

§ 8. Коэффициенты корреляции Пирсона и Спирмена

Цель темы: освоение основных понятий и процедуры использования коэффициентов корреляции при проведении психологического исследования.

Задачи темы:

- знакомство с основными понятиями корреляционного анализа;
- анализ особенностей применения и процедуры подсчета коэффициентов корреляции;
- освоение пошаговой процедуры подсчета коэффициентов корреляции.

Теоретическое описание

Корреляция (от лат. *correlatio* – соотношение). **1.** Мера степени и направления связи между значениями двух переменных [8. С. 377]. **2.** Статистический показатель вероятности связи между двумя переменными, измеренными в количественной шкале [8. С. 138].

Корреляционный анализ. Проверка гипотез о связях между переменными с использованием коэффициентов корреляции [6. С. 147].

Направление связи. Определяется прямым или обратным соотношением значений двух переменных: если возрастанию значений одной переменной соответствует возрастание значений другой переменной, то взаимосвязь называется прямой (положительной); если возрастанию значений одной переменной соответствует убывание значений другой переменной, то взаимосвязь является обратной (отрицательной). Показателем направления связи является знак коэффициента корреляции [7. С. 67].

Коэффициент корреляции (r). **1.** Количественная мера силы и направления вероятностной взаимосвязи двух переменных; принимает значения в диапазоне от -1 до +1 [6. С. 67]. **2.** Мера прямой или обратной пропорциональности между двумя переменными [6. С. 147]. **3.** Двумерная описательная статистика, количественная мера взаимосвязи (совместимой изменчивости) двух переменных [6. С. 64].

Коэффициент Пирсона (r -Pearson). Мера корреляции, подходящая для двух непрерывных (метрических переменных), измеренных на одной и той же выборке [6. С. 67].

Коэффициент Спирмена (r -Spearman's). Мера корреляции, подходящая для двух переменных, измеренных в ранговой шкале [6. С. 64].

Сила связи. Достигает максимума при условии взаимно однозначного соответствия: когда каждому значению одной переменной соответствует только одно значение другой переменной (и наоборот), эмпирическая взаимосвязь при этом совпадает с функциональной линейной связью. [7. С. 67].

Вероятность (p -level; p ; от англ. *probability* – вероятность). **1.** Вероятность рассматривают как числовую меру объективной возможности осуществления события A при единичном испытании и обозначают символом P . Согласно классическому определению, вероятность события A выражается отношением числа благоприятствующих осуществлению этого события исходов m к числу всех равновероятных и несовместных исходов n , то есть $P(A)=m/n$ [3. С. 68]. **2.** Значение p – это, буквально, мера вероятности получения статистически значимого результата.

Статистическая значимость (*sig.*; от англ. *statistical signification* – статистическая значимость). **1.** Мера случайности полученного результата, равная вероятности того, что в генеральной совокупности этот результат (различия, связь) отсутствует. Чем меньше эта вероятность (значение p -уровня), тем выше статистическая значимость результата. Результат считается статистически достоверным (значимым), если p -уровень не превышает 0,05 [8. С. 376].

Линейная корреляция. Это изменение одной переменной на одну единицу, приводящее к изменению другой (график представляет прямую линию) [6. С. 65].

Ранговая корреляция. Связь, вычисляемая между переменными в случае если они представлены, или одна из них – в порядковой шкале, а другая в метрической [6. С. 77].

Монотонная связь. Направление изменения одной переменной не меняется с возрастанием (убыванием) другой переменной [6. С. 65].

Немонотонная связь. Направление изменения одной переменной меняется с возрастанием (убыванием) другой переменной [6. С. 65].

Исследовательские задачи:

Коэффициент Пирсона:

- анализ связи между уровнем тревожности школьников (тест тревожности *R. Temml, M. Dorca, V. Amen*) и уровнем протекции в процессе воспитания (опросник «Анализ семейных взаимоотношений» (АСВ) *Э.Г. Эйдемиллера, В.В. Юстицкиса*);
- анализ связи между уровнем интернальности-экстернальности (опросник парциальных позиций интернальности-экстернальности личности *Е.Ф. Бажина, Е.А. Голынкиной и А.М. Эткинда*) и уровнем эмоционального выгорания у педагогов (опросник эмоционального выгорания личности *В.В. Бойко*);
- анализ связи (структуры) элементов вербального интеллекта у школьников разных профилей обучения (методика «Школьный тест умственного развития» (ШТУР) *К.М. Гуревича* и др.).

