

КОНТЕНТ-АНАЛИЗ МЕТОДИКИ «НЕЗАКОНЧЕННЫЕ ПРЕДЛОЖЕНИЯ»: ПРЕДСТАВЛЕНИЕ И ОБРАБОТКА ОЦЕНОК НЕСКОЛЬКИХ ЭКСПЕРТОВ*

М. К. АБРАМЯН¹, Т. В. ТУЛУПЬЕВА²

¹Санкт-Петербургский государственный университет, ²Санкт-Петербургский институт информатики и автоматизации РАН

¹Мат.-мех. ф-т СПбГУ, Библиотечная площадь, д. 2, Санкт-Петербург, 198904;

²СПИИРАН, 14-я линия ВО, д. 39, Санкт-Петербург, 199178

¹<abramian@yandex.ru>, ²<tvvt@iiias.spb.su>

УДК 681.3

Абрамян М. К., Тулупьева Т. В. **Контент-анализ методики «Незаконченные предложения»: представление и обработка оценок нескольких экспертов** // Труды СПИИРАН. Вып. 3, т. 2. — СПб.: Наука, 2006.

Аннотация. В статье описываются процесс сбора и анализа экспертных оценок результатов теста незаконченных предложений (9480 высказываний, 5 экспертных оценок на каждое высказывание), разработанные средства его автоматизации, а также приводятся полученные статистические выводы. — Библ. 6 назв.

UDC 681.3

Abramyan M. K., Tulupyeva T. V. **Content Analysis of Incomplete Sentences Data Set: Representation and Processing of Several Expert's Estimates** // SPIIRAS Proceedings. Issue 3, vol. 2. — SPb.: Nauka, 2006.

Abstract. We describe collecting and further analysis of expert's estimates regarding the results of an incomplete sentences test (9480 sentences, 5 expert's estimate for each) as well as the means used for automation of these processes. We also present the most important statistical conclusions. — Bibl. 6 items.

1. Введение

Устойчивый интерес психологов к проективной диагностике сохраняется уже более века. Различные проективные методики широко используются в практике исследования личности во всех областях современной психологии. С их помощью не только оценивают психологические особенности личности в процессе консультирования или практического исследования, но нередко они служат рабочим инструментом для проверки тех или иных теоретических положений [1].

Проективные методики представляют собой специальную технику клинико-экспериментального исследования тех особенностей личности, которые наименее доступны непосредственному наблюдению или опросу [2].

Общими для всех проективных методик являются следующие признаки:

- 1) неопределенность, неоднозначность стимульного материала или инструкции к заданию;
- 2) отсутствие ограничений в выборе ответа;

*Работа выполнена в рамках исследовательского проекта РГНФ **05-06-06322а** «Исследование психологической защиты в юношеском возрасте».

3) отсутствие оценки ответов испытуемых как «правильных» и «ошибочных» [1].

Одним из видов проективных методик является тест незаконченных предложений Дж. М. Сакса и С. Леви. Тест состоит из 60 предложений и измеряет 15 показателей: 1) отношение к отцу, 2) отношение к себе, 3) нереализованные возможности, 4) отношение к подчиненным, 5) отношение к будущему, 6) отношение к вышестоящим лицам, 7) страхи и опасения, 8) отношение к друзьям, 9) отношение к своему прошлому, 10) отношение к лицам противоположного пола, 11) сексуальные отношения, 12) отношение к семье, 13) отношение к сотрудникам, 14) отношение к матери, 15) чувство вины. В тесте предлагается серия незаконченных предложений, например: «Будущее кажется мне...», «Думаю, что настоящий друг...» и т. п. Испытуемому предлагается продолжить и завершить предложение.

Анализ и обработка большого объема данных, полученных в ходе опроса по этой методике, сопряжено с определенными трудностями: невозможно «напрямую» обработать статистическими методами набор ответов, представленных испытуемыми на естественном языке. Требуется произвести контент-анализ, т.е. привлечь экспертов, которые на основе своего *понимания* текста оценили бы каждый ответ по заданной шкале.

В данном исследовании были автоматизированы: внесение ответов респондентов, их интеграция в единую базу данных, тиражирование базы данных с ответами, предъявляемыми нескольким экспертам, сама работа экспертов по оцениванию, интеграция оценок экспертов в единую базу данных, анализ сходства и различий полученных экспертных оценок, расчёт общей оценки по оценкам экспертов. Описание программного обеспечения и алгоритмов, потребовавшихся на указанных этапах автоматизации, и является целью настоящей статьи. Кроме того, приводятся результаты статистического анализа по результатам теста незаконченных предложений и их краткая интерпретация с точки зрения предметной области [4, 5, 6].

2. Опрос и полученные данные

В рамках исследовательского проекта «Исследование психологической защиты в юношеском возрасте» было опрошено 158 студентов высшего учебного заведения (как технических, так и гуманитарных специальностей): 166 юношей и 71 девушка. Им был предложен тест незаконченных предложений Дж. М. Сакса и С. Леви.

Анализ ответов и отнесение их к категориям происходил следующим образом. Продолжение предложения оценивалось экспертом по следующей шкале: –2 (резко негативное высказывание), –1, 0 (нейтральное высказывание), +1, +2 (очень позитивное высказывание).

