**Рабочая программа дисциплины**

«Измерения в психологии и образовании»

для направления подготовки научно-педагогических кадров в аспирантуре

44.06.01 Образование и педагогические науки,

профиль Общая педагогика, история педагогики и образования

Автор программы:

Тюменева Юлия Алексеевна, к. психол.наук, jutu@yandex.ru

Согласована на заседании Академического совета аспирантской школы по образованию

«\_18\_»\_октября\_ 2016 г., Протокол № 16

Общая трудоемкость 3 з.е.

Часов по учебному плану 114

В том числе:

Аудиторные занятия 36

из них:

лекции 14

семинары 22

Самостоятельная работа 66

Москва, 2016

*Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения кафедры-разработчика программы.*

# Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям аспиранта и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и аспирантов направления 44.06.01 Образование и педагогические науки, обучающихся в Аспирантской школе по образованию и изучающих дисциплину «Измерения в психологии и образовании».

Программа разработана в соответствии с:

* Образовательным стандартом НИУ ВШЭ по направлению 44.06.01 Образование и педагогические науки;
* Учебным планом подготовки аспирантов по направлению 44.06.01 Образование и педагогические науки, профиль Общая педагогика, история педагогики и образования.

# Цели освоения дисциплины

Целями освоения дисциплины «Измерения в психологии и образовании» являются:

* обеспечить аспирантов прочным пониманием принципов измерения в социальных науках: роли выбора конструкта и операционального определения для измерения, связи шкалы измерения с возможностями обработки данных, основными положениями классической теории тестирования; необходимых требований к измерению в той или иной тестовой ситуации и в связи с предназначением инструмента измерения;
* выработать представление об этапах и необходимых процедурах валидизации инструмента оценки;
* сформировать умение видеть разработку или использование инструмента оценки как необходимо звена внутри всего исследовательского процесса, подчиняющегося замыслу и логике последнего.

# Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины аспирант должен:

*Знать:*

* методологические основания различных подходов к измерению в социальных науках;
* основные психометрические процедуры, необходимые для проведения и/или оценки проведенного измерения, конструирования инструментария измерения, а также для оценки научного статуса используемых методик и процедур;
* основные типы измерительных инструментов и процедур, цели и особенности их применения.
* роль и место измерения и анализа данных в едином процессе взаимодействия с людьми, принимающими решения в той или иной области (образовании, социальной политике, управлении персоналом и пр.)
* этические нормы и правила, регулирующие исследовательскую и оценочную деятельность.

*Уметь:*

* различать операциональное определение от теоретического конструкта и предлагать различные варианты операционализации;
* выделять психометрические требования для измерительных процедур в соответствии с задачами их применения
* видеть исследовательские шаги, необходимые для обоснования надежности и валидности инструмента измерения
* делать психометрический анализ инструмента в рамках классической теории тестирования, используя пакеты программ SPSS, Iteman
* делать интерпретации психометрического анализа и определять условия пригодности использования инструмента
* писать отчет и докладывать результаты исследований надежности и валидности измерительных инструментов

*Иметь навыки (приобрести опыт):*

* делать психометрический анализ инструмента в рамках классической теории тестирования, используя пакеты программ SPSS, Iteman
* делать интерпретации психометрического анализа и определять условия пригодности использования инструмента
* писать отчет и докладывать результаты исследований надежности и валидности измерительных инструментов

В результате освоения дисциплины аспирант осваивает следующие компетенции:

| Компетенция | Код по ОС | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения, способствующие формированию и развитию компетенции |
| --- | --- | --- | --- |
| Способность к критическому анализу и оценке современных научных достижений, в том числе в междисциплинарных областях | УК-1 | Предлагает модель и способ оценки того или иного конструкта, обосновывает выбор. | Лекции, семинары, самостоятельная работы, письменные самостоятельные работы, практические занятия, работа по группам, общегрупповая дискуссия, презентации и их защите, мастер-классы |
| Способность осуществлять комплексные исследования, в том числе междисциплинарные, на основе целостного системного научного мировоззрения | УК-5 | Использует адекватные методы и владеет правилами проверки надежности и валидности инструментария, видит границы их применения демонстрирует. Адекватно применяет статистические процедуры, обосновывает их выбор, интерпретирует и оценивает результаты в более широком контексте. | Лекции, семинары, самостоятельная работы, письменные самостоятельные работы, практические занятия, работа по группам, общегрупповая дискуссия, презентации и их защите, мастер-классы |
| Готовность участвовать в работе российских и международных исследовательских коллективов по решению научных и научно-образовательных задач | УК-6 | Воспроизводит основные особенности многосторонней коммуникации, способен ее организовать.  Оформляет и представляет результаты своей деятельности в виде отчета по научной и практической работе и презентации на русском (государственном) и иностранном языке в публичном и индивидуальном общении. | Лекции, семинары, презентации и их защите, мастер-классы с участием зарубежных экспертов |
| Способность на основе анализа результатов теоретических и прикладных исследований формулировать релевантные основания для решений в области образовательной практики и политики | ОПК-3 | Использует концепции, модели, методы, способы и инструменты работы для решения комплексных задач  Демонстрирует знания о способах принятия управленческих решений и о последствиях неверного планирования | Лекции, семинары, самостоятельная работы, письменные самостоятельные работы, презентации и их защите, мастер-классы |
| Способность к выполнению междисциплинарных исследований в области социально-психологических, социокультурных, социально-экономических аспектов образования | ПК-2 | Использует существующие методы исследования и измерения, руководствуясь общими принципами.  Проявляет поисковое поведение и владеет правилами профессиональной деятельности в поисковой среде. | Лекции, семинары, самостоятельная работы, письменные самостоятельные работы, практические занятия, работа по группам, общегрупповая дискуссия, презентации и их защите, мастер-классы |

# Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к циклу дисциплин по выбору и изучается на 1-м или 2-м году обучения.

# Формы контроля знаний аспирантов:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Тип контроля | Форма контроля | 1 год | | | | Параметры \*\* |
| 1 | 2 | 3 | 4 |
| Текущий  (неделя) | Контрольная работа | \* |  |  |  | письменная работа на занятии |
|  | \* |  |  | письменная работа на занятии |
| Домашнее задание |  | \* |  |  | письменная работа 3-4 стр. |
| Итоговый | Зачет |  |  |  | \* | Тест |

## Критерии оценки знаний, навыков

Контрольная работа 1

|  |  |
| --- | --- |
| **Критерий** | **Количество баллов** |
| Умеет различать операциональные определения и конструкты | 0-2 балла |
| Правильно использует шкалы измерения, обосновывает ответ | 0-2 балла |
| Владеет базовыми статистическими процедурами | 0-3 балла |
| Правильно интерпретирует результаты статистического анализа | 0-3 балла |

Контрольная работа 2

|  |  |
| --- | --- |
| **Критерий** | **Количество баллов** |
| Демонстрирует умение использовать профессиональные компьютерные программы для подсчета надежности и анализа тестовых заданий | 0-5 баллов |
| Интерпретирует, делает выводы о пригодности теста и задний, обосновывает вывод | 0-5 баллов |

Домашняя работа

|  |  |
| --- | --- |
| **Критерий** | **Количество баллов** |
| Воспроизводит основные положения теории валидности | 0-5 баллов |
| Интерпретирует результаты исследований валидности | 0-5 баллов |

Оценки по всем формам текущего контроля выставляются по 10-ти балльной шкале.

Дистанционно осуществляется проверка домашних работ.

# Содержание дисциплины

**Тема 1 Проблема измерения в социальных науках**.

Специфика исследований в психологии и социальных науках в отличие от исследований в точных (естественных) науках. Теоретические конструкты и операциональные определения. Формулировка эмпирических вопросов: необходимость операционализации теоретических положений. Критерии хорошей операционализации. Последствия принятия конкретного операционального определения для последующего измерения и обработки результатов.

