

Панина Н.В. Технология социологического исследования. Курс лекций.

СОДЕРЖАНИЕ

<u>ВВЕДЕНИЕ</u>	3
<u>Лекция 1. ПРОЕКТ И ПРОГРАММА ЭМПИРИЧЕСКОГО ИССЛЕДОВАНИЯ</u>	
1. ЭМПИРИЧЕСКОЕ ИССЛЕДОВАНИЕ	
<i>Роль и место в социологическом анализе</i>	11
<i>Возможности и ограничения</i>	12
2. ПРОЕКТ ИССЛЕДОВАНИЯ	
<i>Функции и структура эскиза проекта</i>	14
<i>Уточнение информационных потребностей</i>	16
<i>Программа исследования</i>	19
<i>Определение генеральной совокупности и выборки</i>	19
<i>Инструментарий</i>	20
<i>Сбор информации</i>	21
<i>Обработка и анализ данных</i>	21
<i>Оформление полученных результатов (отчет)</i>	21
<i>Ресурсы и потребности</i>	21
<i>График, смета и общая стоимость исследования</i>	22
<i>Определение приоритетности элементов проекта</i>	22
3. ТИПЫ ЭМПИРИЧЕСКОГО ИССЛЕДОВАНИЯ	
<i>Опрос общественного мнения</i>	23
<i>Социологическое исследование</i>	25
4. ПРОГРАММА СОЦИОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ	
<i>Методологические и методические функции программы</i>	25
<i>Структура программы</i>	27
<i>Технология подготовки программы</i>	29
5. ЗАКЛЮЧЕНИЕ: технологическая схема подготовки проекта и программы	34
<u>Лекция 2. ПОДГОТОВКА АНКЕТЫ</u>	
1. ОСНОВНЫЕ ТРЕБОВАНИЯ К ИНСТРУМЕНТАРИЮ (АНКЕТЕ)	
	37
<i>Валидность</i>	40
<i>Краткость</i>	41
<i>Однозначность</i>	43
2. ОШИБКИ ФОРМУЛИРОВОК	
<i>Неопределенность понятий</i>	43
<i>Неконкретность</i>	45
<i>Неоднозначность</i>	46
<i>Тенденциозность</i>	48
3. ОШИБКИ ФОРМИРОВАНИЯ ВЕЕРА ОТВЕТОВ	
<i>Нарушение рядоположности</i>	49

<i>Неполнота множества</i>	51
<i>Порядок ответов</i>	54
4. ДОПОЛНИТЕЛЬНЫЕ ВОПРОСЫ	
<i>Вопросы-фильтры</i>	56
<i>Открытые вопросы</i>	57
<i>Вопросы социально-демографического блока анкеты</i>	59
5. ОФОРМЛЕНИЕ АНКЕТЫ	61
6. ПРОЦЕСС ПОДГОТОВКИ АНКЕТЫ	64
<i>График работы над анкетой</i>	65
<i>Апробация и доработка инструментария</i>	67
<u>Лекция 3. ВЫБОРКА: ОПРЕДЕЛЕНИЕ ВЫБОРОЧНОЙ СОВОКУПНОСТИ</u>	
1. ПОНЯТИЕ И ПРОБЛЕМЫ ФОРМИРОВАНИЯ ВЫБОРКИ	
<i>Основные проблемы</i>	71
<i>Понятие «выборка»</i>	74
2. КАЧЕСТВО ИНФОРМАЦИИ	75
<i>Репрезентативность</i>	75
<i>Надежность и валидность</i>	77
3. ПЕРЕХОД ОТ ОБЪЕКТА ИССЛЕДОВАНИЯ К ЕДИНИЦАМ НАБЛЮДЕНИЯ (ОПРОСА)	
<i>Технологическая схема</i>	84
<i>Объект исследования</i>	86
<i>Генеральная совокупность</i>	89
<i>Основа выборки</i>	92
<i>Единицы отбора и единицы наблюдения</i>	94
<u>Лекция 4. ВЫБОРКА: ТИПЫ И ПРОЦЕДУРЫ ОТБОРА</u>	
1. ВЕРОЯТНОСТНЫЙ ПОДХОД	
<i>Одноступенчатая случайная выборка</i>	98
<i>Принцип районирования (стратификации)</i>	101
<i>Принцип кластеризации («гнезд»)</i>	102
<i>Многоступенчатая комбинированная выборка</i>	103
2. ЦЕЛЕНАПРАВЛЕННЫЙ (НАПРАВЛЕННЫЙ) ПОДХОД	
<i>Принцип типичности</i>	111
<i>Принцип квот</i>	112
3. ОБЪЕМ ВЫБОРОЧНОЙ СОВОКУПНОСТИ	117
4. ЗАКЛЮЧЕНИЕ: технологическая схема подготовки выборки	
	122
<u>Лекция 5. ОРГАНИЗАЦИЯ СБОРА ДАННЫХ</u>	
1. ДВА СПОСОБА СБОРА ДАННЫХ: ПРЕИМУЩЕСТВА И НЕДОСТАТКИ	
<i>Организации (фирмы) по сбору данных</i>	124
<i>Сбор данных собственными силами</i>	126
2. ТЕХНОЛОГИЯ РАБОТЫ С ОРГАНИЗАЦИЯМИ (ФИРМАМИ) ПО СБОРУ ДАННЫХ	
<i>Выбор фирмы</i>	128
<i>Этапы работы</i>	130
3. ТЕХНОЛОГИЯ САМОСТОЯТЕЛЬНОЙ ОРГАНИЗАЦИИ СБОРА ДАННЫХ	
<i>Структура рабочей группы</i>	137
<i>Этапы работы</i>	142
<i>Подготовка</i>	142

<i>Инструктаж: принципы работы интервьюера</i>	142
<i>Мониторинг и инспекционный надзор</i>	146
<i>Получение анкет</i>	147
<i>Подготовка массива анкет к обработке</i>	148
<u>Лекция 6. ОБРАБОТКА ДАННЫХ</u>	
1. ПОДГОТОВКА К ОБРАБОТКЕ ДАННЫХ	
<i>Планирование обработки</i>	149
<i>Выбор компьютерной программы обработки и анализа данных</i>	150
<i>Планирование методов статистического анализа</i>	151
<i>Мониторинг получения заполненных анкет</i>	159
2. ПРОЦЕДУРА ОБРАБОТКИ АНКЕТ	
<i>Ручная обработка</i>	160
<i>Редактирование заполненных анкет</i>	161
<i>Оценка заполненности анкет</i>	162
<i>Редактирование анкет с вопросами-фильтрами</i>	163
<i>Ввод данных</i>	165
<i>Статистическая обработка данных</i>	167
<u>Лекция 7. АНАЛИЗ И ПРОБЛЕМА ИЗМЕРЕНИЯ</u>	
1. УРОВНИ СОЦИОЛОГИЧЕСКОГО АНАЛИЗА ЭМПИРИЧЕСКИХ ДАННЫХ	
<i>Описание результатов</i>	172
<i>Объяснение полученных фактов</i>	173
2. ОБОЗНАЧЕНИЕ ШКАЛ	
<i>Номинальная шкала</i>	176
<i>Порядковая шкала</i>	180
<i>Метрическая (интервальная) шкала</i>	183
<i>Преобразование шкал</i>	185
3. СТАТИСТИЧЕСКИЙ АНАЛИЗ	
<i>Выбор методов статистического анализа</i>	191
<i>Ряд распределения</i>	193
<u>Лекция 8. АНАЛИЗ ОДНОМЕРНЫХ РАСПРЕДЕЛЕНИЙ</u>	
1. ДИСКРЕТНЫЕ ШКАЛЫ: НОМИНАЛЬНАЯ, ПОРЯДКОВАЯ (РАНГОВАЯ)	
<i>Частотные и процентные распределения</i>	195
<i>Частоты или проценты?</i>	196
<i>По отношению к какому числу следует рассчитывать процент в одномерных распределениях?</i>	198
<i>С какой степенью точности следует приводить данные в процентах?</i>	202
2. НЕПРЕРЫВНЫЕ ШКАЛЫ: ИНТЕРВАЛЬНАЯ, МЕТРИЧЕСКАЯ	
<i>Меры центральной тенденции</i>	204
<i>Меры изменчивости</i>	208
<i>Форма распределения</i>	210
<u>Лекция 9. КРОСС-ТАБУЛЯЦИЯ: АНАЛИЗ ДВУМЕРНЫХ ТАБЛИЦ</u>	
1. РОЛЬ СТАТИСТИКИ В ПРОЦЕССЕ СОЦИОЛОГИЧЕСКОГО АНАЛИЗА	
<i>Возможности и ограниченность</i>	213
<i>Роль исследовательской гипотезы в процессе анализа данных</i>	

	213
2. КРОСС-ТАБУЛЯЦИЯ	
<i>Зависимые и независимые переменные</i>	214
<i>Описание и статистическая экспертиза двумерных таблиц</i>	216
3. СТАТИСТИЧЕСКИЙ АНАЛИЗ СВЯЗИ МЕЖДУ ДВУМЯ ПЕРЕМЕННЫМИ	
<i>Стадии анализа</i>	224
<i>Установление связи между признаками</i>	225
<i>Измерение связи между переменными</i>	228
<i>Объяснение связи между переменными</i>	229
<i>Трехмерный анализ: введение контрольной переменной</i>	230
4. ЗАКЛЮЧЕНИЕ: технологическая схема анализа социологических данных	
<u>Лекция 10. ИТОГОВЫЕ ДОКУМЕНТЫ</u>	237
1. ВИДЫ ИТОГОВЫХ ДОКУМЕНТОВ	239
2. ОТЧЕТ	239
<i>Общие требования к отчету</i>	239
<i>Структура отчета</i>	240
<i>Композиция основной (содержательной) части отчета</i>	242
<i>Оформление таблиц</i>	243
<i>Анализ информации, выводы и рекомендации</i>	245
3. НАУЧНЫЕ ПУБЛИКАЦИИ	249
4. ПУБЛИКАЦИИ В СРЕДСТВАХ МАССОВОЙ ИНФОРМАЦИИ	251
<i>Социолог и СМИ</i>	251
<i>СМИ и социологическая информация</i>	252
<i>Вопросы журналиста к социологу</i>	253
<u>ЗАКЛЮЧЕНИЕ</u>	258
<u>ЛИТЕРАТУРА</u>	259

Навчальний посібник

Паніна Наталія Вікторівна

Технологія соціологічного дослідження

Курс лекцій

Редакційно-видавничий відділ
Інституту соціології НАН України

Відповідальний за випуск: Суїменко В.Х.

Коректор: Кожушко О.О.

Підп. до друку 05.10.97. Формат 60x84 Папір офс.
Ум.-друк.арк. 14.7. Наклад 1500 пр.
252021, Київ-21, вул. Шовковична, 12.

Введение

Заметный всплеск общественного интереса к социологии (характерный для любой общественной системы в периоды политической нестабильности) породил потребность в большом количестве эмпирических исследований и опросов общественного мнения, несоизмеримую с возможностями общества в подготовке необходимого числа профессиональных социологов.

В недавнем прошлом развитие эмпирической социологии определялось в основном ее научным (или идеологическим) значением для общества. В настоящее время, когда резко возросло практическое значение социальной информации и повысилась роль общественного мнения как социального института в политических процессах, эмпирическая социология (и в частности, опросы общественного мнения) в значительной своей части функционирует как производство. Многие социологические центры становятся «фабриками» по производству информации — «товара», спрос на который достаточно высок.

Если работники, занятые в научной сфере, по своей квалификации разделяются в основном на два типа: «ученые-специалисты» (лица, получившие высшее базовое образование в своей профессии, владеющие теоретическими основами знаний в определенной области, деятельность которых направлена преимущественно на поиск «нового знания») и «ассистенты» (помощники, выполняющие обязанности, содержание которых в каждый конкретный момент работы определяет ученый-специалист), то в любом производстве достаточно значимую роль играют специалисты среднего звена — **технологи**. Квалификация технолога предполагает владение совокупностью определенных приемов, позволяющих изготовить качественный **стандартный** продукт. К сожалению, в настоящее время наше общество пока не располагает достаточной базой подготовки специалистов-технологов в области эмпирической социологии, способных удовлетворить современный спрос на актуальную оперативную социологическую информацию: катастрофически не хватает ни учебных центров, ни преподавательских кадров ни учебно-методической литературы. Более того, литературой, обучающей современным практическим навыкам проведения эмпирических исследований, не обеспечивается даже учебный процесс в высших учебных заведениях, готовящих социологов.

Наш «краткий курс» практической социологии не претендует на компенсацию недостатка фундаментальных научных знаний в этой области. Свою задачу мы видим прежде всего в том, чтобы помочь человеку, не имеющему опыта участия в организации масштабных эмпирических исследований и оказавшемуся перед необходимостью самостоятельного выполнения этой работы, сориентироваться в требованиях и правилах, соблюдение которых является обязательным условием получения качественной информации. Особое внимание в данной книге уделяется разбору ошибок, приводящих к грубому искажению получаемой информации, которые наиболее часто допускаются в эмпирических социологических исследованиях.

В такой ситуации могут оказаться вчерашние студенты-социологи, которые прошли полный курс систематического обучения и в какой-то мере научились ориентироваться в социологической литературе. Однако впервые приступая к самостоятельному исследованию, они, как правило, оказываются в растерянности, неожиданно сталкиваясь с многочисленными подводными камнями своей профессии. Эта растерянность часто вызвана тем обстоятельством, что свои академические знания они получали от преподавателей и авторов учебных пособий, которые больше знакомы с теоретическими основами методики и техники прикладных исследований, нежели с современной

практикой подготовки и организации массовых широкомасштабных опросов населения, достаточной для определенных обобщений накопленного опыта.

За последние четверть века в нашей стране образовалась пропасть между развитием теоретических основ методологии эмпирического социологического исследования (как в организации и проведении сбора первичной социологической информации, так и в сфере анализа полученного материала) и практикой эмпирических исследований. Причина разрыва между «высокой теорией» и «презренной практикой» состояла преимущественно в том, что база накопления **конкретных методических приемов**, которые, с одной стороны, должны **отвечать необходимым требованиям**, а с другой — **соответствовать реальным возможностям** их выполнения, была очень ограниченной. Эмпирические социологические исследования, которые проводились в нашей стране с середины шестидесятых годов до конца восьмидесятых отдельными научными коллективами или социологическими лабораториями крупных предприятий, в силу ряда причин не могли позволить накопить достаточный практический опыт.

Во-первых, социологические исследования были, как правило, инициативными, т.е. «социальным заказчиком» выступал автор (или авторский коллектив); соответственно и результаты с точки зрения их значимости и качества «по достоинству» оценивались преимущественно самими авторами.

Во-вторых, в силу лимитированных финансово-организационных возможностей и отсутствия заинтересованности реального заказчика исследования были растянуты во времени — одно исследование научным коллективом планировалось на 3-5 лет, что, естественно, ограничивало возможность апробации всех теоретически разработанных требований и накопления различных модификаций организационных и методических приемов. А поскольку сбор информации осуществлялся силами самих сотрудников или мобилизованных «общественников», то и опыт организации работы собственно интервьюеров (внештатных оплачиваемых сотрудников, функции которых должны принципиально отличаться от функций исследователей) был скорее умозрительным, нежели реальным.

И наконец, исследования проводились в основном на локальных объектах — одно или несколько предприятий, школы, вузы и т.п. Региональные и общенациональные репрезентативные опросы в силу идеологических и финансово-организационных причин были малодоступны для исследовательских коллективов. Именно поэтому, на наш взгляд, наиболее заметный разрыв между строгостью теоретико-методических требований и реальными возможностями выполнения этих требований возник в решении проблемы обоснования и формирования выборки: те принципы и методы, которые предлагаются в большинстве учебных пособий по социологии, чаще всего остаются благими намерениями по дороге в «ад» эмпирической социологии.

С начала 90-х годов ситуация кардинально изменилась. Интерес к общественному мнению со стороны различных социальных и политических структур стимулировал значительный рост широкомасштабных региональных и общенациональных опросов населения. Для получения актуальной оперативной информации о состоянии общественного мнения социологи в своей повседневной практике вынуждены были использовать различные методические приемы, позволяющие интенсифицировать работу и сокращать сроки проведения опросов. Однако в подавляющем большинстве случаев они, к сожалению, не находят сил и времени для оценки влияния используемых модификаций на качество получаемой информации. До настоящего времени новый методический опыт —

обобщенный и статистически обоснованный — не освещается в должной мере ни в научных публикациях, ни в учебных пособиях.

Таким образом, вчерашние студенты, даже добросовестно освоившие вузовский курс социологии, впервые приступая к самостоятельному исследованию, нуждаются в некоторых практических советах, которые, мы надеемся, в дополнение к имеющимся знаниям, помогут в организации и проведении эмпирического социологического исследования.

Второй контингент нашей предполагаемой аудитории мы видим в новообращенных социологах — представителях других профессий, вчера еще знавших о социологии лишь по газетным публикациям, а сегодня, в силу жизненных обстоятельств, имеющих запись в трудовой книжке «социолог n-ой категории» или «ведущий социолог» (в зависимости от того, какая ставка была в распоряжении администрации по новому месту работы, и степени ее доверия к общему уровню развития кандидата на эту должность). И если у студентов в процессе исследования может возникнуть растерянность в связи с тем, что в ряде случаев невероятно трудоемко, а порой практически неосуществимо выполнение многих известных им требований, изложенных в учебниках, то здесь наблюдается совершенно иная картина — большинство новоявленных социологов просто не осведомлены об этих требованиях и со святой простотой идут на многочисленные нарушения процедуры исследования, погрешности и ошибки.

Для профессионально неподготовленного «специалиста» социологическое исследование обычно начинается с «составления анкеты», которая напичкивается более или менее грамотно сформулированными вопросами с веером ответов (чаще всего смещенных по шкалам или нерядоположных) и заканчивается получением «из машины» одномерных и двумерных распределений. Отчет по исследованию у таких авторов чаще всего представляет собой пересказ двумерных таблиц, перемежаемый фантазиями «на тему», без всякой связи с имеющимся материалом. Некоторые затруднения иногда осознаются на той стадии, когда анкета готова и возникает вопрос: «Кого же опрашивать?».

Руководствуясь «здравым смыслом» и соображениями личного удобства, «новые» социологи обычно быстро находят ответ и на этот вопрос. Соображения репрезентативности, правомерности переноса выводов на генеральную совокупность, достоверности полученных данных их мало волнуют просто в силу того, что они не подозревают об этих категориях. Славно потрудившись, проведя одно-два подобных «исследования» (а даже и такие «исследования» достаточно трудоемки), человек искренне уверен, что он «занимался социологией». Здесь вводит в заблуждение кажущаяся простота науки: «спросил — получил ответ — посчитал — узнал». Вопрос «Что же узнал?» — чаще всего не возникает.

Чтобы проиллюстрировать подобный стиль работы и соответствующий уровень социологической квалификации, сегодня, к сожалению, нет необходимости проводить трудоемкую изыскательскую работу в поисках примеров. Достаточно включить телевизор или раскрыть один из номеров какой-либо газеты, приводящей на своих страницах результаты опросов общественного мнения. Чаще всего можно увидеть нагромождение социологических ошибок, для обнаружения которых профессиональному социологу нет необходимости знакомиться с первичным материалом и отчетными документами по организации исследования, так как представленный в публикациях материал сам по себе позволяет не только обнаружить ошибки, но и определить их количественное значение.

Однако далеко не все люди, чья профессиональная деятельность была направлена в новое русло, готовы всенародно обнажать свой дилетантизм. Многие из них охотно овладевают новой профессией и испытывают потребность в информации по технологии процесса социологического исследования. Обращаясь к этой аудитории, мы видим свою задачу прежде всего в том, чтобы обратить ее внимание на разработанные социологической наукой основные требования и правила, позволяющие если не избежать, то учесть возможные ошибки социологического исследования при интерпретации полученного материала.

И наконец, третий контингент, которому, по нашему мнению, могли бы быть полезны систематизированные советы и рекомендации в области практической социологии, — это представители различных областей науки и практики, которые при выполнении своей непосредственной профессиональной деятельности испытывают необходимость в определенной информации о мнениях, настроениях, установках, мотивах и особенностях поведения какой-либо части населения (политики и юристы, менеджеры, работники органов здравоохранения и образования, журналисты и организаторы деятельности средств массовой информации, представители властных структур и правоохранительных органов и многие другие). Специалисты, испытывающие необходимость в информации, получение которой может быть обеспечено преимущественно социологическими методами, чаще всего выступают в роли заказчиков социологических исследований. Однако в силу тех или иных причин нередко предпочитают (или вынуждены) прибегать к самостоятельному социологическому исследованию. Выступая в роли заказчиков или даже исполнителей эмпирического исследования, они, как правило, недостаточно ясно представляют себе все возможности и ограничения социологических методов. Неоднократно приходилось убеждаться в том, что по отношению к эмпирической социологии со стороны представителей других профессиональных категорий проявляется либо тотальный скептицизм («метод опроса никогда не позволяет узнать истину»), либо простодушная наивность, что с помощью одного исследования можно получить исчерпывающие ответы на все возникшие вопросы. Свою задачу в данном случае мы видим в раскрытии общих принципов и правил, позволяющих оценить возможности и пределы применения методов социологического исследования.

Представляемый читателю общий курс по технологии эмпирического социологического исследования включает рассмотрение основных вопросов по следующим темам: проект и программа исследования; особенности подготовки инструментария — конструирование анкеты и техника шкалирования; определение объекта исследования и подготовка выборки; сбор данных и особенности работы с интервьюерами; обработка, анализ и интерпретация данных, подготовка итоговых документов.