Каждому из 15 показателей соответствует ровно по 4 предложения. Показатель рассчитывался как сумма баллов, приписанным экспертом 4-м соответствующим показателю предложениям.

Исследование содержало следующие задачи:

- 1) выделение единиц анализа в дополненных предложениях;
- 2) категоризация единиц анализа;
- 3) статистический анализ.

Важным и сложным этапом в обработке данных этого теста является проверка степени согласованности экспертных оценок на основании выделенных критериев оценки. В ходе исследования был создан программный инструмент для упрощения этого этапа.

3. Электронная регистрация ответов

Для электронной регистрации первичных ответов испытуемых на тест незаконченных предложений был создан программный интерфейс (приложение в среде Borland Delphi 7) и база данных (MS Access 2003). Вначале незаконченные предложения заносились отдельно в файл MS Excel, в частности, и для того, чтобы можно было осуществлять проверку спеллинга вносимых ответов (рис. 1). Имя файла соответствует номеру анкеты, в поле А указывается номер предложения, в поле В продолжение предложения (ответ испытуемого). Затем с помощью приложения (рис. 2) ответы из всех анкет объединялись в общую базу данных (рис. 3).

В базе данных созданы 3 таблицы. Первая таблица называется *Persons*. В ней одно ключевое поле — номер анкеты (ID). Вторая таблица *IncSentences* состоит из 8 полей: ID — номер анкеты, N — номер вопроса, R — продолженное предложение, Exp1, Exp2, Exp3, Exp4, Exp5 — оценки соответствующих экспертов. Третья таблица — *SentenceBeginning*. Состоит из двух полей: N — номер предложения, и *Beginning* — начала предложения. Структуры таблиц *IncSentences* и *SentenceBeginning* показаны соответственно на рис. 4 и 5.

4. Процедура получения экспертных оценок

После того как все анкеты были интегрированы в одну базу данных, была проведена экспертная оценка каждого предложения по следующей инструкции:

Оценка «-2» — крайне негативные высказывания. Содержит слова: слишком или очень в применении к негативно окрашенным глаголам, ненавижу, несчастный, и тому подобные эмоционально насыщенные слова. *Оценка «-1»* — негативные высказывания. Фразы, содержащие следующие слова: было бы лучше, могла бы исправить, грустный, плохо, и тому подобные высказывания. Сюда же засчитывается «уход от ответа» и «штампы». *Оценка «0»* — нейтральные, незмоциональные, неинформативные высказывания. *Оценка «+1»* — позитивно окрашенные высказывания. Фразы, содержащие следующие слова: хорошо, меня устраивает, любить, нравиться. *Оценка «+2»* — крайне позитивно окрашенные высказывания. Фразы, содержащие следующие слова: «очень», «слишком» и подобные им в применении к позитивно окрашенным глаголам и словам, перечисленным в пункте для «+1».

Для проведения оценивания экспертами предложений создавались отдельные базы данных по схеме сходные с вышеописанной базой данных, но определенным образом привязанные к номеру эксперта (рис. 6). Число таких баз данных равно числу экспертов. В каждой из них создавалась дополнительная таблица *Expert*, содержащая идентификатор эксперта. Эти отдельные базы данных и приложение для оценивания предложений (рис. 7) распределялись между экспертами.

Microsoft Excel - 10001

Файл Правка Вид Вставка Формат Сервис

ВБ1

	A	B	C	D
1		1 прыгал на скалке		
2		2 я зеленою		
3		3 стать большим и сильным		
4		4 то сделал бы все для сотрудников		
5		5 лучшим чем настоящие		
6		6 часто поступает неправильно		
7		7 больших собак		

Рис. 1. Электронная регистрация ответов на тест в MS Excel

Перенос данных из Экселя в БД

Open folder with MS Excel files

Max Length of Str <>

Current File <>

Current ID <>

Connect DB

Disconnect DB

Expert On Duty: Exp2

Expert's estimates

ID	N	R
10001	1	прыгал на скалке
10002	2	я зеленою
10003	3	стать большим и сильным
10004	4	то сделал бы все для сотрудников
10005	5	лучшим чем настоящие
10006	6	часто поступает неправильно
10007	7	больших собак

Рис. 2. Приложение для объединения всех анкет в одну базу данных.

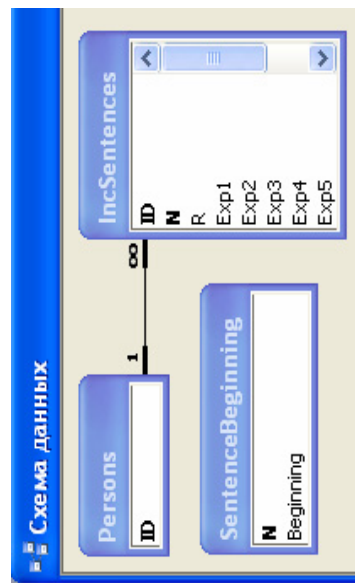


Рис. 3. Схема базы данных

IncSentences : таблица

ID	N	R	Exp1	Exp2	Exp3	Exp4	Exp5
10001	1	прыгал на скак	0	0	0	0	0
10001	2	я зеленою	0	0	0	0	0
10001	3	стать большим	0	0	0	0	0
10001	4	то сделал бы в	0	0	0	0	0
10001	5	лучшим чем н	0	0	0	0	0
10001	6	часто поступае	0	0	0	0	0
10001	7	больших собак	0	0	0	0	0

Рис. 4. Структура таблицы IncSentences.