Концептуализация – необходимый шаг для обобщения и экспликации выводов исследования. Излишне широкая и излишне ограниченные концептуализации. Ограничения и некорректность концептуализации.

Роль теории в эмпирическом исследовании. Формальные и неформальные теории. Последствия теоретического анализа для последующего эмпирического исследования.

**Тема 2. Базовые компоненты измерения – шкалы и язык статистического анализа.**

Непрерывные и дискретные переменные. Понятие о шкале измерения. Виды шкал: номинальная, порядковая, интервальная, шкала отношений. Графики распределения. Допустимые статистики в каждой шкале: описание среднегрупповых данных, возможность сравнения индивида с группой, измерение взаимосвязи в каждой шкале. Сравнительная характеристика шкал измерения.

Язык описательной статистики. Описательная статистика: меры центральной тенденции (мода, медиана, среднее арифметическое) и меры изменчивости (разброс, отклонение, среднее отклонение, стандартное отклонение). Общее представление о корреляции и корреляционных процедурах. Вопросы, на которые может ответить корреляционное исследование.

Оценка корреляции: графическая и математическая, коэффициент корреляции. Специфика оценки корреляционной связи в различных шкалах измерения.

Толкование корреляции. Корреляционная связь и связь причинно-следственная. Коэффициент детерминации. Регрессионный анализ. Линия регрессии. Допущения, связанные с применением и толкование регрессии.

**Тема 3. Статистики для измерения в социальных науках.**

Частотное распределение. Нормальная кривая. Скошенные распределения. Нормальная кривая и вероятность. Выборка, ошибки выборки. Стандартная ошибка измерения. Доверительные интервалы. Z-показатели, линейные преобразования. Процентильные нормы. Нормализация и нелинейные преобразования.

**Тема 4. Классическая теория тестирования**

Основные положения классическая теория тестирования: истинный балл, тестовый балл, ошибка измерения. Определения интервалов для истинного балла. Конструирование и анализ тестовых заданий. Обзор тестовых заданий.

Систематически и несистематические ошибки и их источники.

**Тема 5. Надежность и валидность измерения.**

Понятие о надежности – как одной из основных психометрических характеристик измерения. Интерпретация надежности измерения как надежности тестовых показателей. Модель истинных баллов. Индекс надежности и коэффициент надежности. Ошибка измерения – показатель точности измерения. Альтернативные определения истинного балла и ошибки измерения.

Надежность как согласованность показателей. Виды надежности. Ретестовая надежность. Надежность расщепленных половин. Надежность, установленная методом эквивалентных форм. Внутренняя согласованность тестовых показателей.

Надежность оценок наблюдателей или экспертов – как особый вид надежности, применяемый в ситуации слабоструктурированных инструментов измерения (проективные тесты, интервью, наблюдения, экспертная оценка и пр.).

Коэффициенты надежности для критериально-ориентированных тестов.

Понятие о валидности – как одной из основных психометрических характеристик измерения. Интерпретация валидности как целостной характеристики измерения.

Аспекты валидности: очевидная, критериальная, конструктная, содержательная. Методы оценки различных аспектов валидности.

Статистические процедуры для предсказания и классификации: частная корреляция, множественная регрессия, дискриминантный анализ.

Bias в отборе. Базовые термины. Группы меньшинств. Психометрические определения bias. Процедуры справедливого отбора. Критика моделей.

Факторный анализ. Факторы и факторные нагрузки. Ротация. Коррелирующие факторы. Число факторов. Уникальные вклады. Эксплоративный и конфирматорный факторный анализ.

Связь валидности и надежности измерения.

**Тема 6. Анализ заданий и теория «задание-ответ**

Обзор тестовых заданий. Конструирование и анализ тестовых заданий. Трудность, дискриминативность заданий. Комплексная оценка заданий. Характеристическая кривая задания. Знакомство с программой Iteman. Анализ заданий с помощью Iteman и интерпретация.