В рамках эмпирической социологии накоплен достаточно широкий арсенал методов сбора первичной информации, многие из которых совершенствуются и модернизируются. Наряду с методами массовых опросов развиваются такие перспективные направления, как анализ документов и экспертные методики; социологи не забывают традиционные методы эксперимента и наблюдения; в эмпирической социологии входят в моду интроспекция и фокус-группы. Однако содержание данного учебно-методического пособия полностью сосредоточено на технологии проведения массового опроса населения методом анкетирования. Этот выбор продиктован тем обстоятельством, что именно данная процедура используется в массовом социологическом «производстве» не только в нашей стране, но и в мировой практике опросов общественного мнения. В этом смысле данную книгу точнее было бы назвать «Технология типичного (или стандартного) социологического исследования».

Следует также отметить, что на страницах данного учебника появилась новая для нашей учебно-методической литературы фигура «заказчик». Социологам, имеющим за плечами даже небольшой опыт организации и проведения массовых опросов, не требуется объяснять роль и значение в «производстве социологической информации» заказчика и, соответственно, заинтересованного спонсора. Особенности взаимодействия с заказчиком играют существенную роль, и в первую очередь в процессе крупномасштабного исследования. Во-первых, его интересы целиком сосредоточены на конечном продукте работы; во-вторых, он оплачивает возможность ее осуществления; в-третьих, заказчик, как правило, не является специалистом в области социологии и, следовательно, не имеет достаточно адекватного представления обо всех возможностях и ограничениях эмпирической социологии. Несомненно, учет в технологии проведения эмпирического исследования особенностей взаимоотношения с заказчиком на разных этапах работы играет значительную роль в повышении общей эффективности и качества результатов исследования.

В самом общем виде социологические опросы населения можно разделить на инициативные (выбор проблемы и цели исследования определяется автором) и заказные (выбор проблемы определяется заказчиком, оплачивающим исследование с целью получения определенного рода информации). Однако мы думаем, что инициативные исследования можно рассматривать как разновидность заказных. И хотя в данном случае заказчик и исполнитель выступают в одном лице, исследователь в процессе подготовки должен осознавать эти две роли и рассматривать предлагаемые особенности взаимоотношений между заказчиком и исполнителем как внутренний диалог в процессе исследования. В основном это относится к начальному и заключительному этапам работы.

Выражаю искреннюю признательность моим коллегам Е.И.Головахе и Н.Н.Чурилову, внимательно прочитавшим книгу и своими замечаниями и рекомендациями способствовавшим ее структурной и содержательной доработке.

Лекция 1. Проект и программа эмпирического исследования

1. Эмпирическое исследование

Роль и место в социологическом анализе

Проведение анализа какой-либо социальной проблемы нередко порождает у исследователя потребность в привлечении тех или иных фактов, позволяющих подтвердить (или опровергнуть) гипотетически предполагаемые им закономерности, тенденции развития, механизмы детерминации изучаемого явления. Социальные факты могут быть зафиксированы в документах, в различных явлениях социальной жизни, в работах научных предшественников и т.п. Однако, когда данных недостаточно (они носят скрытый характер или автор не может их получить из доступных ему источников информации), возникает необходимость в проведении специального конкретного

эмпирического исследования для получения необходимой информации. В таком случае эмпирическое исследование является составной частью социологического анализа.

Однако эмпирическое исследование обладает рядом особенностей, специфика которых обусловлена требованием эмпирической верификации довольно сложных, а зачастую и достаточно абстрактных социальных явлений. Приступая к исследованию, социолог, который до определенного момента лишь теоретически анализировал социальную действительность, сразу же сталкивается с рядом проблем.

Эти проблемы можно условно разделить на два типа: 1) проблемы, связанные с неподготовленностью автора к эмпирическому исследованию; 2) проблемы, связанные с ограниченностью эмпирической эксплицированности изучаемых процессов и явлений, — объективные проблемы.

Возможности и ограничения

Что касается проблем первого типа, то они чаще всего порождены заблуждением, что эмпирическое исследование поможет разобраться в тех сложных социальных процессах и явлениях, которые к началу эмпирического исследования представляли лишь социальную, а не научную проблему. Значительная часть трудностей, которые возникают в процессе эмпирического исследования, начинающегося без предварительного теоретического анализа предмета, порождается тем обстоятельством, что **социолог хочет получить ответы на непоставленные вопросы**. Когда автор, не проведя предварительного теоретического анализа, сразу же приступает к эмпирическому сбору информации, в глубине души он надеется, что опрос респондентов поможет разобраться в механизмах самого явления. Поступая таким образом (чаще всего даже не осознавая этого), исследователь фактически пытается свою задачу атрибутировать населению. В результате, получив материал, он готов не к анализу результатов эмпирического исследования, а лишь к представлению полученных данных.

При отсутствии предварительного теоретического анализа объекта исследования социолог сталкивается со многими конкретными проблемами практически на всех этапах эмпирического исследования. Избежать их довольно просто — необходимо лишь твердо усвоить, что эмпирическое исследование является не более чем сбором сырого материала, что само по себе (без целенаправленного структурирования и осмысления) крайне редко приносит новое знание и позволяет разобраться в сути явления. Эмпирическое исследование в основном предназначено для подтверждения (или опровержения) гипотез исследователя; и прежде чем приступить к планированию эмпирического исследования, социолог должен сформулировать гипотезы, которые он хочет проверить. Чем конкретнее сформулированы гипотезы, тем эффективнее общий результат эмпирического исследования.

Основная объективная проблема эмпирического социологического исследования заключается, как правило, в том, что изучаемые социальные явления и процессы определяются понятиями достаточно сложными и абстрактными, сущность которых не может быть непосредственно зафиксирована и измерена одним вопросом анкеты (например, «политическая культура», «межнациональные отношения», «социальная напряженность», «социальная справедливость», «социально-психологическая адаптация», «национальная идентификация» и т.д.). В то же время у исследователя нередко возникает необходимость подтверждения определенных умозрительных конструктов и выдвинутых

гипотез эмпирическими фактами; социально-политическая практика также предъявляет свои запросы к получению эмпирической информации об особенностях социальной ситуации, тенденциях, закономерностях.

Выход из этой проблемной ситуации эмпирическая социология находит в поиске более конкретных характеристик, которые могут рассматриваться как индикаторы абстрактного понятия, описывающие некоторую его часть, эксплицированную в социальных фактах.

Американский социолог П.Лазарсфельд, внесший, пожалуй, самый весомый вклад в развитие измерительных возможностей эмпирической социологии, отмечал, что явление, которое не может быть непосредственно наблюдаемо, все же оставляет следы, которые, если их правильно интерпретировать, сделают возможным его идентификацию и анализ.

Представление сущности изучаемого явления путем всестороннего его описания более конкретными понятиями, поддающимися эмпирической регистрации, позволяет более или менее достоверно и однозначно зафиксировать его в эмпирическом исследовании, направленном на сбор информации о социальных фактах.

Проблема, стоящая перед исследователем, приступающим к эмпирической проверке гипотез, касающихся особенностей изучаемого явления, заключается в том, что, с одной стороны, чем больше характеристик, эксплицирующих сущность предмета исследования, введено в эмпирическое исследование, тем меньше вероятность ошибки измерения. С другой стороны, при подготовке инструментария исследователь вынужден ограничивать набор характеристик, эмпирически представляющих изучаемое явление, так как объем инструментария лимитируется организационными возможностями эмпирического исследования и психологическими ресурсами респондентов.

Прежде чем планировать эмпирическое исследование, социолог должен ясно осознавать, что, с одной стороны, это достаточно трудоемкая работа с колоссальной затратой организационных, человеческих и материальных ресурсов, а с другой — полученная информация носит весьма ограниченный характер.

Чтобы в полной мере оценить и сопоставить потребность в желаемой информации с затратами на эмпирическое исследование, необходима предварительная работа, позволяющая произвести такое сопоставление.

Предварительный этап работы, позволяющий определить потребность в эмпирическом исследовании, возможности получения желаемой информации о тех или иных социальных фактах, а также затраты профессиональных, финансовых, материально-технических и организационных ресурсов, называется подготовкой проекта исследования.

2. Проект исследования

Функции и структура эскиза проекта

Большинство эмпирических исследований в настоящее время характеризуются двумя особенностями: фактором «неожиданности» и фактором «заказчика».

«Неожиданность» определяется как тем, что к профессиональному социологу могут обратиться с предложением (потенциальным заказом) о проведении исследования по

проблеме, которой он раньше не занимался, так и тем, что события, происходящие в обществе, сами по себе могут привлечь внимание социолога к проблемам, над которыми он вчера еще не задумывался, а сегодня считает слишком значительными, чтобы их можно было оставить вне поля зрения в общем анализе изучаемых им социальных явлений (изменение политической структуры, поворот в экономической или социальной политике государства, вспышки межнациональных конфликтов, социальная напряженность и массовый протест и т.п.)

Фактор «заказчика» определяется особенностями взаимодействия с людьми, которые, сами не являясь исследователями, редко бывают осведомлены как о возможностях, так и о границах возможностей эмпирического социального исследования. От правильной стратегии и тактики поведения социолога-исполнителя на начальном этапе сотрудничества зависит основной конечный успех работы.

В такого рода социологических исследованиях (заказных или актуально-оперативных) наиболее важным этапом работы также является подготовка проекта исследования.

Проект исследования позволяет структурировать и провести предварительное оценивание основных элементов и этапов работы, определить стратегию и тактику эмпирического исследования.

Первым этапом в разработке проекта исследования является работа с эскизом проекта.

Эскиз проекта исследования — это перечень пунктов (вопросов), которые необходимо уточнить до начала работы. Именно на основании работы с эскизом готовится проект и, в конечном итоге, принимается решение о том, целесообразно ли проводить данное эмпирическое исследование.

Составление эскиза позволяет:

- 1) уточнить предмет исследования;
- 2) определить реальную стоимость затрат и ресурсов;
- 3) согласовать вопросы, которые нередко становятся предметом спора или конфликта по окончании исследования.

Работа с эскизом исследования позволяет избежать «эффекта отсутствия результата».

Эскиз проекта исследования включает следующие основные пункты.

1. Информационные потребности.
2. Программа исследования.
3. Генеральная совокупность и выборка.
4. Инструментарий (анкета, опросный лист).
5. Сбор информации.
6. Обработка и анализ данных.

7. Отчет (оформление результатов исследования).

8. Потребности и ресурсы.

9. График и смета.

10. Общая стоимость исследования.

Последовательность уточнения пунктов эскиза условна, так как уточнение одних элементов приводит к изменению других.

Работа с эскизом позволяет, во-первых, сформировать достаточно конкретное представление о том, какого рода информация может быть получена в результате работы, а во-вторых, соотнести свои профессиональные, организационные и финансовые возможности с информационным запросом и достаточно основательно подготовиться к качественному выполнению работы.

В процессе работы над эскизом не все пункты разрабатываются с одинаковой степенью конкретности. На первый план здесь выступают те элементы проекта, уточнение которых позволяет определить трудоемкость и стоимость эмпирического исследования. Содержательные же элементы, направленные на уточнение сути социальных фактов, подлежащих анализу, разрабатываются на следующих этапах работы, после окончательного решения о необходимости проведения эмпирического исследования. В работе над эскизом конкретизация содержательных элементов доводится лишь до степени приблизительного определения затрат.

Ниже в общих чертах представлено основное содержание работы над каждым из пунктов эскиза проекта.

Уточнение информационных потребностей

Уточнение информационных потребностей — ключевой этап в проектировании эмпирического исследования. Эта наиболее трудная и творческая часть исследования, к сожалению, чаще всего недооценивается в начале работы и является источником многих недоразумений на заключительном этапе. Ошибки, допущенные на этой стадии, умножаются на всех последующих.

В академической социологии этот элемент называется «определение проблемы, предмета и цели исследования» и в своем разрешении должен дать ответ на вопросы: 1) **что**, собственно, изучается? 2) **для чего** проводится эмпирическое исследование?

Нередко даже профессиональные социологи формально относятся к этому элементу и только на этапе анализа данных начинают задумываться о том, что же, по существу, они могут извлечь из собранного «по теме» материала. Это происходит потому, что возникшую в обществе актуальную проблему так или иначе ощущает практически каждый гражданин. Однако это ощущение носит скорее житейский характер, не имеющий целевой конкретно-научной постановки проблемы. Восприятие проблемы выглядит примерно так: «Переход к рыночной экономике — это проблема», «Преступность среди молодежи — это проблема», «Обострение межнациональных отношений — это проблема».

Однако формулировка, определяющая то, в чем, собственно, заключается проблема, — представляет собой сложную творческую задачу, решение которой составляет половину общего объема интеллектуальной работы исследователя. Следует заметить, что далеко не всем посильно разрешение этой задачи. Однако в проведении многих эмпирических исследований вполне допустимо ее упрощение.

Упрощение такого элемента социологического исследования, как постановка проблемы, сводится к формулировке конкретных вопросов, ответ на которые должен быть получен в результате эмпирического исследования. Таким образом, **информационный запрос** представляет собой вопросы, написанные в произвольной форме, требующие ответа в результате эмпирического социологического исследования.

В процессе формулировки вопросов особое внимание должно быть уделено их **конкретизации**. Например, на первый взгляд, достаточно конкретным представляется следующий информационный запрос: «Получить в результате исследования конкретные рекомендации, как должен поступать руководитель в случае конфликта с подчиненным». Но для социолога подобный информационный запрос не является конкретным. Здесь необходимо уточнить, какого рода конфликты интересуют заказчика (деловые, межличностные, этические и т.д.) — уточнение сфер, предмета или причин конфликтов. Все ли сферы конфликтов должно охватывать исследование или достаточно ограничиться некоторыми из них. Также необходимо уточнить, какой результат разрешения конфликтов волнует заказчика в первую очередь. Информационный запрос переформулируется таким образом: «Как должен поступать руководитель в случае конфликтов с подчиненными, чтобы добиться... а) уважения подчиненных; б) улучшения производственных показателей; в) уменьшения текучести кадров; г) улучшения социально-психологического климата в коллективе и т.п.

Обычно на просьбу социолога уточнить, какие из вопросов интересуют прежде всего, следует первая реакция заказчика: интересуется все! Однако расширение круга вопросов повышает стоимость исследования (за счет увеличения объема анкеты, времени обработки и анализа материала) и снижает общее качество информации, так как слишком большой опросный лист, требующий длительного заполнения, притупляет внимание респондента. Поэтому желательно установить приоритетность различных аспектов проблемы.

Для социолога, планирующего авторское исследование, такой перечень вопросов (даже при наличии сформулированной темы и проблемы исследования) является контрольным регистром на всех этапах работы: контроль конкретности и осознанности проблемы, определение генеральной совокупности, контроль адекватности подготовленного инструментария поставленным задачам (на все ли поставленные вопросы позволит ответить составленная анкета), выбор методов статистического анализа собранного материала, содержание отчета.

Этот перечень вопросов (информационный запрос) принимается за основу уточнения всех остальных пунктов эскиза исследования.

Программа исследования

На этапе подготовки эскиза проекта достаточно определить, кто из специалистов будет привлечен к разработке программы и установить срок, необходимый для этой работы.

Определение генеральной совокупности и выборки

Для определения стоимости и организационных затрат исследования необходимо определить приблизительный объем генеральной совокупности: уточнить, чьи конкретно мнения, настроения, отношения, установки или социальные ориентации будут изучаться. Такое уточнение рационально проводить по следующей схеме: регион — социально-демографическая характеристика — отраслевая специфика.

Например, необходимо изучить отношение молодежи к преступности. Для определения генеральной совокупности сначала уточняется региональная локализация населения, на которое можно будет распространять выводы, полученные в результате исследования: идет ли речь о молодежи страны в целом или какого-либо региона. Затем уточняется, будет ли изучаться лишь городская или сельская молодежь, и если, допустим, городская, то только ли молодежь крупного города или вся городская молодежь. Затем (в рамках социально-демографических уточнений) определяется верхняя и нижняя границы молодежного возраста (например, от 18 до 30 лет или от 14 до 25 лет — конкретные границы зависят от целей и информационных потребностей); и, наконец, уточняется отраслевой контингент — будет ли изучаться мнение всей молодежи, проживающей в данном регионе, или только учащихся, или только работающей молодежи (если работающей, то во всех отраслях или в каких-то определенных) и т.д. Итогом этого обсуждения должно явиться четкое определение генеральной совокупности, например, «учащаяся молодежь в возрасте 18-30 лет западного региона Украины (Львовская, Ивано-Франковская и Тернопольская области)». На основании определения генеральной совокупности (контингента населения, на который можно с определенной вероятностью распространять выводы по полученным в результате исследования данным) исследователь может предварительно определить объем выборки — количество человек, которых необходимо опросить.

На этапе предварительного прикидывания объема выборочной совокупности необходимо учитывать также и аналитические потребности исследования: будет ли последующий анализ ограничиваться представлением о мнении населения в целом или же предполагается сравнение мнений (настроений, установок и т.п.) различных социально-демографических групп. В тех случаях, когда планируется представление полученных данных по различным группам населения, объем выборки должен быть соответствующим образом увеличен. Таким образом, на этом этапе исследователю необходимо уточнить, какие именно социально-демографические группы представляют для заказчика особый интерес. Предварительное определение объема выборки необходимо и для уточнения стоимости исследования.

Своевременное обсуждение генеральной совокупности следует провести также для того, чтобы заранее продумать источники, содержащие основные характеристики генеральной совокупности (общая численность, распределения по основным характеристикам) и пути их получения.

Для ориентировочного определения объема выборочной совокупности исследователь уточняет величину возможной допустимой ошибки, объясняя заказчику, что чем более

точный результат ему необходим, тем большее число людей следует опрашивать. На этом этапе объем выборочной совокупности определяется ориентировочно.

Инструментарий

На этапе работы над эскизом проекта в графе «Инструментарий» достаточно определить примерный объем опросника/анкеты (на основе уточненных информационных потребностей), состав исполнителей и срок работы. Кроме этого целесообразно уточнить технические возможности и базу тиражирования, а также материальные затраты на разработку, подготовку и тиражирование инструментария и методических документов к нему.

Сбор информации

На этапе работы над эскизом проекта решается вопрос о том, какая организация будет осуществлять сбор данных. В том случае, когда сбор данных планируется осуществлять собственными силами, на основе определения выборочной совокупности намечаются «точки опроса», состав и структура группы сбора информации, сроки работы. Подсчитывается примерная стоимость работы по сбору первичной социологической информации.

Обработка и анализ данных

Определяется база обработки первичной социологической информации, технические и программные средства, исходя из которых устанавливаются сроки и стоимость этого этапа работы.

Оформление полученных результатов (отчет)

Следует заранее продумать и обсудить с потенциальными соисполнителями (или заказчиком) форму представления результатов работы (научные статьи, монография, аналитический отчет, аналитические таблицы, графики и т.п.). Если исследование выполняется по заказу, то желательно оговорить объем (если это отчет) или количество таблиц и объем текста комментариев к таблицам.

Ресурсы и потребности

Оценка ресурсов и возможностей проведения исследования включает: определение общего количественного и качественного состава группы, которая будет выполнять исследование, материально-технические потребности и возможности, организационные особенности выполнения каждого из этапов работы (необходимость и пути доступа к ведомственной документации; организации, которые было бы желательно привлечь для

помощи в подготовке к процедуре сбора информации, обеспечении техникой и др). Организационные потребности и возможности необходимо продумать заблаговременно, что позволяет предусмотреть затраты, а также сэкономить время, так как организационную подготовку в таком случае можно осуществлять параллельно с разработкой программы и инструментария.

График, смета и общая стоимость исследования

Общая стоимость исследования — это первоочередная проблема при определении возможности проведения эмпирического исследования. Однако сложность ее разрешения состоит в том, что, не уточнив информационные потребности и, соответственно, объем работы, трудно определить реальную стоимость исследования. Здесь интуиция и накопленный опыт помогут исследователю определить «порядок» общей стоимости.

Окончательная же сумма устанавливается в результате работы с эскизом исследования. Нередко лимит средств, которым связан заказчик (или социолог), приводит не к отказу от исследования в целом, а к корректировке программы (исключение ряда информационных запросов) или к ограничению выборки, или к сужению генеральной совокупности.

Завершение работы над эскизом проекта заключается в составлении таблицы, в которой все элементы эскиза расписаны по срокам исполнения (график) и стоимости (смета). Рабочий документ, представляющий собой перечень основных элементов социологического исследования с указанием стоимости и сроков исполнения, называется проектом социологического исследования.

На основе проекта принимается окончательное решение о том, будет проводиться данное исследование или нет.

Определение приоритетности элементов проекта

После принятия решения о проведении эмпирического исследования начинается второй этап работы над проектом исследования — параллельная подготовка к проведению сбора информации, разработка программы и инструментария, подготовка выборки.

Основными элементами проекта исследования, определяющими его научную значимость и финансовую стоимость, являются программа и выборка. Их соотносительная значимость различна для разных типов исследований.

Если абстрактным идеалом получения социальных фактов эмпирическим путем является сплошной опрос всех представителей исследуемой группы населения по всем возможным характеристикам изучаемой проблемы, то реально в силу вполне понятной лимитированности средств, времени и организационных возможностей в конкретном исследовании необходимо ограничиваться определенным объемом выборки и отдельными аспектами изучаемой проблемы. Чтобы эти ограничения носили достаточно осознанный характер и позволяли оценить значимость и качество полученной информации, приступая к работе, необходимо определить приоритетность элементов исследования.

Первый шаг в оценке приоритетности элементов — определение типа исследования.

Крупномасштабные социологические исследования можно разделить на два типа: **опросы общественного мнения** и собственно **социологические исследования**. Такая классификация носит принципиальный характер для определения приоритетности и глубины разработки основных элементов проекта исследования — программы и выборки.

3. Типы эмпирического исследования

Опрос общественного мнения

Опрос общественного мнения — это сбор информации о совокупном мнении (распределении мнений) определенного контингента населения по какому-либо актуальному вопросу.