SentenceBeginning : таблица

N	Beginning
1	Думаю, что мой отец редко
2	Если все против меня, то
3	Я всегда хотел(а)
4	Если бы я занимал(а) руково
5	Будущее кажется мне
6	Мое начальство
7	Знаю, что глупо, но боюсь

Рис. 5. Структура таблицы SentenceBeginning.

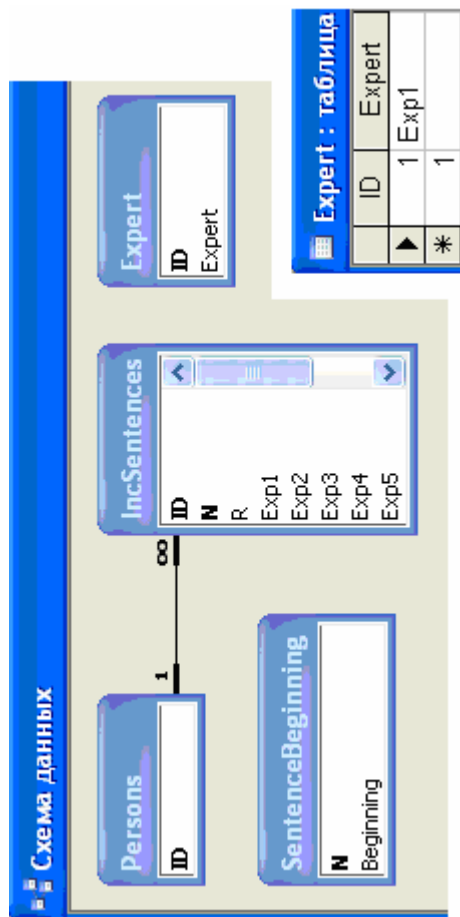


Рис. 6. Схема базы данных, в которую эксперты заносили свои оценки и структура таблицы Expert.

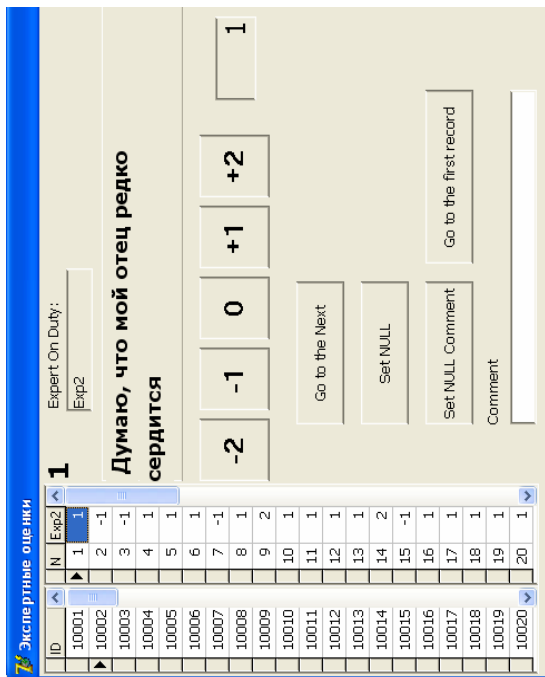


Рис. 7. Приложение для оценивания предложений.

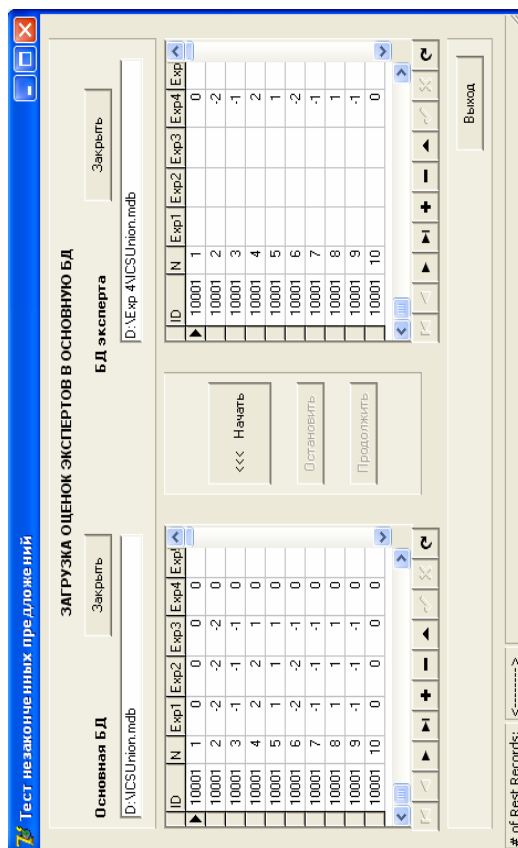


Рис. 8. Приложение для интеграции экспертных оценок в общую БД.