Коэффициент Спирмена:

- анализ связи между уровнем эмпатии (методика диагностики уровня эмпатических способностей *В.В. Бойко*) и удовлетворенностью браком (опросник удовлетворенности браком *В.В. Столина, Т.Л. Романовой, Г.П. Бутенко*) в молодых семьях;
- анализ связи социометрического статуса подростков (социометрический тест *Jacob L. Moreno*) и особенностей стиля семейного воспитания (опросник «Анализ семейных взаимоотношений» (АСВ) *Э.Г. Эйдемиллера, В.В. Юстицкиса*);
- анализ структуры жизненных целей подростков из полных и неполных семей (опросник «Жизненные цели» *Edward L. Deci, Richard M. Ryan Ryan*).

Требования к выборке:

- *объем выборки исследования:*
 - для использования r -Пирсона объем выборки $n_1 \geq 30$ и $n_2 \geq 30$;
 - для использования r -Спирмена объем выборки $n_1 \geq 5$ и $n_2 \geq 5$;
- *распределение:*
 - для использования r -Пирсона должно соответствовать нормальному виду;
 - для использования r -Спирмена соответствие нормальному виду необязательно.

ПОШАГОВОЕ ВЫПОЛНЕНИЕ

Файл-пример:

Откройте файл **SPSS Коэффициенты корреляции.sav** в программе IBM SPSS Statistics 19.

В файле представлены результаты диагностики агрессивности и черт личности учащихся 6 класса (методики: А. Басса-А. Дарки, 16 PF Кеттелла; $n=101$).

В анализе участвуют следующие переменные:

- пол :** 1 – женский, 2 – мужской;
ФА : физическая агрессия;
КА : косвенная агрессия;
Р : раздражительность;
Н : негативизм;
О : обида;
П : подозрительность;
ВА : вербальная агрессия;
ЧВ : чувство вины;
А : доброта;
В : интеллект;

- С : эмоциональная устойчивость;
- Е : настойчивость;
- Ф : беспечность;
- Г : совестливость;
- Н : смелость;
- І : мягкосердечность;
- О : чувство вины;
- Q2 : самостоятельность;
- Q3 : самоконтроль;
- Q4 : напряженность.

I. Коэффициент линейной корреляции Пирсона

Выполните следующий порядок действий:

- Шаг 1** На панели инструментов выберите меню **Анализ**→**Корреляции**→**Парные**.
- Шаг 2** В открытом меню **Парные корреляции** (рис. 34) перенесите из левого окна переменную **вербальная агрессия** в окно **Переменные:** и повторите данный порядок действий для переменных **доброта**, **интеллект**, **эмоциональная устойчивость**, **настойчивость** и **беспечность**.
- Шаг 3** Обратите внимание, что в открытом окне **Парные корреляции** по умолчанию установлены три команды, которые в примере с подсчетом **r-Пирсона** останутся неизменными:
- а) в группе команд **Коэффициенты корреляции** установлена галочка для **r-Пирсона**,
 - б) в группе команд **Критерий значимости** установлена галочка для **Двухсторонний**,
 - в) в самой нижней части окна установлена галочка для команды **Метить значимые корреляции**.

Нажимая команду **ОК**, мы, таким образом, подсчитаем коэффициент корреляции **Пирсона** с предположением о двухстороннем типе связи (то есть как положительной, так и отрицательной связи), и пометим в таблице **Корреляции** окна **Вывод** статистически значимые корреляции.

