональное напряжение или его последствия, для демонстрации валидности процедуры измерения могут быть использованы группы с различными степенями эмоционального рабочего напряжения.

4.2.5 Чувствительность

Чувствительность оборудования должна быть продемонстрирована с помощью воспроизведения экспериментально управляемых вариаций умственной нагрузки, относящихся к интенсивности, продолжительности и их взаимодействию, например количества элементов, которые надо запомнить, или количества умственных операций, которые необходимо выполнить, или с помощью ограничения времени выполнения задания. В реальных условиях эти вариации могут быть получены с помощью отбора условий, в которых должна быть продемонстрирована чувствительность или получение оценки несколькими экспертами с высокой согласованностью оценок различных уровней умственного напряжения.

Градация, используемая при изучении чувствительности, должна быть репрезентативной для предполагаемой области применения измерительного оборудования.

4.2.6 Диагностичность

Для подтверждения диагностичности оборудования должна быть продемонстрирована его способность различать виды умственной нагрузки. Например, в области последствий умственного напряжения необходимо различение усталости, монотонии, умственного пресыщения и пониженной бдительности. То же относится и к требованиям или установленным условиям рабочего напряжения. Проверка диагностичности может быть выполнена экспериментально при помощи изменения условий работы, или в эксплуатационных условиях с помощью выбора известных условий, которые приводят к различным видам умственного напряжения.

Виды рабочей нагрузки, используемые при изучении диагностичности, должны быть указаны в области применения измерительного оборудования.

4.2.7 Обобщаемость

Для демонстрации обобщаемости необходимо использовать экспериментальные или квазиэкспериментальные методы в рабочих и/или лабораторных условиях с изменением различных условий, предоставляющих изменения области применения измерений (например результатов умственной нагрузки), а также условий, способствующих появлению существенной ошибки измерений, например времени суток, условий до работы или эксперимента. Оценка обобщаемости обычно включает оценку дисперсии компонентов в соответствии с дисперсионным анализом, в зависимости от структурной модели, включающей например, источники изменений. Таким образом, модель измерений должна заранее устанавливать, что следует рассматривать в качестве ошибки, а какие изменения соответствуют допустимой дисперсии. Коэффициенты обобщаемости вычисляют в виде отношения допустимой дисперсии к значимой ошибке измерений.

4.2.8 Пригодность использования

4.2.8.1 Общие положения

Методы измерений должны быть эффективными, результативными и удовлетворять потребностям пользователя, т.е. достигать ожидаемого уровня прецизионности в установленной области измерений с минимальными усилиями или объемом требований и приводить к желаемым результатам.

4.2.8.2 Эффективность

Эффективность измерительного оборудования определяет его психометрические свойства, особенно его достоверность, валидность, чувствительность и диагностичность.

4.2.8.3 Результативность

Результативность процедуры измерений зависит от эффективности и усилий, необходимых для выполнения измерений, анализа и интерпретации результатов. Таким образом, процедуры измерений с сопоставимой эффективностью следует сравнить по усилиям, необходимым для выполнения измерений. Поэтому количественные требования не могут быть установлены.

4.2.8.4 Удовлетворенность пользователя

Процедура измерений не только должна быть эффективной, но и должна удовлетворять требованиям пользователя. Жалобы на процедуру измерений являются показателем неудовлетворенности пользователя, однако отсутствие жалоб не может быть доказательством удовлетворенности пользователя. Поэтому важно оценить процедуру измерений по отношению к требованиям пользователя и документировать результаты оценки для ее улучшения. Однако общие количественные требования в этом случае не могут быть установлены, поскольку они зависят от ожиданий пользователя и особенностей области использования.

4.3 Количественные требования к измерительному оборудованию

4.3.1 Общие положения

Следует выбирать такое оборудование, которое лучше других удовлетворяет психометрическим критериям, установленным в настоящем стандарте.

4.3.2 Требования к объективности

Результаты измерений не должны быть подвержены существенному воздействию персонала, выполняющего оценку, анализирующего ее результаты и составляющего отчет. Это должно быть продемонстрировано путем валидации исследований, например, с помощью дисперсионного анализа, который показывает отсутствие существенных воздействий или взаимодействий.

4.3.3 Требования к достоверности

Коэффициенты однородности и устойчивости должны иметь значения не менее приведенных в графе «Достоверность» таблицы 1.