IncSentences : таблица

ID	N	R	Exp1	Exp2	Exp3	Exp4	Exp5
10001	1	прыгал на скак	0	0	0	0	0
10001	2	я зеленою	-2	-2	-2	-2	-2
10001	3	статья большим	-1	-1	-1	-1	-1
10001	4	то сделал бы в	2	2	1	2	2
10001	5	лучшим чем н	1	1	1	1	1
10001	6	часто поступае	-2	-2	-1	-2	-2
10001	7	больших собак	-1	-1	-1	-1	-1

Рис. 9. Таблица IncSentences с заполненными полями Exp1, Exp2, Exp3, Exp4, Exp5.

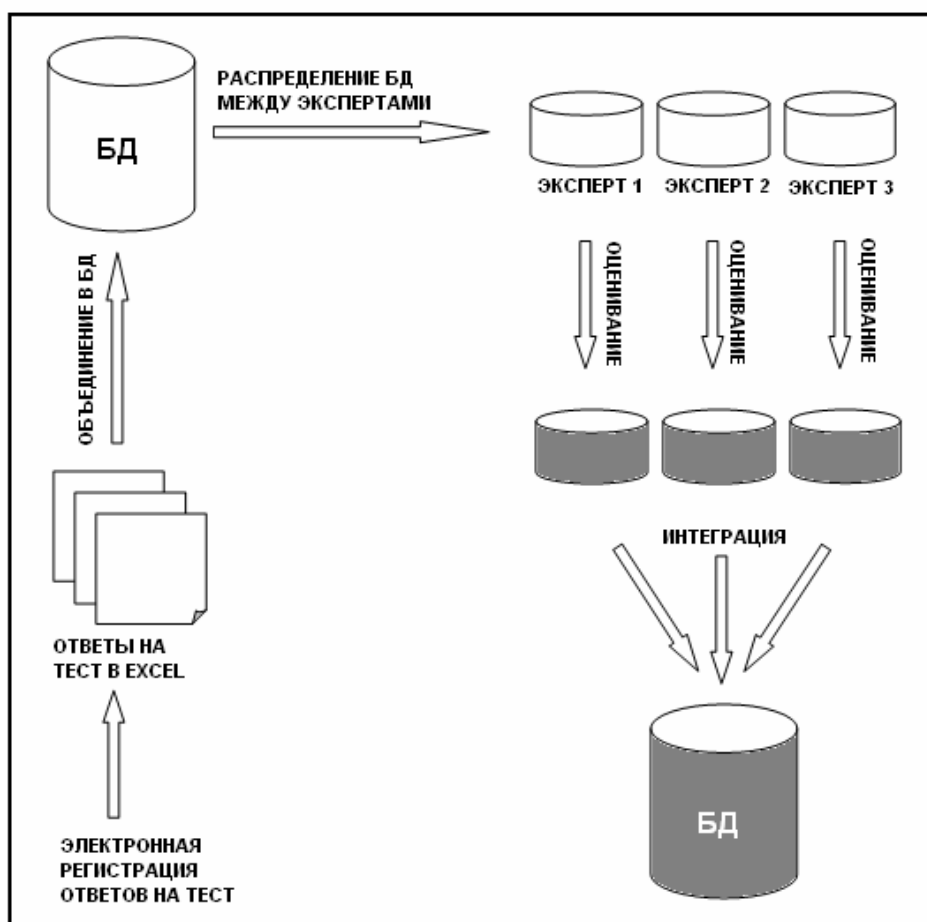


Рисунок 10. Схема этапов работы с данными.

Завершив оценивание предложений, эксперт возвращал свою базу данных. Далее экспертные оценки с помощью специального приложения (рис. 8) интегрировали в общую базу (рис. 3).

В итоге поля Exp1, Exp2, Exp3, Exp4, Exp5 в таблице IncSentences исходной базы данных содержат экспертные оценки (рис.9). Для облегчения восприятия на рис. 10 приведена схема описанных выше этапов работы с данными полученными по методике «незаконченные предложения».

5. Анализ экспертных оценок

В результате интеграции всех экспертных оценок создавалась база, с данными, которые уже можно было анализировать:

- 1) проверять согласованность оценок, обсуждать и вносить изменения в оценки;
- 2) определять значения психологических характеристик приведенных выше 15-ти показателей;
- 3) проводить статистический анализ.

Опишем подробнее этап проверки согласованности экспертных оценок.

Для проверки согласованности экспертных оценок было создано приложение (в среде Delphi 7).

В исследовании предложения оценивались пятью экспертами. Соответственно каждое предложение характеризовалось последовательностью 5 оце-

нок — целых чисел в интервале от -2 до 2 . Для того, чтобы определить несогласованные оценки и вывести их в таблицу для последующего обсуждения, вычислялась метрика близости оценок.

Обозначим число экспертов буквой N ; в нашем случае $N = 5$.