**Тема 7. Подсчет баллов и интерпретация тестовых показателей. Композитные баллы и интерпретация.**

Процедуры подсчета баллов при разных форматах теста. Интерпретация тестовых показателей. Сырые, процентильные, стандартные показатели. Нормальное распределение и нормирование. Нормы. Выравнивание баллов для различных тестов. Определение выравнивания.

Сбор данных для выравнивания. Выравнивание в IRT подходе. Вертикальное выравнивание.

Что такое композитные баллы? Смысл их использования и примеры из международных сравнительных исследований. Специфика подсчета композитного балла. Взвешивание компонентов. Надежность.

**Тема 8. Условия проведения тестирования. Администрирование и этика тестирования.**

Источники систематических ошибок измерения. Условия проведение тестирования и диапазоны их изменчивости. Валидность тестовых баллов и условия тестирования.

Разработка Руководства к инструменту оценки. Процедуры для администрации теста. Типы администрирования. Влияние условий, в которых проводится тестирование. Стандартизация. Специальные мероприятия по обеспечению тестирования в особых группах респондентов. Обеспечение мотивации тестируемых. Тестирование как социальные отношения. Компьютер как тестирующий. Мошенничества при тестировании. Обеспечение безопасности теста и результатов тестирования. Отчет.

Этическиепроблемы связанные с отбором, процедурой, обработкой, интерпретацией, хранением информацией и отчетом.

# Образовательные технологии

Лекции, семинары, самостоятельная работы, письменные самостоятельные работы, практические занятия, работа по группам, общегрупповая дискуссия, презентации и их защита, мастер-классы

# Оценочные средства для текущего контроля и аттестации аспиранта

Для текущего контроля используются контрольные работы и домашние задания.

## Тематика заданий текущего контроля

Примерные вопросы контрольных работ:

1. В чем отличие научных теорий от ненаучных?
2. В чем отличие измерения в социальных науках от измерения в физических науках?
3. Что такое операциональное определение и какова его роль в процедуре и качестве измерения?
4. Как соотносятся шкалы измерения с возможностями последующего анализа данных?
5. Какие статистические процедуры работы с допустимы данными, полученными в разных шкалах измерения?
6. Как интерпретируется корреляция?
7. В чем суть нормо-ориентированного и критериально-ориентированного тестирования и как эти различия отражаются на спецификации теста?
8. Что такое линейные и нелинейные преобразования тестовых баллов и в каких случаях нецелесообразно прибегать к нелинейным преобразованиям?
9. Что такое надежность теста и надежность тестового задания, и каковы процедуры ее установления?
10. Что такое валидность теста и каковы процедуры ее установления?
11. В чем принципиальные отличия классической тестовой теории от IRT?
12. Какие основные проблемы возникают при подсчете тестовых баллов?
13. Каковы современные тенденции в измерении способностей и достижений человека?
14. Какие основные проблемы возникают при валидизации личностных опросников?
15. Какие типы администрации теста существуют и при каких видах измерения обычно их используют?
16. Какова связь смещения (bias) тестовых заданий и этических требований к процедуре измерения?

# Вопросы для оценки качества освоения дисциплины

Задания для итоговой оценки качества освоения дисциплины представляют собой тестовые задания закрытого и полузакрытого вида. Примеры заданий представлены в п.9.1.

## Примеры заданий промежуточного /итогового контроля

1. Если после проведения теста обнаруживается, что тест был слишком сложен для аспирантов, можно «подогнать» результаты тестирования под нормативы, добавляя, скажем, 10 баллов к каждому достигнутому баллу. Какая статистическая мера изменится после этой манипуляции?
2. медиана
3. корреляция баллов с другими переменными
4. стандартное отклонение
5. диапазон полученных баллов
6. Каков 95% доверительный интервал для достигнутого балла 118 по тесту Векслера (среднее 100, SD = 15), если надежность теста 0,95?
7. 109 – 127
8. 115 – 121
9. 111 – 125
10. 117 – 119
11. Какова стандартная ошибка измерения теста, если его надежность 0,00?
    1. ноль
    2. плюс-минус один балл
    3. стандартное отклонение тестовых баллов
    4. равна коэффициенту надежности, каковым бы он ни был
12. Какая статистика дает описание величины согласованности тестовых баллов по двум эквивалентным формам или версий одного и того же теста?