Примечание — Требование достоверности 0.7 (измеренной коэффициентом корреляции) соответствует 50 %-й ошибке и 50 % требуемой дисперсии и является минимальным требованием к достоверности процедуры измерений.

Для оценки взаимной достоверности экспертов необходимы те же самые значения, что и для оценки достоверности.

Стабильность зависит от интервала времени между измерениями происходящих событий/разработок. Требования к стабильности измерений в соседних интервалах времени должны быть сопоставимы с упомянутыми выше требованиями к однородности и устойчивости (за исключением субъективных шкал, где эти требования нужны только для параллельных форм).

4.3.4 Требования к валидности

При использовании метода совместного подтверждения (т.е. при использовании двух или более параллельных подходов измерений, например при использовании двух различных психометрических шкал), должны быть достигнуты значения коэффициентов валидности, приведенные в таблице 1. Эти значения относятся к измерениям одного этапа стресс—напряжение—последствие (см. ИСО 10075) с применением одного вида средств измерений.

Таблица 1 — Количественные требования к методам измерений для различных уровней прецизионности при использовании классической теории испытаний

Уровень прецизионности	Объективность	Достоверность	Валидность	Чувствительность	Диагностичность
1 Для точных измерений	нет влияния	≥ 0,9	≥ 0,5	≥ 5 уровней	< 0,10
2 Для целей контроля	нет влияния	≥ 0,8	≥ 0,4	≥ 3 уровней	< 0,20
3 Для грубой оценки	нет влияния	≥ 0,7	≥ 0,3	≥ 2 уровней	< 0,40

Вычисление коэффициентов валидности на всех этапах цикла стресс—напряжение—последствие является способом прогноза валидности и обычно заканчивается получением меньших значений коэффициентов валидности, чем при их вычислении по одному этапу. Однако при этом действуют одни и те же ограничения для валидности.

Если для средства измерений, выполняющего измерения нескольких факторов, заявлена валидность по каждому фактору, необходимо, чтобы коэффициент корреляции между факторами был менее 0,4.

Если валидность заявлена более чем для одного этапа процесса стресс—напряжение—последствие, то валидность должна быть продемонстрирована для каждого из этих этапов с указанием отдельных оценок валидности.

4.3.5 Требования к чувствительности

Методы измерений должны быть способны различать не только предельные состояния, но и идентифицировать как можно больше степеней рабочей нагрузки в предполагаемой области в процессе стресс—напряжение—последствие, например, различных степеней сложности задания.

Большее количество диагностируемых уровней нагрузки в предполагаемой области измерений обеспечивает большую чувствительность (т.е. большая чувствительность позволяет различать боль-

шее количество состояний). В таблице 1 показаны уровни нагрузки, которые должны быть различимы на каждом уровне прецизионности.

4.3.6 Требования к диагностичности

Методы измерений должны быть способны различать источники рабочего напряжения и разные виды последствий воздействия умственной нагрузки, например личных производственных заданий. Процедура измерений должна быть установлена для конкретной области измерений. Результаты измерений данной области не должны зависеть от результатов других измерений. Корреляция между результатами измерений (например усталости с умственным пресыщением) должна быть как можно меньше. Корреляция > 0,40 указывает на существенную зависимость и, следовательно, на недостаточную диагностичность. В таблице 1 указаны значения коэффициента корреляции, требуемые для каждого уровня точности.

Для демонстрации диагностичности оборудование должно быть применено в ситуациях с различными видами рабочей нагрузки (например различными условиями напряжения), и должно быть отражено, что оборудование показывает только те аспекты рабочей нагрузки, для которых предназначено его использование.

Хорошо известной процедурой демонстрации диагностичности является подход, когда несколько методов измерений применяют к нескольким областям измерений, а диагностичность оценивают по более высокой корреляции внутри одной области измерений, чем между различными областями измерений.

4.3.7 Требования к обобщаемости

Если использован подход теории обобщаемости (G-теории), то исследование обобщаемости (G-анализ) необходимо проводить в условиях как можно более характерных для предполагаемого использования, т.е. все условия, важные при выполнении измерений, должны быть учтены для обеспечения возможности оценки ошибок, измерений и, следовательно, обобщаемости исследуемого метода.