В приложении реализована возможность подсчета следующих метрик близости (показателей несогласованности):

$\max_{1 \leq i < j \leq N} |\text{exp}_i - \text{exp}_j|$ — Максимальная абсолютная разность между всеми экспертными оценками;

$\sum_{1 \leq i < j \leq N} |\text{exp}_i - \text{exp}_j|$ — сумма всех абсолютных разностей

между всеми экспертными оценками; $\frac{\sum_{1 \leq i < j \leq N} |\text{exp}_i - \text{exp}_j|}{N_{\text{diff}}}$ — средний модуль разности, где N_{diff} — количество таких разностей: $N_{\text{diff}} = \frac{(N-1)N}{2}$ (в нашем случае

$N_{\text{diff}} = 10$); $D = \frac{\sum_{i=1}^N (\text{exp}_i - \overline{\text{exp}})^2}{N}$ — дисперсия, где $\overline{\text{exp}} = \frac{1}{N} \sum_{i=1}^N \text{exp}_i$; \sqrt{D} — стандартное отклонение.

дартное отклонение.

Пользователь (эксперт) выбирает необходимый показатель несогласованности оценок (метрику близости), указывает нижнюю границу показателя, после чего приложение отбирает из базы данных и выводит в таблицу предложения с показателем несогласованности больше или равным нижней границы (рис. 11).

Отобранные записи из базы данных могут быть экспортированы в Excel и распечатаны для облегчения просмотра и обсуждения. После обсуждения между экспертами, если необходимо, вносятся изменения в оценки. Если эксперты не могут прийти к общему мнению, то последовательность оценок фиксируется как неустранимое разногласие. В дальнейшем можно задавать вывод всех несогласованных оценок, в том числе с пометкой «неустранимое разногласие», либо вывод всех несогласованных оценок, но без пометки «неустранимое разногласие», а можно вывести для просмотра только неустранимые разногласия в оценках.

Просматривать несогласованные оценки можно либо по анкетам либо по вопросам (рис. 12). В первом случае по заданной анкете выводятся все несогласованные оценки продолжений предложений из 60 возможных. Во втором случае по указанному номеру вопроса теста незаконченных предложений выводятся ответы всех испытуемых, которые требуют согласования экспертных оценок.

Рис. 11. Панель выбора записей.

По анкетам | По вопросам | По записям | Незаполненные поля

Надеюсь на то, что вопросы скоро закончатся

ID	N	Ехр1	Ехр2	Ехр3	Ехр4	Ехр5	С1	С2	С3	С4	С5
▶ 10001	▶ 20	-1	-1	-1	-1	1	2	8	0,8	0,64	0,8
10002	41	2									
10003	48	1									
10004	53	1									
10005	54	-1									

По анкетам | По вопросам | По записям | Незаполненные поля

Думаю, что мой отец редко ругается на окружающих

N	ID	Ехр1	Ехр2	Ехр3	Ехр4	Ехр5	С1	С2	С3	С4	С5
▶ 1	▶ 10004	-1	1	1	1	1	2	8	0,8	0,64	0,8
2	10005	1	-1	1	0	1	2	10	1	0,639	0,799
3	10027	-1	1	-1	0	0	2	10	1	0,56	0,748
4	10031	-1	1	1	-1	1	2	12	1,199	0,96	0,979
5	10033	-1	1	1	1	1	2	8	0,8	0,64	0,8

Рис. 12. Вывод несогласованных оценок по анкетам и по вопросам.

Поскольку можно также задавать просмотр всех имеющихся записей в базе данных, то описанная выше возможность просмотра ответов либо по анкетам, либо по вопросам, либо одновременно всех записей базы данных, позволяет также определять процент распространенных ответов по некоторым вопросам, выделять категории ответов, которые в дальнейшем помогут психологу правильно оценить ответ (т. е. составить более целенаправленную инструкцию для оценки предложений).

6. Анализ несогласованных экспертных оценок

До подсчета итоговых показателей, измеряемых тестом незаконченных предложений, был проведен анализ согласованности оценок, выставленных пятью независимыми экспертами. Каждому из 5 экспертов необходимо было оценить 158 анкет, в каждой по 60 предложений (всего 9480 предложений). Результаты анализа согласованности оценок представлены в табл. 1–5.

Таблица 1

Число совпавших оценок

Эксперты	2	3	4	5
1	7004 (73,8%)	6958 (73,39%)	6777 (71,48%)	7067 (74,54%)
2	—	7326 (77,27%)	6695 (70,62%)	7635 (80,53%)
3	—	—	6567 (69,27%)	7462 (78,71%)
4	—	—	—	7153 (75,45%)

Таблица 2

Число оценок, отличающихся на 1 балл

Эксперты	2	3	4	5
1	1997 (20,85%)	1890 (19,93%)	2385 (25,15%)	1985 (20,93%)
2	—	1074 (11,32%)	2319 (24,46%)	1234 (13,01%)
3	—	—	2298 (24,24%)	1215 (12,81%)
4	—	—	—	1998 (21,07%)

Таблица 3

Число оценок, отличающихся на 2 балла

Эксперты	2	3	4	5
1	491 (5,17%)	604 (6,37%)	308 (3,24%)	422 (4,45%)
2	—	1061 (11,19%)	453 (4,77%)	606 (6,39%)
3	—	—	567 (5,98%)	778 (8,2%)
4	—	—	—	319 (3,36%)