A. надежность параллельных форм

B. прогностическая надежность

C. внутренняя согласованность

D. ретестовая надежность

1. Какой вид надежности говорит об изменениях в экзаменуемых как источнике ошибки, который и снижает надежность оценки?
   1. ретестовая надежность
   2. надежность экспертов
   3. расщепленная надежность
   4. внутренняя согласованность
2. Какое утверждение о тестовых баллах экзаменуемых будет корректно, когда мы говорим о том, что баллы остаются более или менее теми же самыми между двумя проведениями одной и той же формы теста?
   1. высокая точность
   2. надежность баллов высокая
   3. баллы валидны
   4. содержательная валидность высокая
3. Какие из следующих утверждений являются верными в отношении коэффициентов надежности?
   1. Более короткие тесты являются более надежными, чем длинные
   2. Надежность баллов возрастает, когда баллы становятся более гомогенными
   3. Тесты средней трудности являются более надежными, чем очень трудные тесты
   4. При ошибках измерения, таких как угадывание или ошибочные задания, надежность теста будет высока. В этом и заключается надежность.
4. Если надежность теста оценена в 0,70, какой процент дисперсии тестовых баллов объясняется различными источниками ошибок, связанными с: (1) тестом, (2) проведением теста, (3) людьми, участвующими в тестировании?
   1. 49%
   2. 70%
   3. 30%
   4. 51%
5. Хорошо известно, что тестовые баллы не бывают полностью надежными. В связи с этим, какой способ действий из тех, что перечислены ниже, является подходящим для людей, использующих тестовые баллы (учителя, консультанты, психологи)?
   1. Интерпретировать только очень высокие или очень низкие баллы
   2. Использовать интервью и наблюдение вместо тестов, для более обоснованных решений
   3. Повторно тестировать людей, чтобы получать более согласованную информацию
   4. Использовать доверительные интервалы или учитывать ошибку измерения при интерпретации балов

# Порядок формирования оценок по дисциплине

Оценивается работа на занятиях по обязательной литературе, домашние задания и контрольные работы. На оценку влияет: правильность выполнения заданий контрольных и домашних работ (и их своевременная сдача), демонстрация и знания заданного материала, активность на занятиях.

Преподаватель оценивает работу аспирантов на семинарских и практических занятиях: - воспроизведение прочитанного материала, демонстрация способности видеть связи между основными понятиями, задавание вопросов, отвеы на вопросы преподавателя, участие в дискуссиях и в групповой работе. Оценки за работу на семинарских и практических занятиях преподаватель выставляет в рабочую ведомость. Накопленная оценка по 10-ти балльной шкале за работу на семинарских и практических занятиях определяется перед промежуточным или итоговым контролем - *Оаудиторная*.

Преподаватель оценивает самостоятельную работу аспирантов: оценивается правильность выполнения заданий домашних работ и их своевременная сдача, полнота выполнения, ответы на вопросы аудитории при докладах. Оценки за самостоятельную работу аспиранта преподаватель выставляет в рабочую ведомость. Накопленная оценка по 10-ти балльной шкале за самостоятельную работу определяется перед промежуточным или итоговым контролем – *Осам. работа*.

Накопленная оценка за текущий контроль учитывает результаты аспиранта по текущему контролю следующим образом:

*Отекущий* = *n1·Ок/р + n4·Окол + n5·Одз* ;

Способ округления накопленной оценки текущего контроля: арифметический (например, оценка 4,4 округляется до 4, а оценка 4,5 до 5.

Результирующая оценка за итоговый контроль в форме теста выставляется в 10бальной шкале.