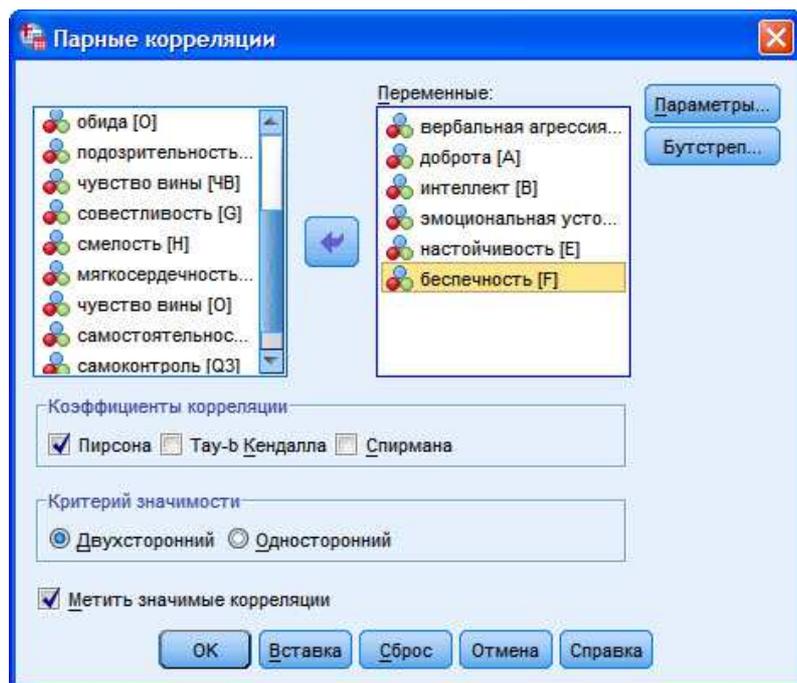


Рис. 34. Меню Парные корреляции

ОПИСАНИЕ И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ

1) В открывшемся окне **Вывод** представлены результаты анализа связи между уровнем развития **вербальной агрессии** (одноименная переменная) и уровнем развития черт личности учащихся 6 классов (переменные **доброта**, **интеллект**, **эмоциональная устойчивость**, **настойчивость** и **беспечность**).

Описанию и интерпретации подлежит таблица под заголовком **Корреляции** (рис. 35; чаще данный тип таблиц называют **Корреляционной матрицей**).

		Корреляции					
		вербальная агрессия	доброта	интеллект	эмоциональная устойчивость	настойчивость	беспечность
вербальная агрессия	Корреляция Пирсона	1	,105	,013	-,157	,441**	,283**
	Знч.(2-сторон)		,294	,895	,117	,000	,004
	N	101	101	101	101	101	101
доброта	Корреляция Пирсона	,105	1	,076	,137	,104	,434**
	Знч.(2-сторон)	,294		,449	,171	,301	,000
	N	101	101	101	101	101	101
интеллект	Корреляция Пирсона	,013	,076	1	,086	,130	,109
	Знч.(2-сторон)	,895	,449		,392	,196	,280
	N	101	101	101	101	101	101
эмоциональная устойчивость	Корреляция Пирсона	-,157	,137	,086	1	-,046	,050
	Знч.(2-сторон)	,117	,171	,392		,645	,619
	N	101	101	101	101	101	101
настойчивость	Корреляция Пирсона	,441**	,104	,130	-,046	1	,302**
	Знч.(2-сторон)	,000	,301	,196	,645		,002
	N	101	101	101	101	101	101
беспечность	Корреляция Пирсона	,283**	,434**	,109	,050	,302**	1
	Знч.(2-сторон)	,004	,000	,280	,619	,002	
	N	101	101	101	101	101	101

** Корреляция значима на уровне 0.01 (2-сторон).

Рис. 35. Корреляционная матрица

Ввиду сложности **Корреляционной матрицы**, остановимся на ее содержании подробнее:

а) В матрице представлены корреляции шести переменных с шестью – в группе строк для изучаемых переменных (например, в строке **вербальная агрессия**) три подстроки обозначают следующее:

- строка **Корреляция Пирсона** – это значения коэффициента корреляции Пирсона,
- строка **Знч.(2-сторон)** – это p -уровень коэффициента корреляции Пирсона,
- строка **N** – это число испытуемых, участвующих в исследовании.

Таким образом, для того, чтобы ответить на вопрос – какой уровень и какова значимость корреляции между переменными **вербальная агрессия** и **настойчивость**, необходимо пересечь подстроки **Корреляция Пирсона** и **Знч.(2-сторон)** в строке **вербальная агрессия** и столбце **настойчивость** (на рисунке это обозначено кругом).

б) Программа SPSS производит подсчет корреляции по принципу «все со всеми», то есть каждая переменная коррелируется с самой собой и в матрице эта корреляция обозначается цифрами **1**. Таким образом по диагонали расположены корреляции, равные единице. Они, естественно, не подлежат анализу ввиду его бессмысленности (на рисунке 35 эта диагональ обозначена пунктирной линией).

в) На **Шаге 3** нами была сохранена по умолчанию установленная галочка для команды **Метить значимые корреляции**. В матрице значимые корреляции (то есть корреляции, p -уровень которых $\leq 0,05$) обозначены звездочкой - *:

- преимущество такого обозначения в том, что при создании большой матрицы (например, 15x15) исследователю хорошо видны значимые корреляции, что ускоряет процедуру анализа результатов;

- недостаток состоит в том, что программа SPSS по умолчанию отмечает только p -уровни двух типов – $p \leq 0,05$ (одна звездочка) и $p \leq 0,01$ (две звездочки).