Таблица 4

Число оценок, отличающихся на 3 балла

Эксперты	2	3	4	5
1	7 (0,07%)	27 (0,28%)	9 (0,09%)	4 (0,04%)
2	—	17 (0,17%)	13 (0,13%)	3 (0,03%)
3	—	—	47 (0,49%)	21 (0,22%)
4	—	—	—	9 (0,09%)

Таблица 5

Число оценок, отличающихся на 4 балла

Эксперты	2	3	4	5
1	1 (0,01%)	1 (0,01%)	1 (0,01%)	2 (0,02%)
2	—	2 (0,02%)	0 (0%)	2 (0,02%)
3	—	—	1 (0,01%)	4 (0,04%)
4	—	—	—	1 (0,01%)

Как видно из приведенных данных почти $\frac{3}{4}$ оценок полностью совпадают, примерно пятая часть всех оценок имеет расхождение только в один балл. Только около 5% оценок имеют расхождения в 2 или 3 балла. Противоположная оценка обнаружилась только в менее, чем 0,05% случаев. Это говорит о том, что оценки высоко согласованы, и что выданные инструкции четко указывают способ оценки продолжения незаконченного предложения.

Согласованность последовательности пяти экспертных оценок можно определять с помощью метрик близости. Гистограмма частот значений метрики «максимальная разность» представлены на рис. 13.

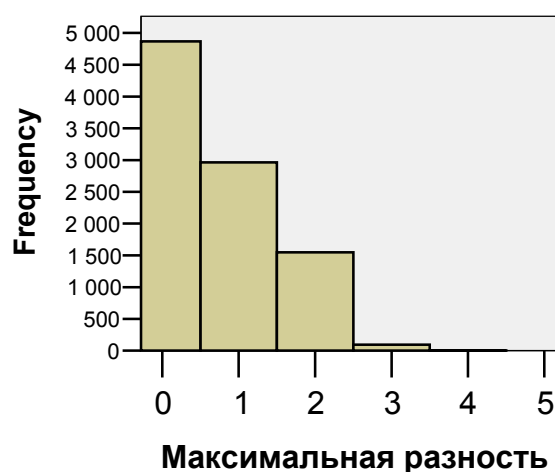


Рис. 13. Гистограмма для метрики согласованности экспертных оценок «максимальная разность».

Для последовательности экспертных оценок также была вычислена минимальная разность, которая показывает, есть ли среди экспертных оценок хотя бы две совпадающие. Оказалось, что для всех предложений этот показатель равен нулю. Это означает, что для всех предложений среди экспертных оценок всегда встречается хотя бы две совпадающие.

По метрике согласованности «максимальная разность» были составлены гистограммы отдельно для каждого вопроса, что позволило выявить «проблемные» предложения, т. е. предложения, в которых значение нуля этого показателя встречается реже, чем другие значения этого показателя. Список предложений, относительно которых несоответствие оценок экспертов случается чаще, чем совпадение, представлен ниже¹:

3. Я всегда хотел(а)...
4. Если бы я занимал(а) руководящий пост...
7. Знаю, что глупо, но боюсь...
9. Когда я был(а) ребенком...
16. Если бы мой отец только захотел...
17. Думаю, что я достаточно способен(на), чтобы...
18. Я мог(ла) бы быть очень счастливым, если бы...
19. Если кто-нибудь работает под моим руководством...
21. В школе мои учителя...
24. Когда-то я...
33. Моим скрытым желанием в жизни...
35. Наступит тот день, когда...
36. Когда ко мне приближается мой начальник...
48. Больше всего я хотел(а) бы в жизни...
52. Мои опасения не раз заставляли меня...
53. Когда меня нет, мои друзья...
54. Моим самым живым воспоминанием детства является...

7. Результаты статистического анализа

Методика «незаконченные предложения» помогла выяснить наличие проблем в системе отношений (табл. 6).

Таблица 6

Гендерные различия в системе личностных отношений

Отношение	Нереализованные возможности	К будущему	К друзьям	К своему прошлому	К лицам противоположного пола	К сотрудникам
Девушки	0.58	2.88	1.46	0.51	-1.65	2.85
Юноши	-0.11	1.72	0.76	-0.23	-1.08	2.21
t	2.62	3.96	3.76	2.46	-1.98	2.26
p	0.0095	0.0001	0.0002	0.015	0.049	0.0254

Выявились различия в системе отношений к нереализованным возможностям, будущему, друзьям, своему прошлому, к лицам противоположного пола и к сотрудникам. Девушки более оптимистично относятся к своему будущему и

¹Нумерация предложений соответствует принятой в методике «Незаконченные предложения».

имеют меньше проблем в отношении с друзьями и сотрудниками. Юноши более негативно, чем девушки, оценивают свои возможности, и менее положительно относятся к своему прошлому. Особо следует отметить, что и юноши, и девушки имеют проблемы в отношении к лицам противоположного пола.

При анализе количества девушек, имеющих проблемы в системе отношений, выявлено, что почти половина девушек (49,2%) негативно оценивает себя, более 90% девушек имеют страхи и опасения и чувство вины. 78% девушек имеют проблемы в отношении к мужчинам. Среди юношей 47,3% имеют проблемы в отношении к себе, 53,8% негативно оценивают свои возможности, 49,5% имеют проблемы в отношении с вышестоящими лицами, 63% негативно оценивают женщин, более 90% юношей имеют страхи и опасения и чувство вины.