Итоговая оценка складывается из суммы ткущей оценки и итогового теста и переводится в 10-бальную шкалу.

Аспирант не имеет возможности пересдать низкие результаты за текущий контроль или работу на занятиях, самостоятельную работу, если за любой из этих видов работы была выставлена оценка.

В случае если аспирант пропустил срок сдачи отчетности по текущему контролю или самостоятельной работе по уважительной причине (к уважительным причинам может быть отнесена болезнь в указанный период, подтвержденная справкой из медицинского учреждения) допускает сдача форм контроля не позднее чем через две недели с момента окончания действия справки, если другие сроки не установлены деканатом. В случае, если аспирант не сдает работу в указанный срок, ему выставляется оценка «0».

За каждый день просрочки сдачи домашней работы по неуважительной причине итоговая оценка за эту домашнюю работу снижается на 1 балл (по 10-бальной шкале)

Условия пересдач

На пересдаче аспиранту не предоставляется возможность получить дополнительный балл для компенсации оценки за текущий контроль. Первая пересдача проводится преподавателем, отвечающим за чтение дисциплины на факультете. При выставлении результирующей оценки учитывается накопленная оценка за текущий контроль, самостоятельную и аудиторную работу

*k2·Отекущий + k3·Осам. работа + k4·Оаудиторная*

Вторая пересдача проводится в присутствие комиссии, включающей не менее трех преподавателей, при выставлении результирующей оценки учитывается накопленная оценка за текущий контроль, самостоятельную и аудиторную работу

*k2·Отекущий + k3·Осам. работа + k4·Оаудиторная*

В исключительных случаях комиссия может принять решение об отказе от учета накопленной аспирантом оценки. В данном случае комиссией используются следующие критерии оценки ответа аспиранта: полнота ответа на вопрос, самостоятельность при ответе на вопрос, точность в ответе на поставленные вопросы.

В диплом выставляет результирующая оценка по учебной дисциплине, которая формируется по следующей формуле:

*Одисциплина = k1·Опромежуточный* + *k2·Опромежуточный* + *k3*·*Оитоговый*

Способ округления результирующей оценки по учебной дисциплине: арифметический (например, оценка 4,4 округляется до 4, а оценка 4,5 до 5).

# Учебно-методическое и информационное обеспечение дисциплины

## Основная литература

1. Анастази А., Урбина С. Психологическое тестирование. СПб.: Питер, 2003.
2. Крокер Л., Алгина Дж. Введение в классическую и современную теорию тестов. Москва,Логос 2010.
3. Поппер К. Предположения и опровержения. Рост научного знания. М.: АСТ, 2008.
4. Kaplan, R.M., & Saccuzzo, D.P. (2008). Psychological testing: Principles, applications, and issues (7th ed.). Wadsworth, Cengage Learning.

## Hambleton, R.K., Swaminathan, H. (1991). Fundamentals of Item Response Theory. CA: Sage.

## Дополнительная литература

1. Тюменева Ю.А. Психологическое измерение. М.,Аспект-Пресс. 2007.
2. American Educational Research Association (AERA), American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: AERA.
3. Gregory, R.J. (2007). Psychological testing: History, principles, and applications (5th edition). New York: Pearson.
4. Miller, M.D., Linn, R.L., & Gronlund, N.E. (2008). Measurement and assessment in teaching (10th ed.). Prentice Hall.
5. Thorndike, R. M., Thorndike-Christ, T.M. (2009). Measurement and evaluation in psychology and education (8th ed.). Pearson Prentice Hall.

## Программные средства

Для успешного освоения дисциплины, аспирант использует следующие программные средства:

* SPSS (версия не ниже 12)
* ITEMAN 4.2 Classical item analysis (студенческая версия)

# Материально-техническое обеспечение дисциплины

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

* ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
* мультимедийный проектор с дистанционным управлением.

Учебные аудитории для самостоятельных занятий по дисциплине оснащены ­­­­­­­­­­­­­­­­­­­­­­­­ ноутбуками, с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.