Однако чаще всего в психологических исследованиях диапазон значимости корреляции $p \leq 0,05 \geq 0,001$ делится на три уровня значимости:

- $p \leq 0,05$ (одна звездочка - *),
- $p \leq 0,01$ (две звездочки - **),
- $p \leq 0,001$ (три звездочки - ***).

Таким образом, ввиду ограничений возможностей программы SPSS, рекомендуем на **Шаге 3** снять галочку для команды **Метить значимые корреляции**.

г) Интерпретация **коэффициента корреляции** производится исходя из уровня силы связи:

- $r > 0,01 \leq 0,29$ – **слабая** положительная связь,
- $r > 0,30 \leq 0,69$ – **умеренная** положительная связь,
- $r > 0,70 \leq 1,00$ – **сильная** положительная связь,
- $r > -0,01 \leq -0,29$ – **слабая** отрицательная связь,
- $r > -0,30 \leq -0,69$ – **умеренная** отрицательная связь,
- $r > -0,70 \leq -1,00$ – **сильная** отрицательная связь.

д) Интерпретация p -уровня **коэффициента корреляции** производится аналогично тому, как это делалось для параметрических и непараметрических критериев:

- если p -уровень $\leq 0,05$, то связь между переменными является статистически значимой;
- если p -уровень $> 0,05$, то связь между переменными является статистически не значимой.

е) Также при интерпретации p -уровня **коэффициента корреляции** важным является не только сам факт значимости, но и ее уровень. Традиционно p -уровень корреляции дифференцируется на три уровня:

- $p \leq 0,05 > 0,01$ – **низкая** статистическая значимость,
- $p \leq 0,01 > 0,001$ – **средней силы** статистическая значимость,
- $p \leq 0,001$ – **высокая** статистическая значимость.

2) В нашем примере результат анализа следующий:

- для корреляции переменных **вербальная агрессия** и **настойчивость**: в процессе корреляционного анализа была выявлена умеренная положительная связь высокой степени значимости между уровнем развития **вербальной агрессии** и **настойчивостью** – r -Пирсона = 0,44 при $p \leq 0,001$.

Описание полученного результата произвести достаточно просто: чем выше вербальная агрессия у детей в 6 классе, тем выше уровень их настойчивости.

Интерпретация описанного результата требует от исследователя учета как минимум двух дополнительных условий: знания возрастных особенностей детей подросткового возраста и полного анализа рассматриваемой в примере **Корреляционной матрицы**.

Тем не менее, попытаемся проинтерпретировать полученную корреляцию: человек с высокими показателями по **фактору Е** в различных ситуациях взаимодействия чаще проявляет независимость, самоуверенность, хвастливость, тщеславность, конфликтность, своенравность, нетрадиционность, мятежность, бесцеремонность, смелость, авантюристичность и т.д. Данные особенности личности особо остро проявляются именно в подростковом возрасте. Учитывая, что здесь ведущей деятельностью является общение со сверстниками (по Д.Б. Эльконину), можно предположить, что вербальные формы общения являются для подростка наиболее удобным средством решения конфликтных ситуаций. Последние, в свою очередь, возникают как результат повышенного в подростковом возрасте максимализма требований к себе и другим, противоречия между желаниями и возможностями и т.д.

Исследователю важно понимать, что здесь он сталкивается с проблемой объяснения в психологии. Не имея в рамках данной работы возможности подробно остановиться на данном вопросе, мы отсылаем заинтересованного читателя к источникам [4; 9; 10; 11; 12; 13].

Логика описания и интерпретации корреляций остальных переменных в общих чертах соответствуют рассмотренному примеру.

II. Коэффициент ранговой корреляции Спирмена

Выполните следующий порядок действий:

- Шаг 1** На панели инструментов выберите меню **Анализ→Корреляции→Парные**.
- Шаг 2** В открытом меню **Парные корреляции** (рис. 34) перенесите из левого окна переменную **вербальная агрессия** в окно **Переменные:** и повторите данный порядок действий для переменных **совестливость, смелость, мягкосердечность, чувство вины и самостоятельность**.
- Шаг 3** а) для того, чтобы использовать коэффициент корреляции **г-Спирмена** в группе команд **Коэффициенты корреляции** снимите галочку для команды **Пирсона** и установите галочку для команды **Спирмана**;
б) в группе команд **Критерий значимости** сохраните выбор для двухстороннего критерия,
в) в самой нижней части окна **Парные корреляции** снимите галочку для команды **Метить значимые корреляции** и нажмите **ОК**.