Высокая обобщаемость метода при применении универсальной шкалы, учитывающей различные условия измерений (время, количество повторений и др.), обеспечивает достоверность результатов G-анализа. Таким образом, G-коэффициенты для уровней прецизионности с 1-го до 3-го должны удовлетворять требованиям, приведенным в таблице 2.

Таблица 2 — Количественные требования к методам измерений для различных уровней прецизионности при использовании G-теории

Уровень прецизионности	G-коэффициенты				
	Объективность	Достоверность	Валидность	Чувствительность	Диагностичность
1 Для точных измерений	0	≥ 0,9	≥ 0,9	≥ 5 уровней	≥ 0,90
2 Для целей контроля	0	≥ 0,8	≥ 0,8	≥ 3 уровней	≥ 0,80
3 Для грубой оценки	0	≥ 0,7	≥ 0,7	≥ 2 уровней	≥ 0,70

Метод совместного подтверждения, использующий различные способы измерений одного и того же объекта, требует демонстрации высокой нагрузки всех средств измерений для исследуемого объекта (универсальной и/или составной шкалы) при многомерном анализе. Валидность конструкции может быть подтверждена высокими значениями G-коэффициентов при использовании чисто экспериментального подхода. При этом должны быть достигнуты значения G-коэффициентов, представленные в таблице 2

Для демонстрации чувствительности количество уровней, приведенное в таблице 2, должно быть дифференцировано. Это требует высокой обобщаемости на все исследуемые уровни нагрузки. Кроме того, должно быть продемонстрировано, что все уровни (не только предельные) могут быть уверенно распознаны с помощью определения 95 %-ных доверительных интервалов для каждого уровня.

Для демонстрации диагностичности исследуемое оборудование следует применять в нескольких ситуациях или экспериментальных условиях, различающихся по видам рабочей нагрузки. Если исследуемое оборудование используют в предполагаемой области применения, следует пользоваться значениями G-коэффициентов, приведенными в таблице 2. Если исследуемое оборудование используют в

других условиях, то G-коэффициенты должны быть установлены ниже (например около 0). При использовании конструкции с несколькими методами многомерного анализа должны быть достигнуты высокие нагрузки в предполагаемых областях применения метода измерений для всего оборудования, предназначенного для этих областей измерений. Низкую нагрузку используют для оборудования, предназначенного для измерений в других областях.

4.3.8 Обзор количественных требований

В таблице 1 представлены количественные требования в соответствии с классической теорией испытаний, а в таблице 2 представлена та же информация для G-теории.

Если оборудования, соответствующего этим критериям, не существует, его необходимо разработать. До его применения должны быть использованы методы с максимально возможными психометрическими критериями (см. 4.1). В любом случае оборудование с достоверностью < 0,5 не следует использовать вообще.

Примечание — Требования таблицы 1 или таблицы 2 следует применять при оценке риска.

4.4 Требования к документации

4.4.1 Требования к документации при разработке оборудования для измерений умственной нагрузки

Документация, относящаяся к разработке оборудования, предназначенного для измерений умственной нагрузки, должна содержать следующую информацию:

- год разработки;
- имя и инициалы разработчика;
- предназначенную область измерений (например умственный стресс, умственное напряжение, умственную усталость (см. ИСО 10075);
- уровень прецизионности оборудования и критерии достоверности, валидности, чувствительности и диагностичности;
- теоретическую основу метода измерений;
- группы населения/условия/области, для которых предназначены измерения;
- группы населения/условия/области, на которых были определены психометрические критерии;
- психометрические свойства оборудования, например достоверность, валидность, коэффициенты обобщаемости;
- конструкцию оборудования, используемого для определения психометрических свойств;
- задания/условия (и их вариации), использованные для определения валидности;
- продолжительность периодов работы и отдыха (и их последовательность), используемые для определения валидности;
- требования/ограничения на условия выполнения измерений;
- требования/ограничения для используемого оборудования;
- требования/ограничения, относящиеся к подготовке и квалификации персонала, выполняющего измерения (например психолога, специалиста по эргономике, физиолога, управляющего персонала, специалиста в области гигиены труда и т.д.);
- стандартные образцы, если имеются, для групп населения, заданий или условий работы;
- процедуры/условия, которые должны быть соблюдены при выполнении измерений;
- процедуры/условия, которые должны быть соблюдены при анализе данных;
- процедуры/условия, которые должны быть выполнены при интерпретации данных;
- оценку усилий (времени, сотрудников, оборудования), требуемых для выполнения измерений;
- норму ответа при анализе опросного листа;
- описание процедуры обработки данных, включающее результаты анализа оценки эффективности оборудования и удовлетворенности им пользователей;
- необходимую информацию для заполнения контрольной таблицы, представленной в приложении В.