При проведении такого рода исследований самым важным элементом проекта является проблема **репрезентативности выборки**; и, соответственно, в организационном плане — установление объема и способа отбора единиц наблюдения, так как результаты опроса ограниченного числа людей при интерпретации результатов переносятся на всю генеральную совокупность. Например, опросив тысячу жителей Киева, исследователь делает выводы о мнениях и настроениях киевлян в целом; опросив две тысячи жителей Украины, распространяет выводы на население Украины в целом и т.д. Поэтому, определяя выборку, исследователь должен достаточно ясно представлять, в какой степени и по каким характеристикам (пол, возраст, образование, социальное положение и т.д.) опрошенные представляют тот контингент, на который распространяются выводы, полученные в результате проведенного опроса.

При опросах общественного мнения значение программы социологического исследования сводится к минимуму, и часто в программе вообще нет необходимости. Например, программа не нужна при опросах в электоральных ситуациях («За кого будете голосовать?», «Будете ли принимать участие в референдуме?» и т.п.) и при оперативных исследованиях острой социальной ситуации (например, реакция населения на те или иные события в стране или на решения органов власти т.п.). В опросах общественного мнения необходима лишь редакция информационного запроса в соответствии с основными техническими требованиями и правилами составления анкеты.

Полученные в результате опроса общественного мнения данные представляются преимущественно в виде одномерных распределений на поставленные вопросы или в виде распределения мнений различных социальных контингентов населения.

Обязательными требованиями при представлении результатов опроса общественного мнения (отчетных материалах, публикациях в средствах массовой информации и т.п.) являются:

- 1) предъявление полной формулировки вопроса в том виде, в котором он предлагался в анкете;

- 2) определение генеральной совокупности и указание объема выборочной совокупности;
- 3) указание ошибки репрезентативности.

Социологическое исследование

Собственно **социологическое исследование** — это изучение сути предмета исследования и его взаимосвязей с другими элементами социального мира.

Из определения ясно, что для перехода от постановки проблемы исследования в самом общем виде («Социально-психологические проблемы Чернобыльской катастрофы», «Обострение межнациональных отношений», «Отношение населения к рыночной экономике», «Социально-политическая ситуация в Крыму» и т.п.) — проблемы, имеющей много аспектов и причинно-следственных связей, — к разработке инструментария необходима определенная схема конкретизации предмета и целей исследования.

Эту функцию (конкретизация предмета исследования) выполняет программа исследования. В проведении собственно социологического исследования подготовка программы исследования является самым важным этапом работы.

4. Программа социологического исследования

Методологические и методические функции программы

Программа социологического исследования — это научный документ, отражающий логически обоснованную схему перехода от теоретического (концептуального) осмысления изучаемой проблемы к инструментарию конкретного эмпирического исследования. Суть программы обусловлена двумя основными взаимосвязанными методологическими функциями. Первую методологическую функцию программа выполняет на этапе подготовки анкеты — методологическое обоснование инструментария исследования; вторую — на этапе анализа полученных данных — обоснование логики анализа и интерпретации полученных в результате сбора первичной социологической информации по разработанному инструментарию данных.

К сожалению, во многих случаях даже профессиональные социологи к программе относятся не как к методологическому инструменту, а как к формальному документу: программа с изложением цели и задач исследования (более или менее семантически связанных с названием темы) подготавливается как официально необходимый документ, а затем (после прохождения этапа утверждения темы или подписания договора) откладывается в сторону, и к ней больше не обращаются. Расплатой за формальное отношение к программе в начале исследования является чувство растерянности и неопределенности на том этапе работы, когда необходимо приступить к анализу полученного эмпирического материала. В этом случае исследователь (часто того не осознавая) вынужден вернуться к составлению программы, чтобы получить более или менее приемлемую схему анализа. Однако последствия попытки подменить программу подобной схемой, как правило, оказываются несостоятельными: одних признаков явно не

хватает, другие, хотя и кажутся, на первый взгляд, интересными, не находят своего места в общем анализе проблемы.

Исследование без предварительной программы (или с программой, подготовленной формально) выглядит обычно следующим образом. Руководитель исследования собирает группу исполнителей и объявляет, что будет исследоваться проблема «межнациональных отношений (в таком-то регионе)» или «социально-психологических последствий аварии на Чернобыльской АЭС», или «отношение населения к рыночной экономике» (приводятся примеры самых распространенных в последнее время заказов на социологическое исследование). Сразу же начинается подготовка анкеты. Можно составить сотни вопросов, так или иначе отражающих социально-психологические последствия аварии на АЭС, и добавить в анкету еще сотни вопросов, определяющих гипотетические факторы, влияющие на эти показатели (процесс составления вопросов — довольно увлекательная часть работы, особенно если она не регламентируется предварительной программой). Значимость тех или иных вопросов чаще обуславливается темпераментом и напористостью автора вопроса, нежели целевым информационным запросом. Если вопросов получилось «слишком много», часть из них отсеивается усилиями «коллективного разума» в соответствии с теми же критериями — авторитет и темперамент отдельных исполнителей.

В результате такого «мозгового штурма» появляется инструментарий исследования — анкета, содержащая несколько десятков или сотен вопросов. Хорошая получилась анкета или плохая? Квалифицированный социолог никогда не возьмется давать оценку анкете (имеется в виду ее содержательная, а не техническая сторона), не сопоставляя ее с программой исследования, так как качество анкеты определяется ее адекватностью конкретному информационному запросу и соответствием индикаторов (вопросов анкеты) выделенным показателям изучаемого явления. Нужен или не нужен конкретный пункт-вопрос анкеты, измеряет или не измеряет он какую-либо характеристику предмета исследования — можно оценить, имея критерии оценивания. Эти критерии и должна содержать программа исследования.

Структура программы

Что же представляет собой программа исследования с точки зрения ее структуры и технологии подготовки? Описание программы социологического исследования содержится практически в каждом пособии по методике и технике социологического исследования.

Здесь мы остановимся на тех элементах программы, без определения которых, по нашему мнению, не рационально приступать к социологическому исследованию в целом и к подготовке инструментария в частности.

Структура такой «программы-минимум» содержит четыре основных взаимосвязанных элемента:

- 1) предмет исследования;
- 2) показатели;
- 3) цель исследования;

4) гипотезы.

1. Предмет исследования определяется конкретной формулировкой ответа на вопрос: «Что?» («Что будет изучаться?»); например, «социальная напряженность» или «политическая культура» и т.п.

Кроме наименования предмета исследования необходимо дать его определение в той формулировке, которую автор принимает за основную. Без определения предмета исследования, даже после проведения эмпирического этапа, нельзя сказать, что была изучена проблема, к примеру, политической культуры, так как непонятно, что автор подразумевал под «политической культурой», если в программе не было дано определение этому понятию.

2. Показатели определяются ответом на вопрос: «Чем?» («Чем будет измеряться предмет исследования?»).

Как правило, изучаемые социальные явления или процессы достаточно сложны и абстрактны, в силу чего не поддаются полностью эмпирической интерпретации. Определение переменных, которые позволяют измерять сложные явления, так как поддаются эмпирической интерпретации, представляет собой процедуру определения показателей изучаемого явления.

3. Цель эмпирического исследования определяется ответом на вопрос «Для чего?» («Для чего проводится эмпирическое исследование?»). Если ответ на этот вопрос у исследователя начинается со слов: «Чтобы узнать...», то это означает, что цель еще недостаточно конкретизирована. Дальнейшая конкретизация может быть продолжена ответом на следующий вопрос: «Для чего узнать?»

Конкретизация цели исследования необходима не только (и даже не столько) для прагматического обоснования затрат на проведение эмпирического исследования, сколько для методологического обоснования критериев «целенаправленного» отбора или отсева показателей и индикаторов, а также гипотетических факторов. Можно предложить достаточно большое количество показателей изучаемого явления, они будут убедительно обоснованы с точки зрения раскрытия сущности предмета, но нельзя забывать об эмпирических лимитах (организационные возможности исследования и психологические возможности респондентов, финансовые и временные ограничения и т.п.) Обоснованно сузить круг возможных показателей позволяет четко сформулированная цель исследования.

4. Гипотезы — обоснованные предположения о причинах и факторах, определяющих характер изучаемого явления, требующие эмпирического подтверждения или опровержения.

Технология подготовки программы

Рабочий вариант программы удобно оформлять в виде схемы, включающей показатели изучаемого явления (с обозначением их структурной взаимосвязи) и показатели гипотетических факторов.

В общем виде схема программы представлена на рисунке 1.

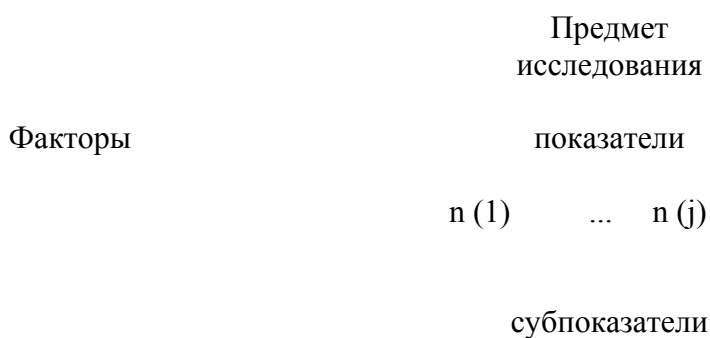


Рис. 1. Схема программы исследования

Эмпирическая интерпретация теоретического показателя может иметь несколько ступеней (субпоказателей), последняя из которых должна быть доведена до такой степени конкретности, чтобы каждый из субпоказателей последней ступени мог быть измерен шкалой ответов на конкретный вопрос анкеты (индикатор показателя).

Фактически подготовка анкеты должна состоять в том, чтобы перевести показатели в адекватные им вопросы.

Даже в тех случаях, когда формулировка вопроса рождается спонтанно (а в процессе работы над анкетой это случается очень часто), необходимо, исходя из этой формулировки, определить соответствующий вопросу показатель и найти его место в общей схеме программы, что в дальнейшем значительно облегчит анализ полученной информации. И вместо того, чтобы каждый раз говорить, к примеру, о «респондентах, утвердительно ответивших на вопрос: “Считаете ли Вы, что социально-экономическое положение страны в будущем году улучшится?”», в тексте анализа достаточно будет сказать о мнении «социальных оптимистов», поскольку уже в программе была обоснована связь этого вопроса с показателем «социального оптимизма».

Взаимосвязь показателя и формулировки вопроса может быть очевидной (прямой), когда формулировка логически вытекает из определения показателя. Так, уверенность в улучшении социально-экономической ситуации непосредственно связана с показателем социального оптимизма и не требует дополнительного обоснования. Но связь может быть и не столь очевидной, когда, например, ответы на упомянутый выше вопрос предлагается рассматривать в качестве показателя «поддержки экономической политики правительства». И хотя логически вполне допустима интерпретация уверенности в улучшении экономического положения как показателя поддержки экономической политики правительства, желательно провести эмпирическое обоснование этого соответствия.

Стрелки между показателями указывают на гипотезы исследователя. Гипотезы исследователя представляют собой схему анализа первичного материала. Гипотезы могут быть выражены как текстуально (например, «люди, имеющие детей, чаще хотят переселиться из зоны радиоактивного поражения»), так и «стрелками» на схеме программы между показателем такого социально-демографического фактора, как «структура семьи», и таким социально-психологическим показателем, как «отношение к переселению».

Если внесение гипотетических связей слишком загромождает схему, гипотезы можно оформить и в виде заказа на обработку данных. (например, 100x28, 35,...45, где числа означают кодовый номер пункта анкеты).

В качестве примера подготовки программы приведем некоторые фрагменты работы над проблемой «Социальная политика в условиях крупномасштабной аварии на АЭС».

Первоначальный заказ поступил в традиционно неопределенной формулировке — «социально-психологические последствия аварии на ЧАЭС».

1. Совместно с заказчиком конкретизируется цель исследования — определение приоритетности мер социальной политики в случаях подобных аварий на атомных электростанциях. Речь идет о том, чтобы определить и обосновать, какие из возможных мер социальной политики (отселение людей, медицинская помощь на месте, психопрофилактика, пропаганда и т.п.) необходимо осуществлять в первую очередь в подобных ситуациях. (Декларация о том, что подобное не должно повториться, имеет скорее морально-этическую, нежели практическую ценность, ибо доскональное изучение горького опыта явится большей гарантией, нежели абстрактные декларации).

2. Конкретизация и классификация (совместно с заказчиком) возможных мер (меры, направленные на сохранение здоровья; меры, направленные на оптимизацию психологического состояния; меры, направленные на снижение социальной напряженности) позволила определить три основных показателя социально-психологического состояния населения: 1) физическое самочувствие; 2) психологическое самочувствие; 3) социальное самочувствие.

3. Далее исследователь определяет, является ли каждый из показателей простым (измеряемым одним вопросом анкеты или тестом) или составным (в свою очередь, состоит из ряда субпоказателей). Например, «физическое самочувствие» включает в себя, по мнению исследователя: 1) объективные показатели здоровья, 2) самооценку здоровья. В свою очередь, «объективное состояние здоровья» в рамках социологического опроса может быть измерено по таким показателям: 1) наличие хронических заболеваний, 2) частота и длительность заболеваний, 3) частота головных болей, 4) утомляемость (суточная динамика физического состояния) и т.д.

Ниже (в качестве примера) представлена иллюстрация схематической конкретизации показателей «физического самочувствия».

Социально-психологическое
состояние

Физическое
самочувствие

Психологическое
самочувствие

Социальное
самочувствие

1. Самооценка
здоровья

Состояние
здоровья

2. Наличие
хронических
заболеваний

3.
Длительность
заболевания

4. Частота
головных
болей

5.
Утомляемость

Заштрихованные субпоказатели рассматриваются в программе в качестве индикаторов, которые непосредственно трансформируются в вопросы анкеты:

1. «Как Вы оцениваете состояние своего здоровья в целом?»
2. «Есть ли у Вас хронические заболевания?»
3. «Сколько дней за последние 12 месяцев Вы болели так, что приходилось брать бюллетень или соблюдать постельный режим?»
4. «Часто ли у Вас бывают головные боли?»
5. «Какое физическое состояние у Вас бывает чаще всего...
5.А. ... утром. 5.Б. ... вечером

Допустим, наиболее общая схема предстоящего анализа проблемы «Социальная политика в условиях крупномасштабной аварии на АЭС» определяется следующими достаточно альтернативными четырьмя гипотезами.

1. Радиооблучение вызывает ухудшение физического самочувствия, что приводит к ухудшению психологического состояния и социального самочувствия.
2. Ухудшение здоровья населения (независимо от радиации) связывается в сознании людей с ухудшением ситуации и это приводит к ухудшению психологического самочувствия.
3. Предубежденное отношение к ситуации повышает пессимизм, что приводит к психосоматическим расстройствам и реально ухудшает физическое самочувствие и состояние здоровья.
4. Ухудшение общей социальной ситуации (экономического положения, политической ситуации, общей экологической обстановки) приводит к росту психологической

напряженности и плохому самочувствию людей, которое связывается в их сознании с радиационным фактором.

Сбор эмпирического материала и последующая статистическая обработка позволяет определить количественную оценку подтверждения (или опровержения) каждой из выдвинутых гипотез. Последующий логический анализ позволяет предложить приоритетность тех или иных социальных мер.

Так, например, при подтверждении первой гипотезы приоритетом социальной политики должно стать решение проблемы отселения людей с загрязненных территорий.

При подтверждении второй гипотезы в основе социальной политики должны находиться вопросы улучшения медицинского обслуживания, качества информирования населения о причинах ухудшения состояния здоровья.

Если наиболее значимо подтверждается третья гипотеза, то социальная политика должна быть сосредоточена прежде всего на вопросах психопрофилактики и психотерапевтической помощи населению.

И наконец, при подтверждении четвертой гипотезы речь должна идти прежде всего об изменении наиболее общих условий жизни людей, что применительно к жителям регионов, пострадавших от аварии, может иметь частное проявление — улучшение экономических условий жизни посредством дифференцированной системы льгот для населения регионов, пострадавших от аварии.

Разумеется, рассматривая предложенный пример, читатель может не согласиться ни с предложенными показателями, ни с выдвинутыми гипотезами, обоснованно их опровергнуть или дополнить. В этом, собственно, и должно заключаться обсуждение программы исследования авторским коллективом.

В данном же контексте мы лишь попытались убедить начинающего социолога (а «неначинающие» могли убедиться на собственном опыте) в необходимости подготовки программы до начала исследования.

Своевременно подготовленная программа конкретного эмпирического исследования, наряду с общеметодологическими функциями (анализа предмета исследования при подготовке инструментария и синтеза разрозненных индикаторов — пунктов анкеты — в систему показателей изучаемого явления на этапе анализа и обобщения полученных данных), выполняет ряд конкретных методических функций на различных этапах исследования.

Методические функции программы

1. Систематизация плана анализа предмета исследования.
2. Определение структуры понятийного аппарата.
3. Формирование критериев составления и оценки инструментария.
4. Определение выбора статистических методов анализа.

5. Формирование структуры и содержания отчета.

В заключении первой лекции приводится технологическая схема подготовки проекта и программы социологического исследования. Под технологической схемой здесь и в дальнейших разделах понимается последовательность действий, необходимых для выполнения определенного этапа работы. Такое «краткое руководство» призвано выполнять функцию напоминания о наиболее необходимых, с нашей точки зрения, действиях при практическом осуществлении эмпирического социологического исследования.

5. Заключение: технологическая схема подготовки проекта и программы

1. Работа над эскизом проекта: совместно с заказчиком оформляется информационный запрос, уточняются объект, генеральная совокупность и объем выборки, структура и состав исполнителей на различных этапах эмпирического исследования, организационные потребности и возможности, финансовые затраты и сроки проведения исследования, прикидывается общая стоимость работ по проведению исследования.
2. Определение целесообразности проведения эмпирического исследования: сопоставляется необходимость в получении предполагаемой информации с финансовыми затратами на ее получение и принимается одно из возможных решений: 1) от эмпирического исследования целесообразно отказаться в целом, так как затраты на его проведение не окупятся той информацией, которую можно получить в результате; 2) проект необходимо корректировать (сбалансировать объем работы с финансовыми возможностями: сузить круг информационных потребностей, сузить объект исследования, уменьшить объем выборки за счет увеличения ошибки репрезентативности и т.д.); 3) исследование целесообразно проводить в соответствии с подготовленным эскизом проекта.
3. Оформляется и подписывается проект исследования — документ, представляющий собой перечень основных этапов исследования с указанием результатов, стоимости и сроков исполнения каждого из этапов.
4. Разработку содержательной части исследования целесообразно начинать с установления типа исследования: **опрос общественного мнения** или собственно **социологическое исследование**.
5. При подготовке к опросу общественного мнения основное внимание следует уделять подготовке репрезентативной выборочной совокупности; вместо программы исследования вполне допустимо ограничиться информационным запросом.
6. В том случае, когда предстоит выполнять собственно социологическое исследование, особое внимание следует обратить на разработку программы. Разработка программы включает следующие шаги ее подготовки.
 - 6.1. Определение цели исследования. Даже если в первоначальном замысле нет реальной практической цели, автору программы необходимо «целенаправленно» ограничить предмет исследования и круг показателей.

- 6.2. Определение и формулировка предмета (предметов) исследования.
- 6.3. Составление структурной схемы показателей предмета (предметов) исследования.
- 6.4. Оценка возможности измерения каждого из показателей конкретным индикатором (вопросом анкеты). Поиск структурных субпоказателей.

Доведение конкретизации показателей до такой степени, чтобы каждому из них мог соответствовать вопрос анкеты (или тест).

- 6.5. Внесение в схему гипотетических факторов. Определяя фактор, необходимо сразу же давать группировку ответов по этому признаку (например, «возраст»: 1) до 30 лет, 2) 30-55 лет, 3) >55 лет; или «должность»: 1) руководитель, 2) исполнитель и т.п.).
- 6.6. Группировка гипотетических факторов по типам (например, «демографические», «социально-экономические», «социально-политические» и т.п.).
- 6.7. Подготовка плана анализа: обозначение гипотез, которые будут проверяться в ходе статистического анализа. Гипотезы могут быть обозначены в программе одним из наиболее удобных для авторов способом: а) стрелки на схеме — от показателя фактора к субпоказателю предмета; б) текстуально; в) в виде заказа на статистическую обработку.

Лекция 2.

Подготовка анкеты

1. Основные требования к инструментарию (анкете)

Основным инструментарием сбора первичного материала в массовых широкомасштабных социологических опросах является стандартизированная анкета (опросник, опросный лист). Стандартизированным является опросный лист (анкета), отвечающий следующим требованиям: а) каждый из вопросов (пунктов) анкеты предьявляется всем без исключения респондентам в одной и той же фиксированной формулировке; б) определен порядок предьявления вопросов, каждый из которых имеет свой зафиксированный номер (в соответствии с порядком предьявления); в) к каждому из содержательных вопросов прилагается перечень фиксированных ответов. В большинстве случаев варианты ответов предьявляются респонденту; такие вопросы называются закрытыми. В некоторых случаях исследователь не открывает респонденту варианты ответов, предлагая ему отвечать на вопрос произвольно; в таких случаях возможные варианты ответов заносятся в кодификатор, в соответствии с которым на этапе обработки данных произвольные ответы респондентов кодируются перед их вводом в ЭВМ.

Существуют два основных способа сбора информации в массовых опросах населения: самозаполнение опросников (анкет) респондентами и интервью. При использовании первого из способов опроса интервьюер оставляет анкету респонденту, объясняет основные правила ее заполнения и через определенное время забирает заполненную анкету, проверяя качество заполнения. При использовании интервью (опрос «лицом к лицу») вопросы по анкете респонденту задает интервьюер, и он же отмечает в бланке анкеты выбранный респондентом вариант ответа.