ОПИСАНИЕ И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ

1) Описание и интерпретация результатов **корреляционного анализа** с использованием коэффициента **г-Спирмена** идентичны анализу результатов подсчета коэффициента **г-Пирсона**.

ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Задание 1.

Файл-пример: SPSS Коэффициенты корреляции.sav

Откройте файл **SPSS Коэффициенты корреляции.sav** в программе IBM SPSS Statistics 19.

- 1) Подсчитайте корреляцию **г-Пирсона** между переменными **раздражительность, самостоятельность, самоконтроль и напряженность**.
- 2) Опишите и проинтерпретируйте полученный результат исходя из рассмотренного в параграфе способа анализа данных.

Задание 2.

Файл-пример: SPSS Коэффициенты корреляции.sav

Откройте файл **SPSS Коэффициенты корреляции.sav** в программе IBM SPSS Statistics 19.

- 1) Подсчитайте корреляцию **г-Спирмена** между переменными **обида, доброта, интеллект и эмоциональная устойчивость**.
- 2) Опишите и проинтерпретируйте полученный результат исходя из рассмотренного в параграфе способа анализа данных.

КРАТКОЕ ОПИСАНИЕ ИСПОЛЬЗУЕМЫХ ПРОЦЕДУР АНАЛИЗА

1) Условия выбора коэффициентов корреляции.

Принятие решения о выборе того или иного **коэффициента корреляции** основывается на двух уже рассмотренных нами в § 5 и § 6 критериях, а также на особенностях их взаимосвязи.

Первым критерием является нормальность распределения:

- если распределение значений является нормальным, то принимается решение о выборе коэффициента линейной корреляции **г-Пирсона**;
- если распределение значений не является нормальным, то принимается решение о выборе коэффициента ранговой корреляции **г-Спирмена**.

Вторым критерием является объем выборки:

- если $n \geq 30$, то принимается решение о выборе коэффициента линейной корреляции **г-Пирсона**;
- если $n < 30$, то принимается решение о выборе коэффициента ранговой корреляции **г-Спирмена**.

В тех случаях, когда $n \geq 30$, но распределение значений не является нормальным, принимается решение о выборе коэффициента ранговой корреляции **г-Спирмена**.

Когда же объем выборки $n < 30$, то проводить проверку на нормальность распределения значений не имеет смысла. Поэтому в данном случае принимается решение о выборе коэффициента ранговой корреляции **г-Спирмена**.

2) Линейные и ранговые корреляции.

Принимая решение о выборе типа корреляции, при интерпретации результатов важно помнить и учитывать, что линейные корреляции являются более точными, чем ранговые. Ранжирование значений при использовании г-Спирмена естественным образом снижает меру индивидуальной изменчивости измеренного на испытуемых показателя.

3) Сила связи.

Принято считать, что сила коэффициента корреляции, как один показателей меры связи, дифференцируется на три уровня как для положительных, так и для отрицательных корреляций:

- $r > 0,01 \leq 0,29$ – **слабая** положительная связь,
- $r > 0,30 \leq 0,69$ – **умеренная** положительная связь,
- $r > 0,70 \leq 1,00$ – **сильная** положительная связь,
- $r > -0,01 \leq -0,29$ – **слабая** отрицательная связь,
- $r > -0,30 \leq -0,69$ – **умеренная** отрицательная связь,
- $r > -0,70 \leq -1,00$ – **сильная** отрицательная связь.

4) «Связь» или «Влияние»?

Зачастую исследователи при интерпретации результатов корреляционного анализа используют термин **влияние**, описывая значение показателя **г**. Например, такая фраза, как «было выявлено влияние уровня интеллекта на личностную тревожность», является следствием понимания буквального снижения уровня личностной тревожности при повышении уровня IQ. Здесь исследователь принимает уровень IQ за причину, а личностную тревожность за следствие.