4.4.2 Требования к отчету по измерению и оценке умственной нагрузки

Отчет о результатах измерений должен содержать следующую информацию:

- цель измерений;
- описание измерительного оборудования, включая краткое описание теоретической основы измерений, области измерений (например умственного стресса, умственного напряжения, умственной усталости), уровня прецизионности;

- имя и квалификацию сотрудника, ответственного за выполнение измерений;
- дату и время выполнения измерений;
- количество, возраст, пол и опыт участников или сотрудников, если они являются объектом измерений (описание испытательного образца);
- описание нормы ответа и представителей опрашиваемой группы (групп) при анализе опросного листа;
- принятые меры для сохранения объективности процедуры измерений;
- принятые меры для сохранения анонимности/конфиденциальности (где применимо);
- принятые меры для достижения информационной согласованности;
- данные контрольной таблицы в соответствии с приложением В;
- описание процедуры измерений;
- описание условий рабочей нагрузки (рабочее место, задание, и т.д.);
- ограничения, касающиеся выполнения измерений;
- комментарии участников или сотрудников;
- особенности, выявленные при выполнении измерений;
- описание результатов измерений;
- обсуждение и интерпретацию результатов измерений для будущих действий;
- подпись и дату.

Приложение А
(справочное)

Дополнительная информация, относящаяся к обобщаемости способа измерений

Обобщаемость связана с достоверностью измерений, которые могут быть распространены на группы населения и/или условия выполнения измерений. Должны быть учтены все условия, важные при выполнении измерений (например время дня, различные варианты конструкции дисплея и др.), а также условия, измерения в которых формируют репрезентативную выборку для проектирования рабочего места.

Обычно оценку в баллах используют для оборудования или метода при принятии решения о качестве конструкции рабочего места или о необходимости выполнения корректирующих измерений. Возможны два вида решений: относительные и абсолютные. Относительные решения предполагают сравнительный анализ, например сравнение вариантов конструкции. В этом случае решение основано на сравнении или ранжировании вариантов конструкции, позволяющих сделать вывод о том, насколько один вариант лучше другого. Если не существует альтернативной конструкции, то существует только один вариант, и решение может быть основано только на абсолютной оценке. Такие решения позволяют судить о том, насколько хорошо разработано рабочее место в отношении умственной нагрузки без учета других вариантов. Теория обобщаемости позволяет рассчитать два коэффициента, характеризующих достоверность каждого вида решений, а именно:

a) ρ_{δ}^{2} — относительный G-коэффициент;

b) Φ или ρ_{Δ}^{2} — абсолютный G-коэффициент.

G-коэффициент описывает точность, с которой можно распространить наблюдаемый результат на измерения из группы условий.

Теория обобщаемости превосходит классическую теорию испытаний по:

a) возможности работы с несколькими источниками погрешностей измерений;

b) возможности работы с составляющими погрешности измерения;

c) представлению коэффициентов достоверности измерений, пригодных для предназначенного использования измерений;

d) дифференциации между относительными и абсолютными решениями.

Кроме того, многие из вышеприведенных показателей, таких как устойчивость, стабильность или объективность, могут быть определены в пределах одного исследования, если использована адекватная конструкция, включая, например, элементы измерений и выносливость испытателей. Теория обобщаемости может быть применена также для исследования чувствительности и диагностичности измерительного оборудования при различных видах и уровнях рабочей нагрузки и экспериментальной обстановки.

Приложение В
(справочное)

Контрольная таблица выбора оборудования для измерений

При выборе оборудования рекомендуется использовать таблицу В.1.

Таблицу заполняют на основе имеющейся информации об оборудовании и требований к измерениям. Важные требования рекомендуется выделить и проверить их выполнение. Если не представлены количественные требования, следует проверить пригодность оборудования для предназначенной цели измерений.