При подготовке инструментария способ сбора информации имеет значение лишь на заключительной стадии — при оформлении анкеты.

Основные же правила и требования к конструированию стандартизированной анкеты, предназначенной для массового опроса населения, являются общими как для интервью, так и для опроса способом самозаполнения.

В данном разделе будут рассмотрены основные общие требования к анкете (опросному листу), соблюдение которых при работе над инструментарием позволяет избежать многих ошибок, приводящих к искажению получаемой в конечном результате информации.

Саркастическое замечание руководителя Института демоскопии в Алленсбахе (ФРГ) профессора Э.Ноэль о том, что «ни на одном этапе статистически репрезентативного исследования профаны не предлагают свои услуги столь рьяно, как при составлении анкеты», является весьма актуальным в условиях возросшего спроса на социологические исследования.

Неопытных социологов в работе по подготовке анкеты привлекает, во-первых, эвристичность, творческий характер работы. Во-вторых, получение в результате работы научного документа как зримого свидетельства своей сопричастности к профессиональному сообществу. В-третьих, кажущаяся легкость — вопросы анкеты по своей форме максимально приближаются к привычной повседневной речи. И наконец, возможность после каждого составленного вопроса поставить точку в работе и считать ее выполненной (не случайно, в анкетах начинающих социологов самым недоработанным является блок «демографических» вопросов).

Чем опытнее и квалифицированнее социолог, тем больше трудностей он испытывает на этапе подготовки анкеты.

Почему квалифицированный социолог испытывает трудности там, где дилетант справляется с творческой задачей без всяких проблем? Потому что он хорошо представляет себе, какие **исследовательские задачи** приходится решать в процессе **подготовки инструментария**.

1. Формулируя вопрос, квалифицированный социолог не создает «нечто» из «ничего», а выполняет кропотливую работу «переводчика», переводя исследовательский вопрос с языка категорий, определяющих показатели (в которых, собственно, и ведется анализ изучаемого явления), на язык повседневной речи, привычной для респондентов.
2. Квалифицированный социолог видит ошибки и ограниченность того или иного варианта формулировки, а также набора вариантов ответа и старается их избежать.
3. Формулируя вопрос, квалифицированный социолог одновременно обдумывает и учитывает возможности его статистической обработки и корректирует веер ответов с учетом требований, обусловленных выбором метода статистического анализа, так как каждый отдельный вопрос не живет самостоятельной жизнью и в дальнейшем должен найти свое обоснованное место в общем контексте анализа изучаемой проблемы. После окончания опроса анкетный вопрос должен претерпеть «обратный перевод» — с языка анкеты на язык показателей изучаемого явления.

Поскольку анкета не является независимым научным документом, с которого начинается и которым оканчивается творческая работа, то и оценивать ее качество можно, в

основном, сопоставляя с различными элементами программы (целью, предложенными показателями, предполагаемыми методами статистического анализа).

И хотя в целом конкретная формулировка вопросов, формирование веера ответов и общая композиция анкеты практически всегда в большей степени определяются логикой, вкусом, интуицией и опытом исследователя, чем некоторым стандартным набором правил, все же в процессе подготовки анкеты необходимо учитывать основные требования, игнорирование которых обесценивает получаемую на основе данной анкеты социальную информацию.

По своему функциональному назначению вопросы анкеты можно разделить на содержательные и вспомогательные. Содержательные вопросы являются эмпирическими индикаторами признаков изучаемого явления.

Основными требованиями к содержательным вопросам анкеты являются: валидность (обоснованность), краткость и однозначность.

Валидность

Под валидностью (обоснованностью) подразумевается степень соответствия вопроса анкеты изучаемому показателю. Вопросы могут быть прямыми или косвенными. Валидность прямого вопроса определяется точностью перевода показателя в вопрос. Например, показателю «Степень доверия политическому лидеру N» соответствует вопрос: «В какой степени Вы доверяете N?». Однако тот же самый вопрос может быть и косвенным, если в прогностическом анализе электорального поведения он «работает» на другой показатель — «политический выбор». Утверждение, что тот или иной ответ на вопрос «В какой степени Вы доверяете N?» можно интерпретировать как политический выбор, требует обоснованного доказательства — доказательства валидности.

Существуют три основных способа доказательства (или опровержения) валидности: логический, эмпирический и статистический.

Логическое доказательство. В данном примере логическим обоснованием интерпретации степени доверия как предпочтения (выбора) политического лидера может быть следующий аргумент: чем выше степень доверия человека политическому лидеру, тем выше вероятность того, что он отдаст на выборах свой голос именно этому кандидату. Однако автор вопроса может столкнуться и с контраргументом: возможна ситуация, когда, в большей степени доверяя одному кандидату, избиратель на выборах руководствуется другими мотивами — реальностью победы данного кандидата, особенностями политической ситуации, давлением микросоциального окружения и т. п.

На практике, если при обсуждении анкеты обоснованность вопроса вызвала сомнение у ряда соисполнителей, то обычно на аргументы автора вопроса возникают контраргументы оппонентов, и каждой из сторон собственная логика представляется более обоснованной. В таких случаях в анкету лучше включить оба варианта вопроса, чтобы в дальнейшем процессе анализа статистически оценить валидность каждого из предложенных индикаторов.

Эмпирическое доказательство. Проверка обоснованности путем сопоставления прогноза, сделанного на основе полученной информации, с реальными социальными

фактами (массовым поведением, статистическими данными и т.п.). В рассмотренном выше примере можно сопоставить прогноз результатов выборов, сделанный на основе анализа степени доверия, с реальными итогами выборов.

Статистическое доказательство. Определение меры связи между данными, полученными по косвенному вопросу, с другими критериями, обоснованность которых участникам исследования представляется более очевидной. В нашем примере можно подсчитать коэффициент корреляции между данными, полученными в ответах на вопросы о степени доверия и степени готовности голосовать за данного кандидата на выборах.

Если валидность вопроса ничем не подкреплена, кроме логики автора, то в научных документах (статьях, монографиях, аналитических отчетах, публикуемых таблицах и т.п.) наряду с названием показателя (например, «рейтинг политических лидеров») необходимо приводить полную формулировку вопроса, которая фигурировала в анкете. Это позволяет автору не только проявлять методическую деликатность, но и предоставлять возможность другим исследователям предложить иную интерпретацию вопроса или, сопоставив с другими данными, судить об обоснованности авторской интерпретации.

Краткость

Нет необходимости приводить все аргументы, содержащиеся в социологической литературе и направленные на обоснование принципа краткости в формулировках вопросов. Практически каждому исследователю понятно, что чем длиннее вопрос, тем труднее респонденту понять его основной смысл. Обычно при восприятии длинного и сложного (составного) вопроса респондент чаще всего как бы забывает его первую часть, пока дочитывает или слушает последнюю. Читая (или слушая) вопрос вторично, человек пытается выделить основные (ключевые) слова, определяющие суть вопроса. Если респондент и со второго раза (а тем более — с третьего) не уловил смысл сказанного, он либо пропускает вопрос, либо отмечает любой вариант, создавая видимость ответа.

Эксперименты в области межличностного общения позволили установить, что для большинства людей 11-13 слов являются пределом восприятия фразы без существенного искажения ее основного смысла. Однако потребность исследователя конкретизировать вопрос для получения необходимой информации часто приводит к значительному увеличению длины вопроса.

Что можно предпринять в данном случае? Во-первых, необходимо убедиться в том, что вопрос нельзя сократить за счет вычеркивания отдельных слов (пример 1). Во-вторых, если вопрос представляет собой составное предложение (сложносочиненное или сложноподчиненное), можно разбить его на несколько предложений, и главное из них выделить другим шрифтом (пример 2). В-третьих, подчеркнуть ключевые слова, фокусируя внимание респондента на основном смысле вопроса (пример 3).

1-ый вариант

2-ой вариант

Пример 1

«Сейчас я зачитаю Вам ряд суждений, касающихся России, Украины и СНГ. Затем, пожалуйста, скажите, в какой степени

«Сейчас я зачитаю Вам ряд суждений. Скажите, пожалуйста, в какой степени Вы согласны с каждым из них»

Вы согласны с каждым из утверждений, используя выражения: да, скорее да, скорее нет, нет. И так Вы отвечаете на каждый вопрос.

Пример 2

«Как Вы поступите, если Вам необходимо купить определенную вещь, и Вы специально за ней поехали в магазин, находящийся далеко от Вашего дома, а, зайдя в магазин, увидели, что там большая очередь?»	«Предположим, Вам необходимо купить определенную вещь. Вы специально за ней поехали в магазин, находящийся далеко от вашего дома. Зайдя в магазин, увидели, что там большая очередь.»
---	---

Как Вы поступите?»

Пример 3

«Сколько дней (примерно) за последний год Вы болели так, что приходилось брать бюллетень или соблюдать постельный режим?»	«Сколько дней (примерно) за последний год Вы болели так, что приходилось брать бюллетень или соблюдать постельный режим?»
--	--

Однозначность

Требование однозначности подразумевает одинаковое понимание всеми респондентами того смысла вопроса, который вкладывал в него исследователь. Нарушение однозначности — самая распространенная ошибка в анкетах начинающих социологов.

Первая группа ошибок, проистекающих из нарушения принципа однозначности, — это ошибки, заключающиеся в **формулировках** вопроса и ответов. **Вторая** группа ошибок проистекает из нарушений требований к **шкале измерения** (вееру ответов на вопрос анкеты).

2. Ошибки формулировок

Неопределенность понятий

Если изучаемое понятие не определено и не доведено до эмпирического аналога на этапе подготовки программы исследования, то социолог нередко свою задачу пытается (часто того не осознавая) переложить на плечи респондентов. Так, например, не определив в рамках исследования понятие «адаптация», спрашивает у респондентов, как проходил процесс их адаптации к новому коллективу. Подразумевается, что респондент сам должен не только знать, что такое адаптация, но и дать оценку успешности этого процесса. Но в таких случаях социолог не учитывает, что, оценивая процесс адаптации, разные

респонденты подразумевают различные аспекты своего состояния и условий жизнедеятельности.

В других случаях социологу введенные категории кажутся слишком очевидными и однозначными только потому, что он лично к ним привык и, составляя анкету, не учитывает, что в общественном мнении данная категория может не быть зафиксирована именно в том значении, которое подразумевал автор исследования. Достаточно типичным примером неоднозначного вопроса является следующая формулировка: **«Как Вы оцениваете уровень социальной напряженности в Вашем регионе?»**

В этом вопросе присутствуют сразу два неопределенных термина: «социальная напряженность» и «регион». Понятие «социальная напряженность» в настоящее время не имеет однозначного определения ни в социальной науке, ни в повседневной жизни. Исследователь может по-своему истолковывать эту категорию, а каждый из респондентов по-своему. (На практике, сталкиваясь с неопределенными или неоднозначными терминами, респонденты часто спрашивают интервьюеров — «что имеется в виду», а интервьюеры дают собственное толкование используемому слову). Включая подобную категорию в вопрос анкеты, исследователь фактически перекладывает решение своей задачи (например, «оценка уровня социальной напряженности в различных регионах») на респондентов. Он должен был в программе исследования определить понятие социальной напряженности, выделить составляющие ее компоненты (например, «забастовки», «неудовлетворенность условиями жизни» и т.п.) и по каждому индикатору сформулировать самостоятельный вопрос, а в процессе анализа — интегрировать полученные данные.

Термин «регион», который для автора вопроса, возможно, вполне однозначно определен и типология регионов предусмотрена программой анализа, не может быть одинаково понят разными респондентами. В данном случае необходима конкретизация на административно-территориальной основе (населенный пункт, район, область, республика и т.п.), так как респондент не знает, что под регионом подразумевается в данном случае.

Неконкретность

При изучении фактов для получения сопоставимой информации довольно часто требуется конкретизация места, времени и субъекта, по отношению к которым фиксируется информация. Например, в вопросах типа «Как часто Вы делаете (чувствуете) ...?» необходимо уточнить **временные границы** действий (состояний), о которых требуется получить информацию.

При оценке отношений необходимо четко определять **критерий оценивания**. Например, в вопросе «Как бы Вы оценили нынешнюю политическую ситуацию в Украине?» необходимо уточнить — по какому критерию следует оценить (стабильность, демократичность, перспективы развития или по какому-либо другому критерию, интересующему авторов исследования).

Конкретизация подразумевает сфокусированность вопроса на определенной однозначной информации. Нередко один вопрос содержит сразу несколько неконкретизированных значений. Например, достаточно естественный в обычном разговоре вопрос **«Как Вы распределяете по степени важности следующие статьи расходов семейного бюджета (питание, одежда, лечение и т. д.)?»** совершенно непригоден для

стандартизированной анкеты. В обычном разговоре, как и в нестандартизированном интервью или клинической беседе, подразумевается, что человек, отвечающий на вопрос, может уточнить, о чем собственно идет речь. В рамках же стандартизированного опроса, предполагающего последующую статистическую обработку по стандартной программе, такой вопрос является некачественным, так как он содержит неконкретизированные понятия, и разные респонденты могут его по-разному интерпретировать.

Если вернуться к приведенному выше примеру, то при внимательном прочтении можно увидеть, что один вопрос содержит целый ряд неконкретизированных понятий. Во-первых, неконкретизирован критерий оценки важности (степень важности — по доле в бюджете или по первоочередности приобретения, или по какому-либо другому значению). Во-вторых, не конкретизирован субъект (для кого важно: лично для респондента или для большинства членов семьи, для главы семьи, распоряжающегося бюджетом и т.п.). В-третьих, неконкретизировано слово «распределяете». Распределяет ли респондент реально бюджет в семье или должен проективно распределить в ситуации опроса. Если подразумевалось реальное распределение, то неконкретизировано время — последний год, полгода, месяц, неделя. Даже в стабильных экономических условиях распределение бюджета в семье может со временем претерпевать значительные изменения, а в условиях инфляции распределение семейных бюджетов претерпевает значительные изменения в течение достаточно короткого времени практически во всех семьях.

Если вопрос касается действий респондента, актов поведения, он должен быть конкретизирован отрезком времени. Если речь идет об оценивании, то необходима конкретизация критерия оценки. Часто конкретизация вопроса заключается в предложенных ответах. Однако в любом случае необходимо контролировать, достаточно ли ясно из вопроса и перечня ответов, что конкретно должен выразить респондент.

Сложность подготовки вопросов заключается в том, что требование конкретности и требование краткости вопроса находятся в противоречии и автору приходится искать компромисс между этими двумя требованиями.

Неоднозначность

К сожалению, слишком часто в анкетах встречаются вопросы, формулировки которых подразумевают два, а то и больше собственно вопросов. Ниже приводятся примеры довольно типичных формулировок сочетания двух вопросов, на которые обычно предлагается только одна шкала ответов. Ниже приводятся примеры формулировок вопросов, задуманных и оформленных авторами как один вопрос, а на самом деле заключающих в себе несколько вопросов, на которые респондент не имеет возможности дать ответы.

Пример 1. ***«Часто ли Вы испытываете потребность послушать музыку, сходить в кино, театр, почитать художественную литературу?»***

Пример 2. ***«Каковы Ваши прогнозы относительно последствий аварии на Чернобыльской АЭС для Украины и для Вас лично?»***

Пример 3. ***«Являетесь ли Вы сторонником или противником эксплуатации и дальнейшего строительства АЭС в Вашей области?»***

Обычно подобная ошибка вызвана тем обстоятельством, что для автора исследования не имеет значения, на какую из перечисленных им альтернатив он хочет получить

положительный (или отрицательный) ответ. В предусмотренной им классификации не имеет значения, испытывает ли человек потребность пойти в театр, кино, почитать книгу или послушать музыку. Каждое из этих действий он оценивает как, допустим, потребность в духовной деятельности. Однако он не учитывает, что конкретный человек, заполняющий анкету, может часто испытывать потребность почитать художественную литературу и при этом терпеть не может ходить в кино. Он может иметь разные прогнозы, касающиеся последствий аварии по отношению к себе лично и по отношению к Украине: допустим, по отношению к себе считает, что «все обойдется», а относительно Украины, что «самое худшее еще впереди». Он может быть сторонником эксплуатации уже действующей АЭС и при этом противником дальнейшего строительства. Подобные вопросы очень раздражают респондентов, и, естественно, ответы на них не предоставляют достоверную информацию.

Реже, но все же встречаются случаи, когда автор буквально задает два вопроса в одной формулировке. Например, «Хотите ли Вы уехать из Вашего населенного пункта и собираетесь ли это сделать?». В принципе это допустимо, если из каких-либо методических соображений автор делает такое совмещение сознательно. В таком случае необходимо только проконтролировать, чтобы веер ответов включал все возможные комбинации ответов на первый и на второй вопросы: «хочу и собираюсь», «хочу, но не собираюсь», «не хочу, но собираюсь» и т.д. Чаще всего такое совмещение нецелесообразно — лучше поставить два вопроса.

Тенденциозность

В некоторых случаях пояснения к вопросу являются не необходимой конкретизацией, а подсказкой к ответу, отражающей авторскую гипотезу. Это так называемые «тенденциозные вопросы». Такие вопросы, содержащие тенденциозные пояснения в своей формулировке, являются «слишком однозначными», поскольку их результат можно с уверенностью предсказать заранее. Однако этот результат не представляет собой реальную информацию, а является следствием методической ошибки.

Пример. «Считаете ли Вы достаточной степень социальной защищенности жителей вашего города (района) с учетом работы в Вашей области атомной электростанции?»

В такой формулировке тенденциозность автора проявляется уже в том, что вместо просьбы оценить степень социальной защищенности (нейтральная формулировка, необходимая для получения объективной информации) он использует модальную частицу «ли», основная языковая функция которой — подчеркивание субъективного характера фразы, выражение сомнения. Вторая же часть вопроса — просьба учесть работающую атомную электростанцию — не оставляет сомнений в результатах информации, которая будет получена по этому вопросу.

В подобных формулировках проявляется не только субъективная позиция автора, но и его методическая неопытность. Естественно, что каждый исследователь является человеком и гражданином с определенной личностной позицией. И если автора исследования беспокоит проблема влияния атомной электростанции на чувство социальной защищенности, он вправе исследовать степень этого влияния. Для этого прежде всего необходимо получить объективную информацию о самооценке уровня социальной защищенности у населения в данном регионе, не наталкивая респондентов на те

обстоятельства, которые могут реально на нее влиять (ведь те, кого не опрашивали, могли и не связывать социальную защищенность с работой АЭС). Определить же влияние АЭС на степень социальной защищенности можно в результате сравнительного анализа оценок населения, проживающего вблизи АЭС, с оценками населения, проживающего в другом регионе, сопоставимом по остальным характеристикам. Неквалифицированность социолога в данном случае проявляется в том, что он проблему исследователя — анализ влияния АЭС на уровень социальной защищенности — перекладывает на респондентов.

Наряду с нарушением принципа однозначности, вызванным ошибками в формулировках, часто начинающие социологи допускают ошибки, формируя веер (набор) ответов — основные шкалы измерения изучаемых признаков.

3. Ошибки формирования веера ответов

Нарушение рядоположности

Самой распространенной при формировании веера ответов является логическая ошибка — нарушение рядоположности. Нарушение рядоположности ответов заключается в том, что некоторые варианты ответов не соответствуют ключевому слову вопроса и фактически являются вариантами другого ряда ответов (ответами не на тот вопрос, который поставлен, а на другой).

Пример 1. *«Собираетесь ли Вы уехать из Вашего населенного пункта?»*

1 — да, я хочу уехать

2 — нет, я останусь жить здесь

3 — я еще не принял окончательного решения по этому вопросу

В приведенном выше примере первый вариант, по сути, является ответом на вопрос о желании респондента уехать (хотя в вопросе спрашивается не о желании, а о намерениях). Вполне вероятна ситуация, когда человек не хочет уезжать, но в силу обстоятельств (в частности, радиоактивное или химическое загрязнение, семейные обстоятельства и т.д.) собирается это сделать.

Второй вариант ответа в целом соответствует вопросу, хотя более точной формулировкой была бы следующая — «я не собираюсь уезжать», поскольку формулировка «я останусь жить здесь» слишком категорична и может смущать слишком педантичного респондента — «вообще-то не собираюсь, но мало ли как сложится жизнь».

Третий вариант — это ответ уже на третий вопрос — о твердости намерений.

В конечном итоге, на альтернативный, по замыслу автора, вопрос, подразумевающий выбор только одного варианта ответа, один респондент может отметить несколько ответов («да, я хочу уехать», но «еще не принял окончательного решения»), другой уточнит — «Я хочу уехать, но вынужден остаться жить здесь», а третий спросит: «А если я не хочу уезжать, а обстоятельства меня вынуждают, какой ответ я должен отметить?»

Пример 2 также иллюстрирует нерядоположность вариантов ответов на поставленный вопрос.

Пример 2. «Как регулярно Вы следите за деятельностью Верховного Совета Украины?»

- 1 — постоянно получаю информацию
- 2 — слежу только тогда, когда на сессии обсуждаются интересующие меня вопросы
- 3 — получаю информацию от случая к случаю
- 4 — деятельность Верховного Совета меня не интересует

Если первый и третий варианты соответствуют вопросу о регулярности, то второй вариант является скорее ответом на вопрос о причинах, а четвертый — об интересе к деятельности Верховного Совета. Таким образом, можно видеть, что предъявленные варианты представляют три неполных шкалы оценивания (регулярность, причины, интерес). Человека может не интересовать деятельность Верховного Совета, но он постоянно следит за его работой, так как считает необходимым быть в курсе политических решений и принятых законов. Другой, тоже не интересуясь деятельностью ВС, может получать информацию от случая к случаю, когда другие члены семьи смотрят или слушают трансляцию и т.д. Ответ «слежу только тогда, когда на сессии обсуждаются интересующие меня вопросы» вообще ничего не говорит о регулярности получения информации и просто не может не переkreщиваться либо с первым вариантом («постоянно»), либо с третьим («от случая к случаю»). Формулировка «от случая к случаю» к тому же чаще всего нуждается в конкретизации.