Тем не менее, даже если показатель $r \geq 0,70$ при $p \leq 0,001$, говорить о строгой детерминации, или о наличии функциональной связи нельзя. Это связано с рядом принципиальных моментов:

а) Функциональные связи характерны только для объектов, изучаемых естественными науками. В зоологии это, например, влияние периодичности корма на удои коров; в ботанике – влияние частоты и объема полива на длину стебля; в физиологии – влияние роста на массу тела и т.д.

Однако, даже здесь наличие строгих функциональных связей ($r > 0,70$) ограничивается действием промежуточных, неучтенных переменных, снижающих влияние независимой переменной (периодичность корма животных, частота и объем полива растения, рост человека) на зависимую.

б) В психологии предмет исследования гораздо более сложен, чем в естественных науках. Однако поиск связей между психологическими переменными заставляет исследователей искать аналогии в естественных науках при интерпретации результатов, например:

- выделять независимые и зависимые переменные,
- говорить о большей или меньшей значимости той или иной независимой переменной,
- найдя высокий коэффициент корреляции, говорить о влиянии независимой переменной на зависимую.

В связи с тем, что предмет психологии практически не поддается однозначному объяснению, выделять строгие функциональные, причинно-следственные связи между переменными не имеет смысла. Если одна зависимая переменная (например, настроение) статистически значимо сильно коррелирует с другой независимой переменной (например, принятие себя), то обязательно, как на зависимую, так и на независимую переменные будут оказывать воздействие другие независимые переменные (например, на настроение может воздействовать тип темперамента, на принятие себя – самооценка). Поэтому однозначного, прямого влияния уровня принятия себя на настроение нет.

в) Употреблять термин влияние вместо связи при интерпретации результатов исследования можно только в том случае, если при предварительном теоретическом анализе логически была установлена строгая причинно-следственная связь между двумя переменными. В иных случаях использование термина влияние является некорректным.

г) В связи с тем, что корреляционный анализ не позволяет в полной мере устанавливать однозначные связи между переменными, следует обращаться либо к структурному анализу данных, либо к многомерным методам (например, множественный регрессионный анализ).

ЛИТЕРАТУРА ДЛЯ ДОПОЛНИТЕЛЬНОЙ ПОДГОТОВКИ

1. Афанасьев, В. В. Теория вероятностей [Текст] / В. В. Афанасьев. – М.: ВЛАДОС, 2007. – 350 с.
2. Бурлачук, Л. Ф. Словарь-справочник по психодиагностике [Текст] / Л. Ф. Бурлачук, С. М. Морозов. – СПб.: Питер, 2001. – 528 с.
3. Лакин, Г. Ф. Биометрия [Текст] / Г. Ф. Лакин; изд. 4-е, перераб. и доп. – М.: Высшая школа, 1990. – 352 с.
4. Мазиллов, В. А. Научная психология [Текст]: проблема объяснения / В. А. Мазиллов // Методология и история психологии. – 2008. – Том 3. – Вып. 1. – С. 58-73.
5. Математическая энциклопедия [Текст] / гл. ред. И. М. Виноградов; в 5 тт. – М.: Советская энциклопедия, 1977-1985.
6. Наследов, А. Д. Математические методы психологического исследования. Анализ и интерпретация данных [Текст] / А. Д. Наследов. – СПб.: Речь, 2012. – 392 с.
7. Наследов, А. Д. Математические методы психологического исследования. Анализ и интерпретация данных [Текст] / А. Д. Наследов. – СПб.: Речь, 2004. – 392 с.
8. Наследов, А. Д. SPSS 19. Профессиональный статистический анализ данных [Текст] / А. Д. Наследов. – СПб.: Питер, 2011. – 400 с.
9. Никитин, Е. П. Объяснение – функция науки [Текст] / Е. П. Никитин. – М.: Наука, 1970. – 280 с.
10. Юревич, А. В. Объяснение в психологии [Текст] / А. В. Юревич // Психологический журнал. – 2006. - №1. – С. 97-106.
11. Юревич, А. В. Проблема объяснения в психологии [Текст] / А. В. Юревич // Методология и история психологии. – 2008. – Том 3. – Вып. 1. – С. 74-87.
12. Юревич, А. В. Психология и методология [Текст] / А. В. Юревич. – М.: Изд-во «Институт психологии РАН», 2005. – 312 с.
13. Экспериментальная психология [Текст] / ред.-сост. П. Фресс, Ж. Пиаже. Вып. I и II; общая ред. А.Н. Леонтьева – М.: Издательство «Прогресс», 1966. – Глава III. Характер объяснения в психологии и психофизиологический параллелизм. – С. 157-194.