Результаты первичного статистического анализа обобщенно представлены на рис. 14а–б.

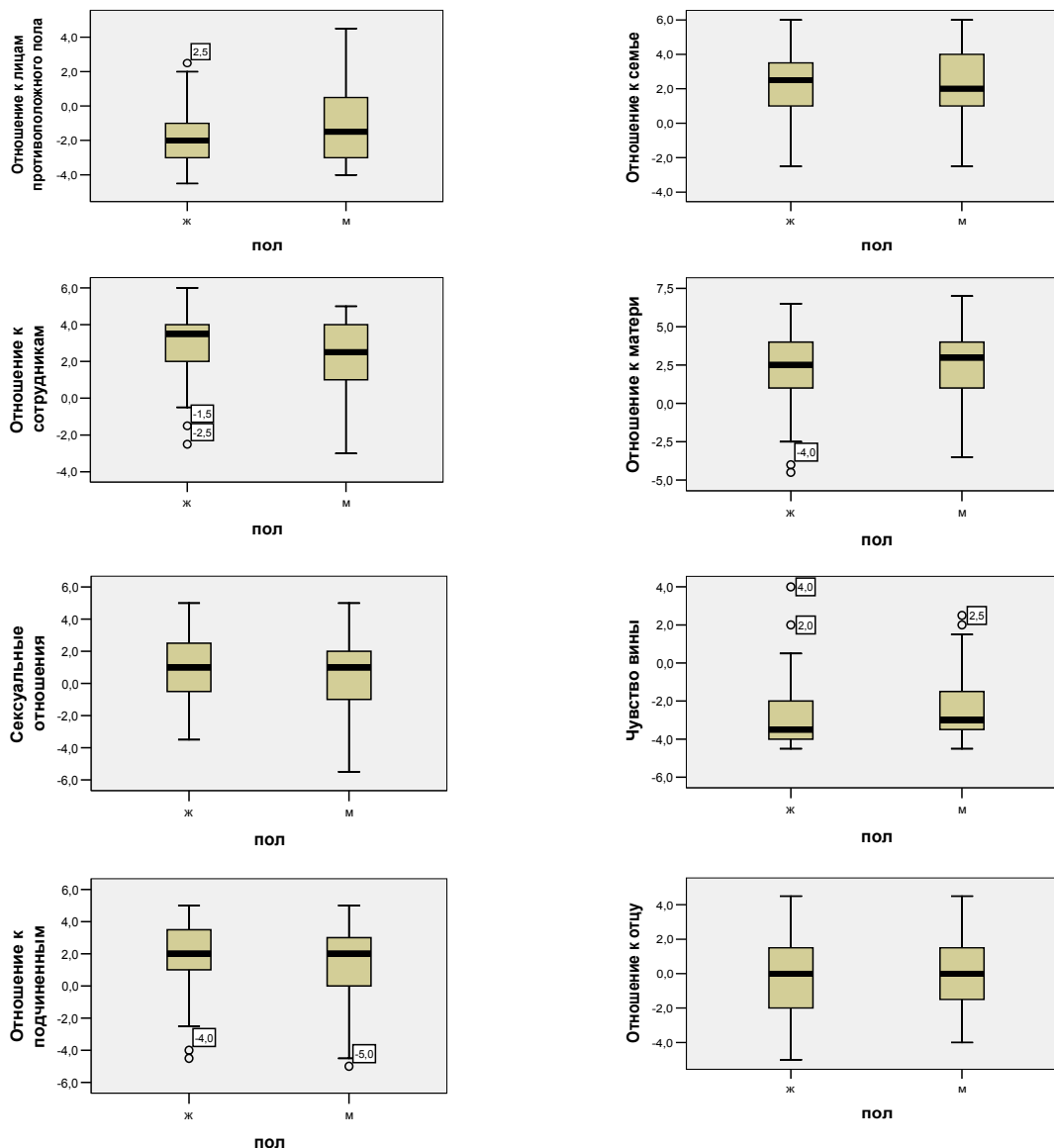


Рис. 14а. Боксплоты по результатам теста «Незаконченные предложения».