Неполнота множества

Неполнота множества вариантов ответа также является достаточно распространенной типичной ошибкой при формировании веера ответов.

Набор ответов должен включать **все** возможные варианты на поставленный вопрос.

В случаях, когда используется оценочная шкала («Положительно-Отрицательно») или шкала согласия («Согласен-Не согласен»), авторы анкеты редко допускают грубые ошибки в веере ответов. Ошибки, возможные при формировании подобного типа измерительных шкал будут подробнее рассмотрены в лекции, посвященной проблеме измерения. В вопросах же, направленных на изучение мотивов, причин, проблем, вызывающих беспокойство («Какие проблемы беспокоят Вас больше всего в настоящее время?» и т.п.), трудно, а зачастую и невозможно предусмотреть все варианты ответов. Однако, не включив какой-либо возможный вариант в перечень ответов, автор допускает двойную ошибку: психологическую и логико-математическую. Не предоставив человеку возможность ответить на вопрос, он формирует у респондента, во-первых, чувство неудовлетворенности, а во-вторых, представление о том, что вопросы можно оставлять и без ответа. Математическая ошибка заключается в том, что в обработку поступает неполное множество.

Чтобы в подобных вопросах исключить ошибку неполноты множества, следует руководствоваться следующим правилом: в конце перечня вариантов ответов использовать **«закрывающую триаду»**:

- а) **дополнение** («другой ответ» _____)
- б) **отрицание** («никакие проблемы не волнуют»)
- в) **уход** («трудно сказать», «затрудняюсь ответить» и т.п.)

Вариант «другой ответ», кроме того, что он в любом случае позволяет каждому респонденту выразить свое мнение, по каким-либо причинам непредусмотренное автором анкеты, помогает еще и оценить качество вопроса в целом. Если в результате опроса этот вариант отметили более 5% респондентов, вопрос считается некачественным и требует доработки — расширения веера ответов на основе изучения конкретных ответов, записанных в строке «другой ответ».

Включение в веер ответов варианта «затрудняюсь ответить» — традиционная форма предоставления респонденту возможности уйти от ответа, если тот по каким-либо причинам не может или не хочет отвечать на данный вопрос. Автор может пойти на «жесткую форму» вопроса и не включить в набор ответов вариант «затрудняюсь ответить». Этот прием применяется для того, чтобы принудить респондентов выбрать тот или иной конкретный ответ. Но в этом случае, во-первых, он должен осознавать, что неизбежно возрастет доля неответов на этот вопрос; а во-вторых, обязательным условием при анализе материала является проверка достоверности данных, полученных с помощью этого вопроса.

Неполнота множества может порождаться разными причинами, среди которых довольно распространенной является элементарная небрежность в работе. Ниже приводятся примеры вопросов с неполными перечнями ответов, очевидность которых проявляется при первых же попытках апробации инструментария.

Пример 1. **«Какая часть людей, живущих по соседству с Вами — русские?»**

- 1 — все
- 2 — большинство
- 3 — некоторые из них
- 4 — мало кто
- 5 — никто
- 6 — не знаю

Пример 2. **«Как часто за последние три года Вы принимали участие в голосовании, выборах?»**

- 1 — ни разу
- 2 — от случая к случаю

3 — часто

4 — не знаю

В примере 1, кроме того, что трудно дифференцировать варианты «некоторые из них» и «мало кто», не предусмотрен такой ответ, как «примерно половина», который будут пытаться найти многие респонденты и, не найдя, могут отказаться отвечать как на этот вопрос, так и на аналогичные последующие вопросы, например, касающиеся коллег, родственников, друзей и т.д.

В примере 2 не предусмотрен вариант ответа «всегда». Учитывая, что участие в выборах обычно принимает большая часть взрослого населения, можно было предвидеть, какая доля респондентов выразит протест, не найдя «своего» ответа на шкале. В то же время вариант ответа — «не знаю» может вызвать закономерное удивление.

Шкала ответов на вопросы, касающиеся частоты (регулярности) каких-либо действий, всегда представляет для авторов анкет большую трудность. Если полярные точки шкалы, такие как «постоянно» («всегда», «регулярно») и «практически никогда», обычно затруднений не вызывают, то промежуточные («часто», «иногда», «изредка» и т.п.) в силу неопределенности единицы измерения часто вызывают бурные дискуссии среди соисполнителей. Поэтому, если исследование не направлено на измерение именно субъективного аспекта восприятия, лучше подобные вопросы (или варианты ответов) конкретизировать — «Сколько раз за последний год (месяц, неделю)...?»

Порядок ответов

В тех случаях, когда исследователь предполагает в последующем анализе считать индексы и коэффициенты связи, он должен обратить внимание на порядок предъявления вариантов ответов, поскольку цифровое обозначение кодовой позиции каждого из вариантов ответа приобретает свое числовое значение. Подробнее на этой проблеме — проблеме соотношения веера ответов с типом шкалы — остановимся ниже (в разделе «Анализ и проблема измерения»). Здесь же хотелось бы только привлечь внимание авторов анкеты к тому, что, когда формируются вопросы оценочного характера (исследуется эмоциональное отношение, доверие, удовлетворенность и т.п.), место такого варианта ответа, как «Затрудняюсь ответить», приобретает тот или иной вес в зависимости от рангового места, на которое автор поставил этот вариант в веере ответов.

При использовании оценочных вопросов («Как Вы относитесь...?», «Как Вы оцениваете...?») вариант ответа «затрудняюсь ответить» лучше ставить не в конце, а в середине шкалы, преобразуя его формулировку в соответствии с общей шкалой.

Пример 1. ***«Удовлетворены ли Вы своей работой?»***

1 — совершенно не удовлетворен

2 — скорее не удовлетворен

3 — трудно сказать, удовлетворен или нет

4 — скорее удовлетворен

5 — полностью удовлетворен

Пример 2. *«Как Вы оцениваете деятельность Верховного Совета Украины?»*

1 — скорее положительно

2 — трудно сказать, положительно или отрицательно

3 — скорее отрицательно

Какую шкалу лучше выбрать — трехбалльную или пятибалльную? Это зависит от того, какой вид анализа будет применен в дальнейшем. Если предполагается расчет индексов и коэффициентов связи — лучше пятибалльную. Если же автору достаточно сопоставить группы людей (удовлетворенных и неудовлетворенных, согласных и не согласных с каким-либо суждением), лучше пользоваться трехбалльной шкалой. Но в этом случае следует смягчить категоричность ответов: «скорее согласен», «в основном положительно».

В последнее время в анкетах часто используются десяти- и одиннадцатibalльные шкалы. Например: «Оцените свое отношение к некоторым известным политическим лидерам. Свою оценку поставьте на шкале термометра (от нуля до десяти). Чем больше число, тем теплее или благожелательнее Вы относитесь к этой личности. Чем меньше число, тем холоднее или неблагоприятнее Ваше отношение. Если Вы не испытываете к ним ни благосклонности, ни неприязни, выберите оценку 5».

Это пример одиннадцатibalльной шкалы с так называемым фиксированным нулем, где, по замыслу исследователя, оценки выше пяти баллов представляют собой положительную ось шкалы, а оценки ниже пяти баллов — отрицательную.

Мы советуем с большой осторожностью использовать одиннадцати и особенно десятибалльные шкалы.

Анализ отчетов интервьюеров показывает, что далеко не все респонденты «чувствуют» такого рода шкалы. И в первую очередь это относится к лицам, не имеющим высшего образования. Например, оценивая Сталина, респондент говорит: «Я отношусь к нему в целом положительно, он много сделал для нашей страны, но совсем немного положительно, совсем немного, так как он был жестоким. Отметьте, пожалуйста, 1 балл». По замыслу автора, «совсем немного положительно» — это 6 баллов. Однако респондент категорически отказался дать 6 баллов — «это слишком много».

Еще сложнее проходят опросы по десятибалльным шкалам, так называемым «шкалам без фиксированного нуля». С одного до пяти баллов шкала отрицательная, с шести до десяти — положительная. Многие респонденты, занимающие нейтральную позицию, не могут найти ее на шкале. Они требуют, чтобы интервьюеры обвели одновременно 5 и 6 баллов, или ставят между этими позициями птичку и обводят ее (если сами заполняют анкету). И если автор не контролирует каждое интервью и ввод данных по каждой анкете, то он получает результаты, «откорректированные» интервьюерами и операторами, занимающимися вводом данных. Поэтому 10 — 11-балльные шкалы лучше оставить для основательных исследований, где время и методическая квалификация исследователя позволяют оценить качество информации, полученной с помощью эти шкал.

4. Дополнительные вопросы

Дополнительные вопросы выполняют в исследовании различные вспомогательные функции: помогают жесткую стандартизированную форму опроса приблизить к диалогу, отфильтровать компетентных респондентов, собрать дополнительную информацию, которая поможет правильно интерпретировать полученные результаты, проверить репрезентативность данных и т.п.

Наиболее распространенными среди дополнительных вопросов являются вопросы-фильтры и открытые вопросы.

Вопросы-фильтры

Под вопросом-фильтром подразумевается вопрос, отсеивающий тех респондентов, к которым не относится последующий вопрос. Например, вопросу *«Посещают ли Ваши дети музыкальную школу?»* должен предшествовать вопрос-фильтр *«Есть ли у Вас дети школьного возраста?»* Вопросу *«Какая национальность у Вашей жены (мужа)?»* должен предшествовать вопрос-фильтр *«Есть ли у Вас жена (муж)?»*

Поскольку любой вопрос занимает определенное физическое место, и перед автором анкеты всегда стоит проблема — какие вопросы можно отсеять без существенного ущерба для содержания и качества информации, то следует учитывать, что во многих случаях нет необходимости ставить прямой вопрос-фильтр. Достаточно, задавая основной вопрос, включить в веер ответов вариант типа «у меня нет детей школьного возраста» или «у меня нет жены (мужа)» и т.п. Такая форма отсеивания позволяет сокращать общий объем анкеты; это удобнее и при анализе полученных данных, так как позволяет без дополнительных сопоставлений сравнивать разные группы респондентов. Кроме того, позволяет дополнительно проконтролировать качество заполнения анкет и ввода данных.

Открытые вопросы

Открытым называется вопрос, в котором автор не предлагает респонденту набор подготовленных ответов, а оставляет место для свободного ответа в произвольной форме.

Пример 1. *«Что такое демократия?»* _____

Пример 2. *«Если Ваше здоровье ухудшилось, то в чем Вы видите главные причины его ухудшения?»*

Обычно социолог ставит открытые вопросы по двум причинам. Во-первых, и чаще всего, потому что не имеет достаточно полного представления о предмете исследования. Он

предполагает, что, получив ответы респондентов, изучит их, на основе изучения составит классификатор и кодификатор, открытые вопросы будут закодированы и таким образом преобразованы в стандартизованный веер ответов для дальнейшего анализа.

Во втором случае социолог заранее имеет кодификатор (перечень ответов), но он хочет скрыть его от респондента для того, чтобы установить, какая доля респондентов напишет определенный ответ, если он не подсказывается анкетой (в этом случае открытый вопрос относится не к вспомогательным, а к содержательным).

И в том, и в другом случае социолог допускает ошибку. Опыт показывает, что на открытые вопросы отвечает обычно не более трети респондентов. Формулировки их ответов носят либо слишком стереотипный и общий характер, либо очень конкретный. Анализировать ответы общего характера нецелесообразно, так как многие респонденты не дали его просто потому, что им было лень писать. Классификация же конкретных формулировок, да еще и без определенного основания (критерия) классификации, настолько трудоемкий и длительный процесс (каждый новый вариант формулировки ставит перед автором анализа проблему — вводить ли новый класс ответов или возможно соотнести данный ответ с предыдущими вариантами), что чаще всего не доводится автором добросовестно до конца, и докодировка оставшихся вариантов ответов перекладывается на технических исполнителей, каждый из которых классифицирует ответы по своему усмотрению.

Но из сказанного не следует, что от открытых вопросов надо отказываться. Открытые вопросы целесообразно использовать в анкетах по следующим причинам.

1. У респондента, отвечающего на закрытые вопросы анкеты, может возникнуть желание более подробно высказаться по поводу предмета опроса, особенно если у него есть определенная позиция или были в жизни конкретные эпизоды, имеющие отношение к предмету исследования (например, причины ухода с работы или обращения в инстанции и т.п.). Он испытывает чувство неудовлетворенности от слишком стандартных ответов, предусмотренных анкетой. Открытый вопрос является своего рода отдушиной, позволяющей человеку высказаться, если у него возникла такая потребность в связи с предметом опроса. (Часто в анкетах, в которых отсутствуют открытые вопросы, люди пишут свои дополнения и комментарии на полях анкеты или в проветах между текстом).

2. Ответы на открытые вопросы не обязательно анализировать по стандартной программе. Можно использовать методы контент-анализа при их обработке. Часто автору анкеты достаточно просто все их внимательно прочитать. Ознакомление с ответами людей на открытые вопросы часто помогает исследователю «почувствовать» живых людей, которые стоят за цифрами, полученными в результате статистического анализа.

Отдельные наиболее типичные или наиболее яркие комментарии могут быть включены в текст отчета (или другого научного документа). Это позволяет не только исследователю, но и читателю «увидеть» живого человека за сухими статистическими выкладками.

3. Ознакомление с ответами на открытые вопросы поможет исследователю в углублении разработки данной проблемы, если он предполагает продолжать ею заниматься.

***Вопросы социально-демографического
блока анкеты***

Блок вопросов, касающихся социально-демографических характеристик респондента, обычно включает анкету, предназначенную для социологического исследования. Какие характеристики следует в него включать? Этот вопрос, которого социологи почему-то избегают в начале работы над программой, неизбежно встает перед исследователем, когда он считает, что работа над анкетой практически закончена. Ощущение завершенности нередко подводит исследователя, и он может небрежно оформить эту часть работы; впоследствии (при анализе полученного материала) испытывает трудности, связанные с недостатком информации, необходимостью перегруппировки данных и т.п., так как значительная часть анализа и представления данных отводится именно влиянию социально-демографических факторов на изучаемое явление.

Выбор демографических характеристик должен определяться в первую очередь гипотезами исследования. Основная проблема при подготовке демографического блока заключается, пожалуй, не только и не столько в отборе демографических характеристик, сколько в классификации и группировке по ним респондентов. Технически это означает, что особое внимание при подготовке блока демографических вопросов следует уделять закрытому вееру ответов. Даже при открытом вопросе анкеты, например, **«Ваш возраст?»** программой должен быть предусмотрен первоначальный вариант группировки данных по этому признаку.

Другим ориентиром при отборе демографических характеристик являются документы государственной статистики и ведомственной отчетности. Сбор информации по тем признакам, которые имеются в статистических документах, всегда позволит социологу сопоставить материал, полученный в результате исследования, с имеющимися статистическими данными, оценить репрезентативность выборки, качество информации и расширить возможности анализа и интерпретации.

Очень сложным является вопрос, касающийся социально-профессионального положения. Тот, кому приходилось сталкиваться с разработкой этого вопроса, знает, как трудно найти такой критерий классификации, который, с одной стороны, был бы обоснован теоретически, а с другой — позволил бы каждому респонденту найти свое место среди предложенных вариантов. В практике социологических опросов существуют эмпирически отработанные варианты данного индикатора, одна из моделей которых отработана в Институте социологии НАН Украины: на вопрос «Род Ваших занятий?» предлагается следующий перечень ответов:

- 1 — Профессиональный политик, руководящий работник госаппарата
- 2 — Руководитель (заместитель руководителя) предприятия, учреждения; председатель колхоза, совхоза
- 3 — Служащий госаппарата (должностное лицо)
- 4 — Специалист технического профиля (с высшим или средним специальным образованием)
- 5 — Специалист в области науки, культуры, здравоохранения, образования, дошкольного воспитания (с высшим или средним специальным образованием)
- 6 — Работник правоохранительных органов, военнослужащий
- 7 — Предприниматель, занятый крупным или средним бизнесом

- 8 — Занимаюсь мелким бизнесом, работаю на бизнесмена
- 9 — Служащий из числа вспомогательного персонала
- 10 — Квалифицированный рабочий
- 11 — Разнорабочий, подсобный рабочий
- 12 — Колхозник, работник совхоза
- 13 — Фермер
- 14 — Учащийся
- 15 — Неработающий пенсионер
- 16 — Домохозяйка
- 17 — Не имею постоянного места работы, но подрабатываю в разных местах в зависимости от обстоятельств
- 18 — Не работаю и не имею никаких источников дохода
- 19 — Другой ответ _____

(напишите)

Вопрос о возрасте лучше задавать в открытой форме: «**Ваш возраст?**» _____ лет. Такая форма вопроса предоставляет больше возможностей для анализа.

Вопрос о национальности тоже лучше оставлять открытым: «**Ваша национальность?** _____», с последующей кодировкой. Это лучше делать по двум причинам. Все национальности в анкете учесть невозможно, да и не целесообразно. Человеку же, национальности которого нет в перечне, неприятно писать себя в графе «другая» (некоторые авторы пишут и «другое»). Многими респондентами болезненно также воспринимается порядок (место), на котором его национальность указана в перечне ответов.

5. Оформление анкеты

Нет особой необходимости подробно останавливаться на проблеме оформления анкеты. Следует только заметить, что многолетний опыт проведения опросов позволяет отметить, что качество полиграфического, технического и эстетического дизайна бланка анкеты, безусловно, повышает интерес, ответственность и качество работы в целом как интервьюеров, так и респондентов.

Несомненно, это обстоятельство понимает каждый исследователь. Однако он нередко бывает лимитирован техническими и финансовыми возможностями. Здесь только хотелось бы напомнить, что не следует экономить средства и стремиться сделать из бланка анкеты «полиграфический шедевр». В конечном итоге, это окупится снижением

количества отказов от участия в опросе и, главное, общим повышением качества заполнения.

На наш взгляд, «научить» оформлять анкету невозможно. Здесь самое широкое поле для проявления вкуса, интуиции и изобретательности автора инструментария. Однако можно предложить наиболее общие рекомендации, которые могут быть полезны на заключительном этапе работы по созданию основного инструментария.

1. Желательно, чтобы бланк опросника имел обложку с названием анкеты. Это название не обязательно должно отражать название темы (или цели) исследования. Важно, чтобы оно согласовывалось с содержанием вопросов «в глазах респондентов». На обложке обычно вверху указывается полное название организации, проводящей опрос, внизу — год проведения опроса.
2. Вопросы анкеты лучше разбить на «тематические» блоки (вопросы, связанные одной проблемой); каждый из блоков следует предварять либо названием, либо небольшим вступлением, либо общеустановочным вопросом.
3. Порядок следования как блоков, так и вопросов внутри каждого блока должен устанавливаться не логикой исследователя (зависящей от логики последующего анализа), а логикой респондента — человека, отвечающего на вопросы. Другими словами, ему должно быть «как бы понятно» — почему и для чего ему задают «именно эти вопросы». Поэтому часто в анкете приходится ставить вопросы, которые продиктованы не потребностями анализа (как правило, в последующем анализе они и не используются), а для связки отдельных вопросов или для перехода от одной проблемы к другой. Например, переход к блоку вопросов, направленных на установление политической ориентации людей, может предваряться вопросом: «Интересуетесь ли Вы политикой?» или «В какой степени Вас интересует политика?» и т.п.

Существуют определенные различия в оформлении анкет, предназначенных для самозаполнения и для интервью. В первую очередь это касается общей инструкции по заполнению анкеты. При интервью инструкция по требованиям к заполнению анкеты готовится на отдельном бланке; при самозаполнении — перед вопросами (лучше на отдельной странице) печатается общая инструкция респондентам по требованиям к заполнению опросника. Эта инструкция должна быть короткой. Здесь достаточно сосредоточить внимание респондентов на том, что необходимо дать ответ на все вопросы, а также на том — как именно следует отмечать подходящие ответы (обводить кружком кодовый номер или подчеркивать ответ и т.п.), как производить исправления, если респондент хочет изменить свой ответ.

Каждый из вопросов (пунктов) в большинстве случаев целесообразно сопровождать дополнительной конкретной инструкцией: например, «Отметьте один вариант ответа» или «Отметьте все подходящие, с Вашей точки зрения, варианты ответов», или «Отметьте не более 3-х, наиболее подходящих, с Вашей точки зрения, вариантов ответов» и т.п. Такие конкретные инструкции-напоминания полезны даже интервьюеру, который должен быть хорошо знаком с анкетой до начала интервью, но в процессе беседы с респондентом может упускать из виду некоторые требования. Тем более необходимы такие инструкции-напоминания респонденту, который даже при внимательном ознакомлении с общей инструкцией, конечно же, не сохраняет в памяти все требования, изложенные в начале бланка опросника.

Кроме вопросов и инструкций к ним, бланк анкеты нередко требуется снабжать и другими дополнительными текстами: вступлениями-комментариями к блокам анкеты или отдельным вопросам; после вопросов-фильтров следует указывать, каким группам респондентов следует пропустить несколько вопросов, не имеющих к ним отношение и т.п.

При подготовке сопроводительных текстов к вопросам надо учитывать, что в случаях самозаполнения все комментарии, обращения и инструкции адресуются непосредственно респонденту; в задачу интервьюера входит лишь контроль за тем, чтобы респондент обратил на них внимание и правильно понял. В случаях, когда сбор информации осуществляется способом интервью, все соответствующие ремарки, комментарии и инструкции в тексте бланка анкеты адресуются интервьюеру; кроме того, обычно в интервью используются карточки с текстом вариантов ответов к тем вопросам, перечень ответов к которым слишком длинный и громоздкий для его адекватного восприятия «на слух».