Таблица В.1 — Контрольная таблица для выбора измерительного оборудования

Цель измерений	Пункт настоящего стандарта	Требования к прецизионности			Требования выполнены		Оборудование неприменимо	Комментарии
		Уровень прецизионности 1	Уровень прецизионности 2	Уровень прецизионности 3	Да	Нет		
1 Классическая теория исследований								
1.1 Объективность	4.3.2	Нет влияния	Нет влияния	Нет влияния				
1.2 Достоверность	4.3.3	≥ 0,9	≥ 0,8	≥ 0,7				
1.3 Валидность	4.3.4	≥ 0,5	≥ 0,4	≥ 0,3				
1.4 Чувствительность	4.3.5	≥ 5 уровней	≥ 3 уровней	≥ 2 уровней				
1.5 Диагностичность	4.3.6	< 0,10	< 0,20	< 0,40				
2 Обобщаемость	4.3.7							
2.1 Объективность		0	0	0				
2.2 Достоверность		≥ 0,9	≥ 0,8	≥ 0,7				
2.3 Валидность		≥ 0,9	≥ 0,8	≥ 0,7				
2.4 Чувствительность		≥ 5 уровней	≥ 3 уровней	≥ 2 уровней				
2.5 Диагностичность		≥ 0,90	≥ 0,80	≥ 0,70				
3 Пригодность использования	4.2.8							
3.1 Эффективность	4.2.8.2							
3.2 Результативность	4.2.8.3							
3.3 Удовлетворение	4.2.8.4							

Приложение ДА
(справочное)

Сведения о соответствии ссылочных международных стандартов ссылочным национальным стандартам Российской Федерации (и действующим в этом качестве межгосударственным стандартам)

Таблица ДА.1

Обозначение ссылочного международного стандарта	Степень соответствия	Обозначение и наименование соответствующего национального стандарта
ИСО 6385:2004	IDT	ГОСТ Р ИСО 6385—2007 Эргономика. Применение эргономических принципов при проектировании производственных систем
ИСО 10075:1991	—	*
ИСО 10075-2:1996	IDT	ГОСТ Р ИСО 10075-2—2009 Эргономические принципы обеспечения адекватности умственной нагрузки. Часть 2. Принципы проектирования

* Соответствующий национальный стандарт отсутствует. До его утверждения рекомендуется использовать перевод на русский язык данного международного стандарта. Перевод данного международного стандарта находится в Федеральном информационном фонде технических регламентов и стандартов.

Примечание — В настоящей таблице использованы следующие условные обозначения степени соответствия стандартов:

- IDT — идентичные стандарты.

Библиография

AGARD, Human performance assessment methods, AGARD-AGARDograph-308, Neuilly sur Seine: NATO-AGARD, 1989

BRENNAN, R. L., Generalizability theory, New York: Springer, 2001

CROCKER, L., & ALGINA, J., Introduction to classical and modern test theory. Chicago, Ill.: Holt, Rinehart and Winston, 1986

Shingeldecker, C.A., A task battery for applied human performance assessment research. AFAMRL Technical Report 84-071; Dayton: Air Force Aerospace Medical Research Laboratory, 1984

ISO 9241-11:1998, Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability (ИСО 9241-11:1998 Эргономические требования к офисным работам с использованием видеодисплейных терминалов (VDT). Часть 11. Руководство по применимости)

УДК 658:562.014:006.354 ОКС 03.100.30, 03.120.10 Т59

Ключевые слова: эргономика, эргономические принципы проектирования, умственная нагрузка, объективность, валидность, стабильность, диагностичность, распространимость

Редактор *И.В. Меньших*
Технический редактор *Н.С. Гришанова*
Корректор *Р.А. Ментова*
Компьютерная верстка *В.И. Грищенко*

Сдано в набор 16.04.2010. Подписано в печать 21.05.2010. Формат 60x84 1/8. Бумага офсетная. Гарнитура Ариал.
Печать офсетная. Усл. печ. л. 2,32. Уч.-изд. л. 1,60. Тираж 114 экз. Зак. 418.

ФГУП «СТАНДАРТИНФОРМ», 123995 Москва, Гранатный пер., 4.
gost.e-ecolog.ru gost.e-ecolog.ru
Набрано во ФГУП «СТАНДАРТИНФОРМ» на ПЭВМ
Отпечатано в филиале ФГУП «СТАНДАРТИНФОРМ» — тип. «Московский печатник», 105062 Москва, Лялин пер., 6

gost.e-ecolog.ru