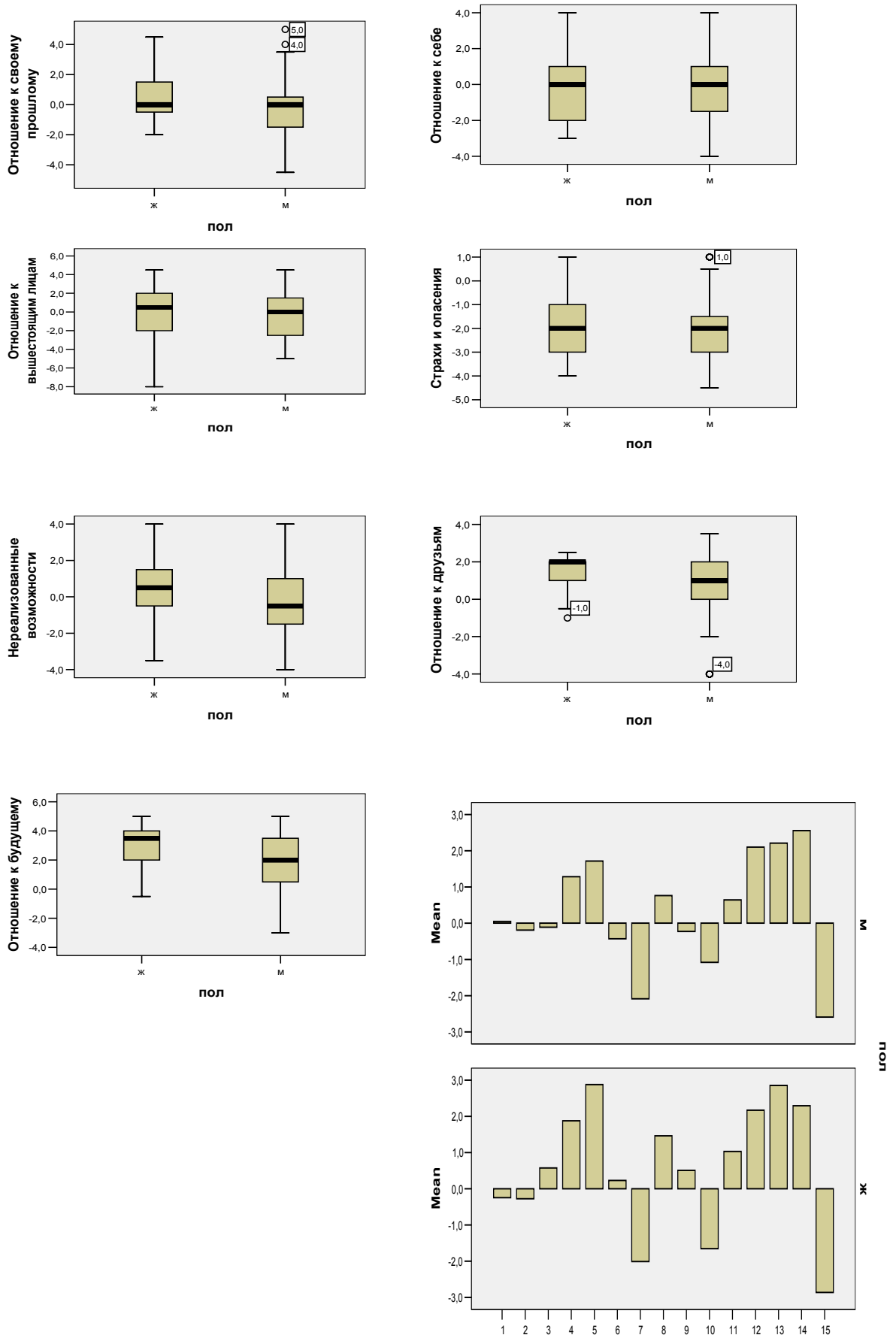


Рис. 146. Боксплоты и гистограмма средних значений по результатам теста «Незаконченные предложения».

8. Заключение

Разработанный комплекс программных приложений позволил не только автоматизировать внесение экспертных оценок и вычислить степени их несогласованности, но также и выявить наиболее проблемные элементы в методике «Незаконченные предложения». Те из незаконченных предложений, относительно которых несоответствие оценок экспертов случается чаще, чем совпадение, потребуют изменений в инструкции, выдаваемой экспертам. Возможно, изменения должны коснуться не только этой инструкции, но также и формулировок начала предложений, составляющих методику.

Отметим, что разработанный комплекс ПО позволит психологам активнее применять проективный тестовый инструментарий — в данном случае методику незаконченных предложений — в консультативной и диагностической практике.

Использование методики «Незаконченные предложения» в процессе диагностики позволит выявить проблемы в отношениях, которые в юношеском возрасте имеют серьезное значение. Наличие трудностей в отношениях с друзьями, родителями, семьей могут создать основания для демонстрации юношей или девушкой девиантного, или даже опасного для общества поведения, могут быть связаны с высоким уровнем рискованного поведения, а также активизацией психологической защиты и неконструктивных копинг-стратегий. Своевременное выявление таких проблем позволит правильно организовать психологическую работу по коррекции системы отношений и, тем самым, уменьшить вероятность демонстрации неконструктивных форм поведения.

Литература

1. Бурлачук Л. Ф. Психодиагностика: учебник для вузов. СПб.: Питер, 2003. 352 с.
2. Соколова Е. Т. Проективные методы исследования личности. М.: МГУ, 1987. 66 с.
3. Практикум по возрастной психологии / Под ред. Головей Л. А., Рыбалко Е. Ф. СПб.: Речь, 2002. 688 с.
4. Калинин С. И. Компьютерная обработка данных для психологов. СПб.: Речь, 2003. 134 с.
5. Наследов А. Д. Математические методы психологического исследования. Анализ и интерпретация данных. СПб.: Речь, 2004. 392 с.
6. Pagano M., Gauvreau K. Principles of Biostatistics. Pacific Grove: Duxbury Press, 2000. 592 p.