Размещая содержание подготовленной анкеты на страницах бланка, социолог обычно сталкивается с проблемой расположения вопросов. Вопрос с прилагаемым к нему перечнем ответов должен быть размещен на одной странице; нельзя переносить на другую страницу часть вариантов ответов, оставляя вопрос вне поля зрения респондента. Часто в результате размещения вопросов на страницах бланка анкеты социолог не удовлетворен полученной «картинкой» — на одних страницах «густо», на других — «пусто». В данном случае следует учитывать, что респонденту легче воспринимать отдельный вопрос — его основной смысл и содержание — тогда, когда он четко выделен на странице из общего текста. Чем больше интервалы между отдельными пунктами анкеты, тем лучше. Однако, учитывая, что увеличение количества страниц одного бланка анкеты приводит к многократному увеличению общего расхода бумаги, можно руководствоваться следующим правилом: в тех случаях, когда возникают проблемы с компактным размещением вопросов, лучше уменьшать интервалы между строчками одного вопроса, но увеличивать интервалы между пунктами анкеты. В целом все же хотелось бы напомнить, что при оформлении анкеты в большей степени следует руководствоваться качеством анкеты, чем соображениями, связанными с возможностью экономии бумаги. «Спрессованная» в результате экономии бумаги анкета может настолько снизить качество информации, что значительно дешевле было бы вообще не проводить опрос.

6. Процесс подготовки анкеты

Почти каждый социолог-исследователь испытывал чувство неудовлетворенности, обнаружив ошибки и погрешности в основном инструменте сбора информации — анкете, тиражированной сотнями (а то и тысячами) экземпляров. Это и содержательные ошибки, вызванные поспешностью и непродуманностью, и чисто технические, возникшие в ходе подготовки заключительного экземпляра анкеты. Типичность подобной ситуации порождается, с одной стороны, определенным психологическим законом эффектом оконченного действия (закончена основная творческая часть работы), с другой стороны, отсутствием у руководителя исследования технологической схемы по завершению процесса подготовки инструментария (тем более, что и в методической литературе этому этапу практически не уделяется внимание).

График работы над анкетой

При составлении общего графика работы по организации социологического исследования следует руководствоваться принципом **«не экономить время»** на подготовке анкеты. Время можно сэкономить на этапе сбора информации за счет привлечения большего числа интервьюеров или на этапе ввода информации, увеличив число операторов, но, создавая основной инструмент получения информации, ради которой, собственно, и выполняется огромная работа в целом, такая экономия крайне нерациональна.

Оптимальный срок подготовки инструментария эмпирического социологического исследования — 3-4 недели. (При проведении экспресс-опросов общественного мнения, основной целью которых является оперативный сбор информации о реакции населения на какие-либо события, политические решения и т.п., анкета, естественно, готовится в более сжатые сроки; но в таких случаях привлекаться к работе может только квалифицированный специалист, имеющий большой практический опыт подготовки инструментария).

Примерный график работы над анкетой включает два основных этапа, которые в соответствии с особенностями работы могут быть условно названы этапом «штурма» и этапом «осады».

Первый этап (этап «штурма») охватывает 9-10 дней. Здесь желательно не терять времени даром и не выбиваться из достаточно жесткого графика работы.

1-ый день. Руководитель на основе программы исследования составляет список тематических блоков анкеты (например, «социальная напряженность», «политические установки» и т.д.) и дает соисполнителям задание. Руководитель может назначить на эту работу ответственного исполнителя, в таком случае в последующем тексте вместо руководителя подразумевается ответственный исполнитель.

2-3-ий день. Каждый из соисполнителей готовит первый вариант вопросов по своему блоку, а также вопросы, касающиеся социально-демографических характеристик (в соответствии с гипотетическими факторами влияния на изучаемое явление). Более двух дней тратить на подготовку первого варианта нецелесообразно. Количество вопросов при подготовке первого варианта не следует искусственно ограничивать, руководствуясь ориентировочно намеченным общим объемом инструментария. В противном случае исследователь может не найти интересные варианты, которые приходят в голову не сразу, а в результате своеобразного «индивидуального мозгового штурма», эвристическая ценность которого как раз и обусловлена снятием первоначальной стереотипности, которая обычно проявляется в самом начале работы по составлению вопросов. Сокращение вопросов с целью оптимизации общего объема анкеты лучше производить на последующих стадиях работы над инструментарием.

4-ый день. Техническая подготовка первого варианта блоков вопросов. Желательно обеспечить перепечатку всех блоков вопросов в течение одного дня, независимо от общего объема работы. Количество отпечатанных экземпляров должно соответствовать числу исполнителей, каждому из которых в конце дня руководитель передает полный набор блоков вопросов. Если исполнитель один, то он должен передать первый экземпляр коллегам, выполняющим в данном случае роль экспертов.

5-ый день. Соисполнители дорабатывают свой блок и конструктивно рецензируют все остальные части анкеты — не только делают замечания, но и вносят свои предложения. Руководитель продумывает структуру анкеты — расположение блоков в бланке анкеты.

6-ой день. Коллектив исполнителей собирается для обсуждения первого варианта анкеты. В ходе обсуждения соисполнители обмениваются редакционными замечаниями, решают, какой вариант вопроса лучше с точки зрения измерения показателя изучаемого явления, какие вопросы следует отсеять, каких индикаторов недостаточно для измерения изучаемого объекта. Особое внимание в процессе обсуждения уделяется вееру ответов, который необходимо оценивать с учетом предполагаемых методов статистического анализа полученной информации, что нередко выпадает из поля зрения исследователей при обсуждении инструментария. Соисполнитель в процессе обсуждения его блока вопросов должен записать все высказанные замечания и рекомендации, а в конце обсуждения получить от каждого из коллег свой блок с ранее отмеченными замечаниями.

7-8-ой день. Доработка блоков вопросов с учетом замечаний. Руководитель дорабатывает блок социально-демографических вопросов. В конце восьмого дня соисполнители передают доработанные блоки вопросов руководителю.

9-ый день. Перепечатка анкеты в целом. Перепечатанную и вычитанную анкету необходимо размножить в количестве 30-40 экземпляров.

Самая большая ошибка при подготовке анкеты — отдавать в тираж первоначальный вариант сразу же после его подготовки.

Апробация и доработка инструментария

После подготовки первого варианта, собственно, и начинается работа, необходимая для разработки инструментария, пригодного для получения качественной информации. В методической литературе этот этап принято называть апробацией анкеты или пилотажным исследованием. Речь идет о микромодели крупномасштабного исследования, по которой опрашивается несколько десятков респондентов, и на основании анализа их ответов вносятся коррективы в анкету или в инструкцию к опросу.

Поскольку эта работа требует дополнительных сил, времени, средств и ресурсов, то исследователи очень часто вообще отказываются от проведения этого этапа (правда, чаще всего об этом скромно умалчивают).

Однако между повышенными академическими требованиями, не учитывающими жизненных реалий, и широко распространяющейся, ничем не обоснованной и не выдерживающей никакой профессиональной критики практикой существует компромиссный путь завершения работы над инструментарием. Он заключается в соблюдении следующих правил.

1. Анкета должна «отлежаться». Первый вариант анкеты, даже если он не вызывает никаких сомнений у разработчиков, не следует сразу же тиражировать. Необходимо время, чтобы проявились латентные (не всегда осознаваемые) творческие возможности участников исследования, которые были инициированы первым форсированным этапом работы. Психологический механизм творческой работы таков, что исследователь, даже завершив какой-либо этап работы, осознанно или неосознанно возвращается к анализу

поставленной проблемы. Во-первых, в голову приходят новые дополнительные варианты редакции вопросов или возможных показателей, неучтенные факторы и т.п. Во-вторых, перечитывая еще и еще раз анкету, исследователь может сам обнаружить ошибки, которые не были замечены в процессе подготовки первоначального варианта, когда внимание распределяется и на содержательную часть, и на возможности верификации, и на особенности формулировок, и на порядок вопросов и т.п. элементы работы над инструментарием.

2. Исследователь должен своими глазами увидеть «живую реакцию» потенциальных респондентов на разработанный инструментарий. Другими словами, он должен сам по подготовленной анкете опросить несколько человек, чтобы увидеть, как «работают вопросы». В качестве таких респондентов могут выступать родственники или знакомые. В определенном отношении хорошо знакомых людей даже лучше опрашивать на этом этапе работы, так как они в меньшей степени реагируют на самого «интервьюера», не стесняются переспросить в том случае, если не понимают смысл вопроса, высказать критические замечания и т.п.

В последующей доработке анкеты исследователь, естественно, будет учитывать не каждое замечание «человека с улицы», тем не менее, кроме отдельных полезных для себя наблюдений, он может установить некоторые типичные реакции, которые необходимо учесть или в самой анкете, или в инструкции интервьюерам.

Такие пробные интервью позволяют также убрать из анкеты «неработающие» вопросы и дополнить полузакрытые вопросы, если в вариативном вопросе вариант «другой ответ» выбирают более 5% опрошенных всеми соисполнителями. «Неработающими» считаются вопросы, не позволяющие дифференцировать респондентов, например, если на какой-либо вопрос более 80% опрошенных дают один и тот же ответ. «Неработающими» также считаются вопросы, на которые более половины людей затрудняются ответить.

Кроме того, апробацию проходят формулировки, которые при обсуждении анкеты ряду соисполнителей казались сомнительными с точки зрения возможности их восприятия «простыми людьми». Нередко предварительно проведенные опросы показывают: то, что кажется слишком трудным для понимания самим социологам, не вызывает трудностей и негативных реакций у опрашиваемых людей.

3. Анкета должна пройти неформальную экспертизу. Неформальная экспертиза инструментария коллегами-социологами часто бывает полезнее официальных обсуждений. Во время формального обсуждения эксперт может, во-первых, из соображений гуманного характера не высказать все критические замечания, а во-вторых, не выносить на всеобщее обсуждения вещи, которые вызывают у него сомнения, но он не полностью уверен в своей позиции, в-третьих, высказав свою точку зрения публично и вполне официально, эксперт в определенной степени чувствует необходимость ее отстаивать даже тогда, когда под влиянием аргументов автора в глубине души уже не полностью уверен в своей правоте. Неформальная же просьба к коллеге посмотреть инструментарий и высказать свои замечания позволяет эксперту в полном объеме высказать все замечания и указать на обнаруженные им ошибки. Естественно, не все замечания и предложения коллег будут полностью приняты. Но, по крайней мере, у авторов анкеты возникнет необходимость либо более аргументированно остановиться на первоначальном варианте, либо с благодарностью за затраченные усилия исправить незамеченную в первоначальном хаосе работы ошибку. Поэтому мы настоятельно рекомендуем подготовленную анкету «подсовывать» коллегам, имеющим опыт подобной работы.

Именно на этот второй этап работы над анкетой не следует жалеть времени: 1-2 недели должны быть потрачены на то, чтобы избавиться от всех возможных ошибок и погрешностей, которые неизбежно возникают на первом этапе в силу специфики этой работы, отличающейся многоуровневым распределением внимания.

На заключительном этапе — перепечатке окончательного варианта особое внимание необходимо уделить вычитыванию (корректуре) тех частей анкеты, где были сделаны содержательные исправления.

Даже анкеты, подготовленные в оперативном порядке для экспресс-опросов, при проведении которых нет времени выдерживать необходимую технологию, следует задним числом (в период сбора информации) подвергнуть прохождению той же процедуры, чтобы по крайней мере при анализе результатов была возможность учесть погрешности в качестве инструментария.

Лекция 3.

Выборка: определение выборочной совокупности

1. Понятие и проблемы формирования выборки

Основные проблемы

Выборочный подход возник в качестве альтернативы **сплошному** опросу, наиболее типичными примерами которого являются перепись населения и референдум.

В практике эмпирических социологических исследований сплошные опросы проводятся крайне редко: в основном в локальных исследованиях, когда объект исследования представлен небольшой группой людей (одно предприятие и т.п.).

Основной особенностью опросов общественного мнения является то, что обычно опрашивается небольшая часть какой-то группы населения, а выводы делаются обо всей группе. Какое соотношение должно быть между этой группой в целом и той частью группы, от которой фактически и была получена информация, — основная проблема выборки.

Когда специалист и неподготовленный читатель смотрят на одни и те же цифры, отражающие результаты опроса населения, они «видят» их во многом по-разному.

Если человек, не обладающий достаточными социологическими знаниями, знакомясь с результатами опроса общественного мнения, видит, что, допустим, на вопрос «Одобряете ли Вы политику президента в области проведения экономических реформ?» 31% ответили «Да», 43% «Нет», а 26% — «Затрудняюсь ответить», то он каждую цифру переносит на все население буквально. Специалист эти данные «читает» примерно следующим образом: «Да» ответили 31% (плюс-минус 3% или плюс-минус 5%, или плюс-минус 10% и т.д.). Под этим подразумевается, что если бы на данный вопрос отвечала не часть населения, которая была опрошена, а все население, то полученный результат с определенной вероятностью находился бы в указанных пределах.

Величина, которая определяет в данном случае разброс возможного значения данных, полученных в результате опроса, называется *допустимой ошибкой выборки*, которую всегда необходимо учитывать, когда данные опроса **части** населения переносятся на население **в целом**.

Величина ошибки выборки зависит от ряда условий, и при строгом соблюдении процедуры формирования выборки эту величину можно вычислить. Без учета ошибки выборки полученные результаты можно относить **только** к тем людям, которые непосредственно были опрошены. Понятно, что проектируя выборку, исследователь стремится снизить величину ошибки. Для этого ему необходимо ясно представлять себе те проблемы, которые возникают на пути отбора респондентов, и знать различные способы, позволяющие снизить величину ошибки, а также избежать погрешностей, которые приводят к возникновению неучтенных ошибок.

Неучтенные ошибки и погрешности в ее решении обычно приводят к искажению информации в целом. Ярким примером влияния особенностей выборки на полученную в результате опроса информацию может служить публикация в газете «Советская Россия» (3 октября, 1992) — «Мы тоже провели опрос».

«Мы тоже провели опрос»

«В последнее время все чаще появляются в полторанинской прессе и на яковлевском телевидении результаты «социологических опросов» населения в различных регионах России. Если им верить, люди прямо-таки в восторге от нынешней власти. Мы с сыном тоже решили сделать опрос.

Выбрали два места:

- 1) На кольце автобусной остановки (опросили 58 человек, кроме детей в возрасте до 15-17 лет)
- 2) В очереди в магазине за молоком и вареной колбасой (опросили 44 человека).

Опрос вели по форме «Известий» (N210 за 22 сентября 1992 г.)

1. «Как Вы живете?»

	Пассажиры автобуса	Очередь в гос. магазин	Известия № 210 от 22.09.
очень хорошо	—	—	1%
хорошо	2%	—	10%
средне	6%	11%	51%
плохо	38%	61%	28%
очень плохо	54%	28%	9%

2. «Экономические реформы сейчас нужно продолжать или их следует прекратить?»

	Пассажиры автобуса	Очередь в гос. магазин	Известия № 210 от 22.09.
нужно продолжать	8%	6%	53%
следует прекратить	64%	59%	20%
затрудняюсь ответить	28%	35%	27%

Л.В.Ковалев, ветеран труда, 61 г.,

С.Л.Ковалев, студент, 20 лет

Калининград.

Следует обратить внимание на то, что социологи-любители привели достаточно подробную характеристику контингента людей, которых они опросили (что, к сожалению, далеко не всегда делают даже профессиональные социологи). Уже одно это позволяет установить источники искажения полученной таким способом информации.

Эта публикация самодеятельных «социологов», несмотря на очевидную наивность и выраженную политическую тенденциозность авторов, представляет собой удобный пример, иллюстрирующий влияние выборки на содержательное значение результатов проведенных опросов. Сравнение данных двух «исследований» (газеты «Известия» и читателей Ковалевых) позволяет убедиться, что в результате опросов, проведенных по одной и той же анкете, можно получить принципиально различающиеся цифры, приводящие к прямо противоположным выводам. Наверное, ни у кого не вызывает сомнения, что основной причиной расхождения результатов в данном случае является принцип и способы отбора респондентов — конкретных людей, которые отвечали на анкету.

Так, кого же надо опрашивать, чтобы получить реальную картину общественного мнения, как их отобрать? Эти вопросы являются ключевыми при построении выборки. Квалификация социолога в значительной мере определяется его способностью сформировать и обосновать выборку.

Надеемся, что изложенные ниже основные требования к формированию выборки позволят читателю не только определить, какие именно ошибки и погрешности допустили дилетанты в приведенном выше примере, но и научиться самому избегать подобных ошибок в организации выборочных исследований.

Понятие «выборка»

В настоящее время в эмпирической социологии массовых опросов термин «выборка» имеет два основных достаточно устоявшихся значения.

1. В одних случаях понятие «выборка» используется в качестве синонима понятия «выборочная совокупность» — часть населения (популяции), которая, собственно, и подвергается опросу, с тем чтобы полученные результаты могли быть распространены на изучаемую популяцию в целом.

2. В других случаях под «выборкой» подразумевают процесс отбора единиц наблюдения (технология, способы и методы такого отбора определяют тип выборки) — в данных случаях слово «выборка» является синонимом процедурной категории — «отбор единиц наблюдения».

Построение выборки при подготовке проекта конкретного социологического исследования сводится к решению трех основных проблем:

- 1) определение объема выборки — количества человек, которых следует опросить для получения качественной информации;
- 2) определение типа выборки — построение конкретной схемы процедуры отбора;
- 3) оценка качества выборки — определение, с какой вероятностью и степенью точности результаты опроса выборочной совокупности можно будет переносить на ту или иную часть населения.

Проблемы перечислены в том порядке, в котором они чаще всего возникают перед разработчиками проекта. Но решаться они должны в обратной последовательности. Проблема качества выборки является ключевой — с нее начинается разработка выборки (так как именно она определяет подходы к решению вопроса об объеме и типе выборки), и ответ на этот вопрос после окончания сбора данных ложится в основу анализа полученной информации и интерпретации результатов.

2. Качество информации

Определение качества выборки включает описание ряда условий, позволяющих оценить **валидность, надежность и репрезентативность** выборочной совокупности по отношению к объекту исследования.

Репрезентативность

Репрезентативностью называется свойство выборочной совокупности воспроизводить характеристики генеральной. Отклонения выборочной совокупности от генеральной по каким-либо основным характеристикам называется ошибкой репрезентативности: чем больше величина этих отклонений, тем больше ошибка репрезентативности. Понятно, что выборка не может буквально воспроизвести генеральную совокупность. Основная задача исследователя — учитывать ошибку репрезентативности при интерпретации и обобщении результатов опроса, проведенного на основе выборочной совокупности. Поэтому исследователь должен уметь количественно определять эту ошибку, величина которой рассчитывается в зависимости от типа выборки.

Репрезентативность — необходимое требование к выборке в тех случаях, когда изучается общественное мнение. Однако при проведении многих специальных социологических исследований, направленных на углубленный анализ каких-либо свойств социального объекта, нет необходимости в репрезентативной выборке. Исследователю нередко достаточно набрать основную и контрольную группы людей и провести сравнительный анализ. Например, если основная цель исследования заключается в установлении особенностей личностных качеств предпринимателей и мотивационных побуждений к предпринимательской деятельности, нет никакой необходимости в формировании репрезентативной выборки: достаточно набрать группу лиц, занимающихся предпринимательской деятельностью, а также контрольную группу — людей, не занимающихся предпринимательством. Такой — **аналитический** — подход применим в большинстве случаев, когда исследование направлено на изучение социально-профессиональной специфики предмета.

При аналитическом подходе к выборке следует учитывать следующие необходимые требования: 1) группы должны различаться по основному дифференцирующему признаку (гипотетическому фактору); 2) по всем остальным параметрам (возрасту, полу, уровню образования, характеру деятельности и т.д.) группы должны быть максимально идентичны. Если исследователю удастся набрать достаточно идентичные группы, различающиеся лишь по одному дифференцирующему признаку, то численность таких групп может ограничиваться небольшим объемом (50-100 человек).

Если гипотетических факторов больше одного, то в рамках каждой из групп (основной и контрольной) можно набирать равночисленные подгруппы (например, мужчин и женщин; молодого, среднего и старшего возраста и т.п.). Введение каждого нового фактора требует соответствующего увеличения каждой из групп. Например, при вводе такого признака, как пол, группу предпринимателей (которая первоначально планировалась численностью, допустим, в 50 человек) следует увеличить до 100 (50 мужчин и 50 женщин); точно так же следует увеличить и контрольную группу. Если автор проекта предполагает, что существенным дифференцирующим параметром может являться возраст, и выделяет, предположим, три типа возрастной градации («молодые», «среднего возраста» и «старшего возраста»), то соответственно утраивается и каждая из групп таким образом, чтобы каждая из образованных подгрупп (например, «предприниматели мужчины молодого возраста» и т.д.) была не менее 50 человек. Чем больше по численности будет каждая набранная подгруппа, тем выше вероятность того, что удастся статистически обосновать гипотезы исследования.

Надежность и валидность

Надежность и валидность — основные параметры определения качества информации, полученной в результате проведения эмпирического исследования. Под надежностью подразумевается определенная гарантия, что полученный результат правильно отражает изучаемую действительность. Валидностью (обоснованностью) информации называется подтверждение (доказательство), что исследовались (измерялись) именно те явления, которые социолог предполагал исследовать.

На качестве информации (надежности и валидности данных) сказываются особенности различных этапов эмпирического исследования (исходные теоретические предпосылки, качество инструментария, организация контроля работы интервьюеров и т.д.). Особую роль в обеспечении надежности и валидности данных играют особенности формирования

выборки. В этом разделе мы остановимся на проблеме влияния особенностей выборочного подхода на надежность и валидность получаемой в результате исследования информации.

Надежность информации на этапе формирования выборки обеспечивается учетом **случайных ошибок**.

Случайные ошибки выборки неизбежны вследствие **неоднородности** исследуемого контингента. Качество исследования определяется способностью социолога вычислить величину случайной ошибки и учесть ее при распространении выводов, полученных на основе опроса части определенной группы людей, на всю ее совокупность.

Если бы все люди по всем своим характеристикам были совершенно идентичными, то достаточно было опросить одного человека для вывода, что именно так считают (или, по крайней мере, ответили бы) все остальные. Но социолог именно потому и проводит опрос, что предполагает у обследуемых людей разные точки зрения.

Характеристик, по которым люди различаются, очень много. Часть таких характеристик формализована и фиксируется в определенных документах (пол, возраст, образование, семейное положение, национальность, специальность, должность, заработная плата и т.д.). Другая, значительно большая часть характеристик (личностные особенности, отношение к различным социальным явлениям, мнения и т.п.) не формализована и не зафиксирована в каких-либо документах.

Систематизацией и обобщением формализованных характеристик населения занимается статистика (государственная и ведомственная). Исследование же социолога направлено на то, чтобы формализовать и систематизировать информацию о тех характеристиках людей, которые исследователю представляются наиболее важными с социальной точки зрения, но о них нет сведений в документальных источниках.

Фактически, каждый пункт анкеты является определенной характеристикой (признаком), и исследователь хочет установить, как и в какой степени люди различаются по этой характеристике. Другими словами, исследователь с самого начала исходит из принципа неоднородности обследуемого контингента.

Когда исследователь отбирает часть людей из их общей совокупности для обследования (чтобы по избранной части делать вывод обо всех), то в его выборку могут не попасть люди с какими-либо значениями исследуемых характеристик.

Менее вероятно, что в выборку попадут люди по какой-либо характеристике «нетипичные». Со статистической точки зрения это означает, что их удельный вес в общей совокупности невысок. Чем меньше удельный вес лиц с определенным значением характеристики в общей совокупности, тем меньше вероятность того, что они попадут в выборочное обследование.

Другой особенностью выборки является то, что в выборочную совокупность могут попасть лица со всеми возможными значениями характеристики, но в выборке они будут представлены не в той пропорции, которая характерна для общей совокупности. Например, в общей совокупности 13% лиц, имеющих высшее образование, а в выборку попадает 37%; или в общей совокупности 25% лиц, придерживающихся коммунистической ориентации, а в выборку попадет только 4% и т.п.

Если из общей совокупности отбирать некоторую ее часть **случайным** образом, то на основе теории вероятности можно определить величину, на которую может отклоняться значение признака в выборке по отношению к данному значению совокупности в целом. Эта величина называется **случайной ошибкой** или **ошибкой выборки**.

Величина случайной ошибки зависит от двух основных параметров: от объема выборки и степени изменчивости признака (однородности контингента по какой-либо характеристике).

Если исследователь хочет обеспечить представительство в выборке людей, которые по своим характеристикам нетипичны (не распространены в данной совокупности), он должен увеличивать объем выборки.

Таким образом, чем более однородный контингент подлежит обследованию и чем больше объем выборочной совокупности, тем меньше случайная ошибка выборки и, соответственно, выше надежность полученных в результате опроса данных.

Основным эмпирическим критерием надежности является **устойчивость** результатов, характеризующая свойство получать данные с одним и тем же результатом в разных опросах, проведенных в одинаковых условиях. Например, если бы исследователь повторил свой опрос по той же анкете, примерно в то же время и по аналогичной (по принципам, процедурам и объему) выборке, но на других людях, и получил такие же данные, как и в первом опросе, то можно считать, что результаты его опроса достаточно надежны.

Конечно, случайные ошибки, снижающие устойчивость данных, могут быть результатом и других особенностей исследования: инструментария или работы интервьюеров. Например, если в вопросах анкеты используются термины, совершенно не знакомые многим респондентам, и они будут скорее «угадывать» их значение, то их ответы в этом случае во многом будут лишь «случайной догадкой», и надежность данных по этому пункту анкеты будет снижена.

При тестировании на устойчивость исследователь может делать ориентировочные предположения — привели ли к снижению надежности погрешности выборки или погрешности инструментария: если данные неустойчивы по нескольким (определенным) пунктам анкеты, то, вероятно, к снижению надежности привели погрешности инструментария; если данные неустойчивы по всем пунктам анкеты, то искажение информации происходит скорее всего из-за случайных ошибок выборки. Чем больше объем выборки, тем в меньшей степени случайные ошибки сказываются на общем результате исследования, так как основное свойство случайных ошибок заключается в том, что они «погашают» друг друга.

Валидность (обоснованность) данных, как отмечалось выше, характеризует, в какой степени было измерено именно то, что предполагалось измерить. К снижению валидности могут привести не только ошибки инструментария (о чем шла речь в соответствующем разделе), но и погрешности выборки, в частности — **систематическая ошибка** или **отклонение выборки**. Например, если отбор респондентов будет производиться в районе расположения воинской части, то можно предположить, что в выборке военнослужащих может быть больше, чем это характерно для генеральной совокупности. При повторном исследовании данные могут быть достаточно устойчивыми, но как в первом, так и во втором случае результаты будут подвержены влиянию данного фактора в большей степени, чем его испытывает генеральная совокупность. То же, вернувшись к

приведенному выше примеру, можно сказать и о таких местах опроса, как очередь (в рабочее время суток в очереди значительно больше представителей неработающих категорий населения, чем в структуре населения в целом) или остановка автобуса. Необоснованный выбор точек опроса чаще всего приводит именно к систематической ошибке.

Следует заметить, что надежность и валидность — довольно независимые друг от друга параметры, поэтому при оценке качества информации исследователь должен учитывать обе эти характеристики; данные могут быть валидными, но ненадежными, и наоборот, исследователь может получать достаточно устойчивые (надежные) результаты, но они не валидны по отношению к предмету анализа.

Наглядной иллюстрацией характеристик надежности и валидности, обеспечиваемых особенностями выборки, является аналогия с мишенью, которую приводят в своей работе американские авторы P. Alreck, R. Settle. На рисунке 1 представлены четыре типа результатов выстрелов по мишени. В качестве центра каждой мишени выступает среднее реальное значение некоторого признака (например, средняя заработная плата) изучаемой генеральной совокупности; точками на рисунке отмечены значения полученных результатов повторных исследований по подготовленным выборочным совокупностям.

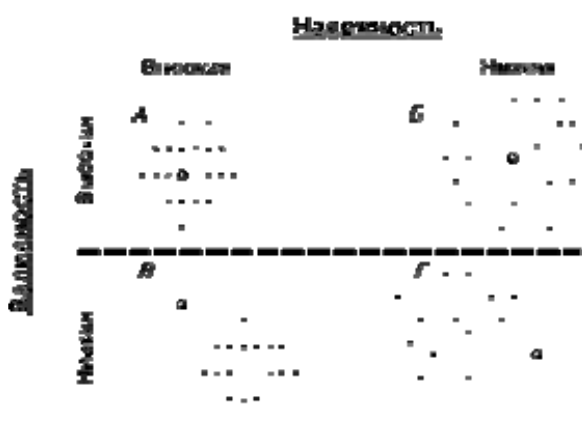


Рис. 1. Иллюстрация надежности и валидности результатов

В квадрате А результаты сгущены вокруг центра (истинного значения измеряемого признака). При каждом повторном исследовании полученная по результатам опроса выборки величина незначительно отличается от истинного значения признака. В данном случае можно говорить о высокой надежности и валидности выборки. Чем меньше общая площадь рассеяния, тем меньше величина случайных ошибок и, соответственно, выше надежность. Чем ближе эти значения к истинному результату, тем выше валидность (обоснованность) данных (исследователь измерял именно то, что предполагал измерять).

В квадрате Б точки также группируются вокруг истинного значения, но результаты повторных измерений значительно отстоят друг от друга. В тех случаях, когда проводится только одно исследование (а чаще всего автор проводит одно исследование), результат, полученный на основе опроса выборочной совокупности, может быть достаточно отдален от истинного значения (из-за влияния случайных ошибок). На практике такой тип ошибок чаще всего возникает в случаях недостаточного объема выборки. Квадрат Б иллюстрирует

результаты исследования, характеризующиеся высокой валидностью, но низкой надежностью.

В квадрате В представлены результаты, характеризующиеся высокой надежностью и низкой валидностью. Исследователь при повторных опросах будет получать достаточно схожие данные, но они не отражают истинную картину характеристики той совокупности, о которой делаются выводы. Например, делаются выводы о средней заработной плате всего населения, а опрашивались респонденты в магазине в рабочее время, когда там выше доля людей неработающих, доход которых в среднем значительно ниже, чем у населения в целом; или опрос проводится в районе какого-либо предприятия, средняя зарплата работников которого в среднем выше, чем у населения в целом, о котором исследователь делает выводы. Систематическая ошибка приводит к искаженному, но достаточно устойчивому результату.

И наконец, квадрат Г иллюстрирует результаты исследования с низкой валидностью и низкой надежностью: результаты исследования претерпевают влияние как случайных, так и систематических ошибок.

Чтобы избежать влияния случайных и систематических ошибок (или, по крайней мере, вычислить их величину и учесть ее при интерпретации результатов), необходимо знать основные характеристики генеральной совокупности, на которую исследователь будет распространять выводы, полученные в результате опроса выборочной совокупности.

В принципе, как остроумно заметил социолог В.Шляпентох: «В этом мире нет нерепрезентативных данных — надо только восстановить ту генеральную совокупность, которую эти данные представляют» (В.Шляпентох. Открывая Америку, 1989, с.74). Другими словами, результаты любого опроса позволяют получить определенную информацию; если исследователь достаточно корректно интерпретирует результаты опроса и делает выводы только относительно той части населения, которую репрезентирует его выборка.

Однако, когда перед социологом стоит достаточно конкретная цель, своей постановкой определяющая объект исследования, модель его выборки по своим основным параметрам должна соответствовать модели генеральной совокупности. Поэтому при формировании выборки он обязан исходить из поставленной цели.

Главная задача исследователя на этапе составления выборки правильно построить модель выборочной совокупности (так, чтобы по своей структуре она соответствовала генеральной совокупности), а на этапе интерпретации результатов — учесть все погрешности, которые были допущены при формировании выборки. Качество информации в конечном итоге определяется не отсутствием ошибок и погрешностей (избежать которых практически невозможно), а создаваемыми исследователем условиями, позволяющими учесть влияние этих ошибок на конечный результат. Более строго это положение формулирует В.Паниотто: «Надежной называют информацию, в которой, во-первых, учтенные ошибки не превышают некоторой заданной исследователем величины; во-вторых, отсутствуют неучтенные ошибки, т.е. ошибки, величину которых исследователь не в состоянии оценить» (Социологический справочник, 1990, с.140).

Решению проблемы учета ошибок построения выборки может помочь технологическая схема перехода от определения объекта исследования (анализировать который предполагает автор на основании данных эмпирического исследования) к определению выборочной совокупности (объема и способов отбора).

3. Переход от объекта исследования к единицам наблюдения (опроса)

Технологическая схема

Переход от объекта исследования к выборочной совокупности в конкретном социологическом исследовании целесообразно осуществлять по нижеприведенной схеме.

1. Объект исследования

2. Генеральная совокупность

3. Основа выборки

4. Единица отбора

5. Единица наблюдения

Объект исследования, генеральная совокупность, основа выборки и выборочная совокупность (совокупность единиц наблюдения) — взаимосвязанные категории, представляющие собой ступени конкретизации объекта исследования, позволяющие обоснованно определить степень репрезентативности выборочной совокупности к заявленному в проекте объекту исследования, с осознанием всех препятствий на пути от определения интересующего объекта исследования к эмпирической проверке. Осознанный и описанный переход от объекта исследования к единицам наблюдения позволяет автору проекта увидеть различные стороны ограниченности возможностей эмпирической проверки выдвинутых положений и гипотез, а также вынуждает исследователя быть более корректным в своих выводах, основанных на результатах опроса.

На первый взгляд может показаться, что подготовка такой схемы затрудняет работу исследователя, увеличивая объем его работы. Но на самом деле это не так. Если социолог принимает решение, кого же в конечном счете он будет опрашивать, с той или иной степенью осознанности всех своих логических шагов, он так или иначе проходит этот путь. Данная схема позволит ему пройти этот путь более рационально и компетентно, и, главное, своевременно — на этапе подготовки к исследованию. Тогда в процессе полевого этапа и этапа анализа данных ему не придется возвращаться к этим же вопросам.

Основная ошибка начинающих социологов заключается обычно в том, что они приступают к определению объема выборки и процедуры (типа выборки), не определив генеральную совокупность и не установив возможностей подготовки основы выборки. Поскольку без этих обойтись в исследовании невозможно, то неквалифицированный социолог, почувствовав собственную беспомощность, нередко берет за основу процедуру, с которой он ознакомился в каком-либо исследовании, где ему доводилось участвовать в качестве рядового исполнителя. Но даже в тех случаях, когда такой перенос допустим, при анализе и представлении полученных данных неквалифицированный социолог все равно проявит свою некомпетентность, поскольку он не понимает возможностей и лимитов, которые определяет данный тип выборки.

Для освоения процедуры подготовки, которая, напомним, является основой оценивания качества выборки и, в конечном итоге, качества информации, определим все вышеуказанные категории («Объект исследования», «Генеральная совокупность», «Основа выборки», «Выборочная совокупность», «Единица отбора», «Единица наблюдения»), являющиеся последовательными ступенями формирования выборочной совокупности; охарактеризуем их взаимосвязь и разберем особенности перехода со ступени на ступень.

Объект исследования

Объект исследования. Под объектом исследования обычно подразумевается совокупность людей, обладающих определенными свойствами и качествами, которые интересуют заказчика и исследователя. Чтобы суть этого определения стала понятнее, поясним это определение примерами из практики социологических исследований. Например, когда в средствах массовой информации сообщают о каких-либо результатах опросов общественного мнения, обычно начинают со слов: «Социологи изучали отношение населения к экономическим реформам...» или «Как показали результаты опросов общественного мнения, уровень доверия населения президенту...» и т.п. При этом не уточняется, население каких именно регионов опрашивалось, не уточняются возрастные и социальные характеристики опрашиваемых, так как подразумевается, что и потребителя информации, и исследователя интересует именно мнение населения в целом.

В качестве объекта исследования подразумеваются все потенциальные носители изучаемого свойства, независимо от того, был ли у них шанс попасть в конкретное эмпирическое исследование. Объект исследования определяется поставленной проблемой и теоретически охватывает всю совокупность людей, которых эта проблема касается, так как данные эмпирического исследования, по крайней мере косвенно, но экстраполируются и на них. Объектом исследования не обязательно может выступать население в целом; это могут быть определенные социальные группы, которых касается изучаемая проблема. Объектом исследования могут выступать определенные социально-демографические группы населения («молодежь», «женщины», «пенсионеры»); социально-профессиональные группы, если именно с ними связана изучаемая проблема (например, «предприниматели», «рабочие», «преподаватели», «журналисты», «деятели искусства», «работники завода N» и т.д.). Объект исследования может определяться социальными ролями и социальным включением — «солдатские матери», «пострадавшие от сталинских репрессий», «пострадавшие от Чернобыльской катастрофы», «читатели газеты «Киевские ведомости», «телезрители», «мафиозные структуры», «одинокие» и т.д.; социальным статусом — «элита», «представители властных структур», «директорат» и т.д. Объект исследования может быть определен национально-этническими характеристиками —

«украинцы», «русские», «евреи» и т.д.; более того, объектом исследования может выступать определенная часть этнической группы, объединяемая общей социальной проблемой, например, «украинская диаспора», «крымские татары, пережившие депортацию», «евреи, собирающиеся выехать в Израиль,» и т.п.

Общей особенностью формулировки объекта исследования является то, что в его определение априорно включаются все лица, которых, по мнению исследователя (или потребителя социальной информации), касается анализируемая проблема. На этапе формулировки объекта исследования разработчика проекта не интересуют возможности эмпирической проверки анализа проблемы. Самое главное для него установить, какие именно контингенты населения являются носителями информации о сути изучаемого явления. Поэтому объект исследования обычно не конкретизируется формальными характеристиками (территориальными, возрастными и т.п.)

В практике освещения полученных данных (средствах массовой информации, публичных и даже научных выступлениях) такие категории населения, как «молодежь», «население», «рабочие» и т.п., обычно фигурируют без уточнения критериев отбора в эту популяцию. И это не является оплошностью авторов выступлений. Просто речь идет о том, что рассматриваемая проблема касается всех представителей данной популяции.

Объект исследования следует рассматривать в качестве базовой категории теоретического **анализа** изучаемой проблемы. Но чтобы провести эмпирическое исследование, объект необходимо операционализировать — сузить таким образом, чтобы, с одной стороны, данные, полученные в результате опроса, можно было в определенной мере отнести к объекту в целом, а с другой стороны, — определить изучаемый контингент так, чтобы его возможно было охватить в конкретном исследовании, которое всегда лимитировано временем, средствами, организационными возможностями и другими факторами.

Следовательно, начальной стадией работы над выборкой является точная формулировка объекта исследования. В заказном исследовании, когда социолог только приступил к разработке проблемы, он не всегда может точно определить объект исследования, так как просто не осведомлен — кого же данная проблема касается. В этих случаях с целью упрощения постановки вопроса он может уточнить у заказчика: «Чье конкретно мнение (отношение и т.п.) вас интересует в целом?» Отредактированный ответ на этот вопрос может быть положен в основу определения объекта исследования. Для уточнения (правильно ли поняли друг друга социолог и заказчик) целесообразно в качестве рабочего подзаголовка название темы сформулировать как сочетание проблемы и объекта (например, «Социально-психологические последствия Чернобыльской катастрофы» — название проблемы; рабочее название может быть сформулировано в следующем виде: «Социально-психологические последствия Чернобыльской катастрофы у лиц, эвакуированных из 30-тикилометровой зоны», или «...у ликвидаторов последствий аварии на Чернобыльской АЭС» и т.п.).

Формулировка объекта исследования в основном названии проекта исследования может как содержаться, так и отсутствовать. Например, тема проекта может называться «Отношение к атомной энергетике», а может называться «Отношение к атомной энергетике населения, проживающего в районе АЭС». Несомненно, что в тех случаях, когда формулировка объекта исследования содержится в названии проекта, вероятность более высокого уровня корректности анализа значительно выше. Но иногда введение объекта в формулировку темы проекта приводит к слишком длинному и громоздкому названию; в таком случае лучше официальное название темы сократить, а при разработке проекта исследования, прежде чем перейти к формированию выборки, следует уточнить

определение объекта исследования. Например, проект называется «Отношение к атомной энергетике». Исследователь уточняет «Отношение населения к атомной энергетике». Заказчик возражает: «Но нам требуется выяснить отношение не только населения, но и лиц, которые могут повлиять на принятие решений о закрытии строительства АЭС». Следовательно, в данном случае речь идет, по крайней мере, о двух объектах исследования, поскольку население и представители управленческих структур, влияющие на решения, касающиеся АЭС, являются носителями различных проявлений и свойств предмета исследования (в данном примере — «отношения к атомной энергетике»). В подобных случаях, конечно, не стоит утяжелять название проекта перечислением всех объектов исследования. Однако в рабочем плане следует четко сформулировать определения всех объектов. Формулировка названия объекта исследования является основой определения генеральной совокупности — второй ступени на пути подготовки выборки.

Генеральная совокупность

Генеральная совокупность — это та часть объекта исследования, на которую социолог с **определяемой** степенью точности сможет распространять выводы, полученные в результате обследования некоторой ее части. Определение генеральной совокупности — это поиск компромисса между теоретически определяемым объектом исследования и ограниченными возможностями эмпирической проверки гипотез.

Переход к определению генеральной совокупности — это конкретизация характеристик объекта, выделение из их бесчисленного множества, во-первых, наиболее важных с точки зрения исследователя; во-вторых, поддающихся фиксации. На этой ступени, социолог начинает осознавать ограниченность возможностей эмпирической проверки некоторых положений и гипотез, возникших в результате теоретического анализа изучаемой проблемы. Для иллюстрации проблем, которые возникают на этом пути, приведем следующий пример. По заказу газеты социолог должен изучить структуру читателей, их интерес к различным материалам, политические ориентации и т.п. Объектом исследования в данном случае выступают «читатели газеты». При определении генеральной совокупности возникает ряд вопросов. По каким критериям определить контингент генеральной совокупности? Где можно получить списки всех единиц генеральной совокупности? Но возможные ответы на эти вопросы приводят к возникновению новых вопросов. Допустим, к читателям газеты можно отнести всех подписчиков. В таком случае следует учитывать только основного подписчика или всех членов семьи? Относить ли к читателям газеты только тех, кто ее читает постоянно, или можно отнести и тех, кто читает данную газету «время от времени»? Тех, кто читает все материалы, или достаточно, что человек просматривает газету и читает только то, что его заинтересовало? Следует ли учитывать тех, кто покупает газету в розницу? Есть ли у газеты постоянные читатели, которые не подписались на нее, но регулярно получают через другие каналы, много ли таких читателей и возможно ли в таком случае их как-то зафиксировать?

Можно видеть, что в данном примере достаточно сложно перейти от объекта исследования, определение которого достаточно понятно в обычном житейском рассуждении, к генеральной совокупности, полностью представляющей объект исследования. В любом случае приходится сужать объект исследования до совокупности людей, параметры которых поддаются формализации, с тем чтобы у них был шанс

попасть в выборочный опрос. Например, это могут быть подписчики газеты (по одному представителю из каждой семьи) и лица, покупающие газету в розницу.

Осознавая ограниченность такого определения генеральной совокупности по отношению к объекту исследования, автор проекта все же имеет определенные преимущества перед тем социологом, который не пытался осуществить переход от объекта исследования к определению генеральной совокупности. Его преимущества заключаются, во-первых, в том, что при анализе полученных данных он будет осознавать, в какой мере выводы его исследования соотносятся с объектом, и соответственно корректировать категоричность тех или иных выводов, а во-вторых, при определении генеральной совокупности он продумывает источники получения информации для составления списка генеральной совокупности — основы выборки, а также — пути доступа к этим источникам. В данном случае это могут быть списки адресов подписчиков, которые находятся в почтовых отделениях связи, а также покупатели у газетных киосков (зафиксированные в определенном временном интервале, например, в течение дня или недели). Данный пример иллюстрирует достаточно сложный случай идентификации генеральной совокупности по отношению к объекту исследования. Значительно чаще в социологических опросах объект исследования включает представителей определенных регионов или социальных слоев населения. Но и здесь требуются определенные усилия для перехода от объекта исследования к определению генеральной совокупности. Например, при изучении тех или иных проблем молодежи объектом исследования выступает, как несложно догадаться, молодежь. Но при определении генеральной совокупности требуется ряд уточнений. В первую очередь это касается возрастных границ. Но даже определив возрастные границы, исследователь должен уточнить — касается ли изучаемая проблема всей молодежи страны или только каких-либо определенных контингентов.

Как правило, конкретизацию генеральной совокупности рационально проводить по следующей схеме:

1. Регион

2. Социально-демографическая характеристика

3. Отраслевая специфика.

Например, необходимо изучить отношение молодежи к преступности. Объектом исследования выступает молодежь. Для определения генеральной совокупности сначала уточняется региональная локализация населения, на которое можно будет распространять выводы, полученные в результате исследования: идет ли речь о молодежи страны в целом или какого-либо региона. Затем уточняется, будет ли изучаться только городская или сельская молодежь, и если, допустим, городская, то только ли молодежь крупного города или вся городская молодежь. Затем (в рамках социально-демографических уточнений) определяется верхняя и нижняя границы молодежного возраста (например, от 18 до 30 лет или от 14 до 25 лет — конкретные границы зависят от концептуального подхода, целей и информационных потребностей); и наконец, уточняется отраслевой контингент — будет ли изучаться мнение всей молодежи, проживающей в данном регионе, или только

учащейся или работающей молодежи (если работающей, то во всех отраслях или в каких-то определенных) и т.д. Итогом этой работы должно явиться четкое определение генеральной совокупности, например, «учащаяся молодежь в возрасте 18-30 лет западного региона Украины (Львовская, Ивано-Франковская и Тернопольская области)».

Своевременно сформулированное определение генеральной совокупности позволяет заблаговременно установить источники, содержащие основные характеристики генеральной совокупности (общая численность, распределения по основным характеристикам), и пути их получения. Следующей ступенью работы над проектированием выборки является подготовка основы выборки.

Основа выборки

Под основой выборки понимается перечень всех элементов генеральной совокупности, из которого, собственно, и производится отбор единиц наблюдения.

Существует два типа основы выборки.

1. Список (перечень) единиц отбора с указанием адресов и основных характеристик. Но нередко подготовка такого списка крайне затруднена или вообще невозможна. В таких случаях используется второй тип основы выборки.
2. Полный перечень всех характеристик, на основании которых можно определять, подходит тот или иной объект под определение генеральной совокупности. В таком случае основа выборки очень близка к определению генеральной совокупности. Она отличается от генеральной совокупности двумя особенностями: а) характеристик больше, и они обязательно включают все границы; б) в основе выборки указываются единицы отбора, которые должны быть исключены из генеральной совокупности по тем или иным причинам.

Например, генеральная совокупность определяется как «все избиратели страны (округа, населенного пункта и т.п.)». Такое определение подразумевает возрастную границу, гражданство и дееспособность. Но при подготовке основы выборки целесообразно заблаговременно (до начала исследования) исключить определенные контингенты лиц, адекватное участие в исследовании которых заранее можно подвергнуть обоснованному сомнению. В число таких контингентов исследователь может отнести, например, людей, имеющих физические (плохой слух, тяжелая болезнь и т.п.) или психические недостатки, которые, как можно предвидеть, не позволят им выступать источниками адекватной информации. Из основы выборки могут быть также исключены лица, находящиеся в предварительном заключении; формально они не лишены избирательных прав, но они априорно являются слишком труднодоступными единицами наблюдения в исследовании, цель которого не сосредоточена на изучении именно данного контингента. Исследователь может ввести ограничение и верхней границы возраста (хотя избирательным законодательством это и не предусмотрено). Из основы выборки могут быть исключены закрытые военные предприятия и т.д. Перед исследователем будет стоять проблема, включать ли в основу выборки избирателей, которые находятся в длительной командировке, и, наоборот, тех, кто в момент опроса будет находиться по предусмотренному выборкой адресу, но как избиратель зарегистрирован в другом месте и т.д.

Общая задача исследователя на этапе подготовки основы выборки — составить как можно более полный перечень, на основе которого каждую потенциальную единицу отбора непосредственные исполнители отбора единиц наблюдения смогут без колебаний и субъективных предпочтений идентифицировать — подходит ли она под генеральную совокупность или нет.

Подготовка основы выборки первого типа — составление полного списка единиц отбора, если такая возможность у исследователя есть, обеспечивает более высокий уровень культуры выборки. Как уже отмечалось, такой список должен включать перечень всех единиц отбора, их адреса и основные характеристики.

Какие характеристики-индикаторы объекта должны быть включены в основу выборки? В первую очередь это определяется поставленной проблемой и исследовательскими гипотезами. Те параметры объекта, которые в соответствии с гипотезами исследователя должны влиять на изучаемое явление, выступают в качестве основных характеристик-индикаторов; именно по этим характеристикам выборочная совокупность должна репрезентировать генеральную.

В том случае, когда у исследователя нет четких и достаточно обоснованных гипотез (исследование носит поисковый характер), можно обходиться стандартным минимумом основных характеристик. Для различных единиц отбора существуют различные стандартные наборы-минимумы характеристик-индикаторов. Не останавливаясь здесь на определении единиц отбора (так как этот вопрос будет изложен чуть ниже), перечислим наиболее распространенные стандартные наборы-минимумы необходимых характеристик.

Если единицами отбора являются люди, то обычно в минимальный набор таких характеристик включаются сведения о поле, возрасте, образовании респондентов.

Если единицами отбора являются семьи, то основу выборки составляет полный перечень всех адресов, а минимальный набор характеристик — тип региона и района проживания.

Если единицей отбора является место работы, то основа выборки список всех предприятий, входящих в генеральную совокупность, а набор характеристик — тип предприятия-учреждения (в основу типологии закладываются характеристики предприятия, которые гипотетически влияют на предмет исследования).

Единицы отбора и единицы наблюдения

Единицей наблюдения в социологических опросах является конкретный человек, который будет отвечать на вопросы анкеты. При одноступенчатой выборке единицы отбора и единицы наблюдения совпадают. (Одноступенчатой является выборка, когда исследователь отбирает единицы наблюдения из общего списка основы выборки). Но одноступенчатая выборка в практике социологических опросов возможна в тех случаях, когда генеральная совокупность небольшая по объему и достаточно плотно локализована. Например, обследуются жители одного района или небольшого населенного пункта, работники одного предприятия или одной отрасли. Такие исследования обычно проводятся при осуществлении социального эксперимента или изучении отношения людей к какому-либо социально-политическому решению, касающемуся жителей небольшой региональной зоны; или предприятие заинтересовано в изучении причин текучести кадров или преодолении возникшего социального конфликта и т.п.

Значительно чаще в социологических исследованиях работники предприятия или жители района выступают в качестве единиц наблюдения не как представители только своего предприятия, а значительно большей социальной группы. В тех случаях, когда генеральная совокупность достаточно большая (ориентировочно — более двух тысяч человек) или информация о ее основных характеристиках рассеяна (не существует общих списков с указанием характеристик), используют многоступенчатую выборку: сначала отбирают регионы, затем населенные пункты, затем предприятия или адреса проживания семей и лишь на последней ступени — конкретные единицы наблюдения. При использовании многоступенчатой выборки производится отбор конкретных объектов, в рамках которых проводится последующий отбор; эти объекты и называются единицами отбора. На последней ступени — единицы отбора и являются единицами наблюдения.

В некоторых учебных пособиях по методике проведения социологических опросов указывается, что единицами наблюдения чаще всего являются люди, но могут выступать и другие социальные объекты: семья (например, при исследовании бюджетов) или трудовой коллектив (например, при исследовании эффективности работы предприятий с различным типом собственности). Но следует заметить, что в данном случае происходит неправомерный перенос единиц анализа, который в дальнейшем будет осуществлять исследователь, на практику выборки. Не следует забывать, что если даже собирается информация о семье в целом (бюджет семьи) или о коллективе (показатели эффективности работы или социально-психологического климата), непосредственным источником информации будут выступать какие-то конкретные представители обследуемой группы. В любом случае интервьюеру требуется конкретная инструкция — к кому он должен обратиться со своей анкетой. Даже если это не имеет значения с точки зрения авторов проекта выборки, интервьюер должен получить конкретную инструкцию («первый встречный на предприятии»; «тот, кто откроет дверь», «тот, кто считает себя наиболее компетентным»). Но такие инструкции крайне редки, потому что, формулируя инструкцию по отбору единицы наблюдения, автор проекта начинает догадываться, что отобранные таким образом источники информации могут существенно исказить истинную картину.

Сколько же человек следует опрашивать и как проводить отбор единиц наблюдений из основы выборки? Этой проблеме посвящена следующая лекция.

Лекция 4.

Выборка: типы и процедуры отбора

В эмпирической социологии развиваются два основных подхода, определяющих тип выборки: вероятностный и целенаправленный (направленный)*.

Прежде чем рассмотреть принципиальные особенности каждого из типов, следует заметить, что обычно в социологическом исследовании редко используется **одноступенчатая** выборка, при которой единицы наблюдения отбираются непосредственно из генеральной совокупности. Одноступенчатые выборки, как уже отмечалось, обычно строятся при локальных исследованиях в тех случаях, когда объем генеральной совокупности невелик, генеральная совокупность достаточно компактно локализована и существует возможность составить общий список всех ее единиц,

содержащий основные характеристики и адреса. В практике социологических исследований обычно прибегают к **многоступенчатым** выборкам, конкретизируя и локализуя на каждой ступени единицу отбора, постепенно «добираясь» до единиц наблюдения — потенциальных респондентов. На каждой из ступеней используются различные процедуры. Но в основе построения каждой из ступеней лежит один из двух основных принципов: вероятностный или целенаправленный (направленный). Какой принцип следует выбирать — решает автор исследования, исходя из преимуществ и недостатков каждого из них. В самом общем виде можно сказать, что вероятностный принцип обладает теоретическими преимуществами, а целенаправленный — практическими. Кроме того, вероятностный отбор лучше осуществлять в тех случаях, когда исследователь имеет возможность составить полный список единиц отбора (и не имеет возможности получить сведения о распределениях по основным характеристикам), а направленный отбор лучше осуществлять, когда исследователь располагает сведениями о характеристиках (например, по данным государственной или ведомственной статистики), а составление (получение) перечня единиц отбора крайне затруднено.

1. Вероятностный подход

Одноступенчатая случайная выборка

Вероятностный подход нередко называется также случайным или стохастическим, поскольку подразумевает **случайный** отбор единиц наблюдения из общего перечня (списка) единиц генеральной совокупности. Основное правило при случайном отборе — **равная вероятность каждой единицы генеральной совокупности попасть в выборку**. При педантичном соблюдении этого условия реализуется основное преимущество случайной выборки: элементы генеральной совокупности оказываются представленными в выборке с вероятностями, которые приближаются к их распределению в генеральной совокупности. При этом, чем меньше признак рассеян в генеральной совокупности и чем больше объем выборочной, тем меньше ошибка репрезентативности. Такой результат является проявлением закона больших чисел, что позволяет при знании дисперсии по фиксируемым характеристикам, заранее определив допустимую ошибку репрезентативности, рассчитать нужный объем выборочной совокупности.

Отработаны различные технические процедуры случайного отбора: механический отбор, использование таблицы случайных чисел, расчет шага отбора и др.

Механический отбор подразумевает составление карточек с номерами, каждый из которых соответствует номеру единиц отбора в общем перечне. Все карточки перемешиваются в барабане и вытаскиваются из него в случайном порядке. В современной практике эта процедура практически не используется, так как требует определенных дополнительных организационных затрат (подготовка карточек (или шариков), обязательно идентичных по своим физическим параметрам, механическое устройство для их равномерного перемешивания и т.д.). Однако именно такая процедура обеспечивает случайную **повторную** выборку. Повторной называется такая выборка, когда выбранный элемент (карточка) вновь возвращается в барабан, и есть определенная вероятность, что она будет выбрана еще раз. Строго говоря, теоретически только такая выборка позволяет рассчитывать при анализе данных статистические коэффициенты связей и значимости. Но на практике, понятно, никто не будет опрашивать одного и того же респондента дважды, даже если его номер дважды попадет в выборку.

Таблицы случайных чисел позволяют с помощью компьютерных программ автоматизировать случайный отбор единиц наблюдения на основе таблицы случайных чисел.

Наиболее часто для осуществления случайного отбора используют расчет шага отбора. Это делается следующим образом. Единицы отбора из списка генеральной совокупности обозначаются с помощью сплошной нумерации. Исследователь должен обратить внимание на то, чтобы в очередности не было никакой закономерности. Общее число единиц генеральной совокупности (допустим, 10 000 единиц) делится на число единиц выборочной совокупности (допустим, 2000 человек); шаг отбора равен 5. Из списка генеральной совокупности отбирается каждый пятый номер. Обычно отбор начинают не с первого номера. Величина шага отбора делится пополам; если эта величина не является целым числом, ее округляют до большего значения. В данном случае следует пять разделить на два, получается 2.5; следовательно, начинать выбор надо с третьего номера; следующий номер отбора будет восьмой ($3+5=8$), следующий — тринадцатый ($8+5=13$) и т.д.

Процедуры случайного отбора достаточно отработаны и с технической точки зрения не представляют для социологов особой трудности. Основные проблемы осуществления одноступенчатого случайного отбора вызваны практическими препятствиями на пути выполнения теоретических требований.

Основная проблема заключается в составлении общего списка генеральной совокупности, в основном, — в получении необходимых сведений для подготовки такого списка.

Второй круг проблем осуществления случайной выборки связан с достижимостью респондентов. Эти проблемы вызваны противоречием между требованием опрашивать именно тех людей, которые попали в случайный отбор (на основе равной возможности каждого попасть в выборку), с одной стороны, и практической невозможностью опросить определенную часть полученной выборочной совокупности — с другой. Обычно недостижимость части выборочной совокупности (до 10-15%) вызвана следующими причинами: отсутствие респондентов (временное или постоянное) по указанным адресам, отказы от участия в опросе и т.д.

Третья проблема одноступенчатой случайной выборки определяется слишком рассеянным (в пространстве) полем респондентов. Если в случайном порядке отобрать 1000 человек, проживающих на территории страны, то полученные адреса будут находиться достаточно далеко друг от друга, и организаторам трудно найти интервьюеров, которые согласились бы опрашивать людей, расположенных таким образом (в лучшем случае эта проблема может решаться за счет повышения оплаты работы интервьюеров и транспортных расходов, что, естественно, приводит к существенному повышению общей стоимости исследования).

Таким образом, требование равной вероятности каждой единицы генеральной совокупности попасть в выборку является самым большим достоинством вероятностного подхода с теоретической точки зрения (поскольку позволяет свести к минимуму случайные и систематические ошибки и определить общую ошибку репрезентативности) и самым большим недостатком — с практической точки зрения (так как в практике массовых опросов это требование часто невыполнимо или настолько трудоемко, что организационные затраты становятся несоразмерными ожидаемым результатам).

Разрешение этого противоречия возможно при осуществлении процедуры многоступенчатого случайного отбора.

При многоступенчатом случайном отборе, в результате предварительного анализа объекта исследования, генеральная совокупность разбивается на подсовокупности — единицы отбора; из их числа на первой ступени в случайном порядке отбирается часть подсовокупностей, и из отобранных подсовокупностей отбираются единицы наблюдения. При необходимости число ступеней может быть увеличено.

Осуществление многоступенчатого случайного отбора отличается двумя особенностями.

1. Ошибка репрезентативности на каждой ступени возрастает; но, главное, что эта ошибка может быть учтена исследователем; неучтенные ошибки сводятся к минимуму.

2. Многоступенчатый подход требует предварительного анализа и систематизации объекта исследования. Под систематизацией подразумевается группировка единиц отбора. Другими словами, все единицы генеральной совокупности требуется предварительно разбить на группы. Следующие параграфы посвящены основным способам предварительной классификации (группировки) объекта исследования, получившим наибольшее распространение в современной практике массовых опросов.

Принцип районирования (стратификации)

Существуют два основных вида классификации (группировки) генеральной совокупности: районирование (стратификация) и кластеризация (метод гнезд, серийная выборка).

Районированная (стратификационная) выборка предполагает предварительную группировку единиц генеральной совокупности по *различающимся* между собой типам.

В основу классификации закладывается признак (или совокупность признаков), по которому респонденты существенно отличаются друг от друга, и при этом именно данный признак, по мнению исследователя, может существенно влиять на изучаемое явление. Таким образом, при выборе критерия стратификации (районирования) исследователь руководствуется двумя особенностями:

- 1) данная характеристика должна влиять на изучаемое явление;
- 2) по данной характеристике единицы совокупности должны различаться.

Например, при изучении отношения к приватизации можно предположить, что тип предприятия, на котором работают респонденты (государственное или частное), будет существенно сказываться на их отношении к приватизации. Исследователь также знает, что доля людей, работающих на частных предприятиях, незначительна, и при механическом случайном отборе вероятность их попадания в выборку достаточно низкая. В таком случае все предприятия разделяются на группы в зависимости от формы собственности, а потом из каждого типа отбираются респонденты пропорционально численности данного контингента в генеральной совокупности. В большинстве крупномасштабных опросов общественного мнения по национальным выборкам первичная стратификация осуществляется по географической локализации респондентов (отсюда и синонимичное название этого принципа — районирование).

В статистическом смысле районирование соответствует выделению такого числа и таких статистически однородных групп, чтобы дисперсия заданных параметров внутри полученных групп была меньше, чем между ними.

Принцип кластеризации («гнезд»)

Кластерный («гнездовой») способ предварительной классификации объекта в определенном смысле противоположен районированному. Если при районировании исследователь выделяет разнотипные подсовокупности, то при гнездовой выборке генеральная совокупность разбивается на однотипные группы («гнезда»), внутри которых содержатся разнородные единицы наблюдения. В качестве таких «гнезд» могут выступать школы, предприятия, семьи, коллективы, почтовые отделения, избирательные участки и т.п. организации, интегрирующие тем или иным образом группы людей, если исследователь считает, что эти кластеры достаточно однотипны.

Используя принцип кластеризации, исследователь должен руководствоваться двумя основными положениями:

- 1) **все** единицы генеральной совокупности должны быть распределены между «гнездами»;
- 2) основные характеристики «гнезд» должны быть максимально идентичными по заданным параметрам.

При использовании гнездовой выборки, в рамках «гнезда» можно осуществлять сплошной опрос, а можно выборочный; это зависит от численности и однородности совокупностей «гнезд».

Многоступенчатая комбинированная выборка

Обычно эти подходы (районированный и гнездовой) в научно-методической литературе противопоставляются друг другу. Но в современной практике проведения массовых крупномасштабных исследований при построении выборки чаще всего используются оба этих принципа. Применяя последовательно на разных ступенях отбора каждый из этих принципов, исследователь добивается повышения репрезентативности выборочной совокупности по различным характеристикам, которые, по мнению исследователя, могут оказывать принципиальное влияние на изучаемое явление.

Приведем несколько примеров подготовки выборки с комментариями, поясняющими соображения, которыми руководствуется исследователь, последовательно реализуя описанные выше принципы на каждой из ступеней отбора.

1. Например, социолог предполагает исследовать профессиональные ориентации старшеклассников, в частности, ориентации на поступление в вуз. Казалось бы, в данном случае на первой ступени целесообразно применить гнездовую выборку, и в качестве «гнезд» рассматривать школы, часть которых следует в случайном порядке отобрать как основу будущей выборочной совокупности. Однако исследователь принимает во внимание следующие обстоятельства: 1) существуют специальные школы (математические, с углубленным изучением иностранных языков и т.п.), контингент

которых в большей степени ориентирован на вуз; 2) таких школ немного в общем перечне школ, и поэтому существует достаточно высокая вероятность, что они не попадут в выборку «гнезд». Учитывая, что школы, контингент которых в большей степени ориентирован на вуз, может оказаться за пределами выборки, что, несомненно, скажется на данных, касающихся профессиональных ориентаций старшеклассников, автор проекта принимает решение — в качестве первого шага произвести районирование (стратификацию) по этому показателю. Специальные школы выписываются в отдельный список (перечень). В дальнейшем «гнезда» (школы) выбираются из каждого списка отдельно, пропорционально численности старшеклассников в каждой из двух полученных подсовокупностей (общие школы, специальные школы).

Естественно, что в основу районирования (стратификации) генеральной совокупности старшеклассников исследователь может заложить и другие критерии, например, тип городского района (индустриальный, культурно-административный, «спальный» и т.п.), если он считает, что школы разных районов не являются достаточно идентичными «гнездами»; и различия между ними могут принципиально сказываться на результатах опроса. В итоге, на первой ступени отбора генеральная совокупность стратифицируется (реализуется принцип районирования), а на второй, когда исследователь считает, что в рамках каждой страты «гнезда» достаточно идентичны с точки зрения целей исследования, производится случайный отбор намеченного числа «гнезд» (школ). В дальнейшем исследователь может в каждой школе проводить сплошной опрос старшеклассников, а может, пользуясь принципом гнездового метода, отобрать классы.

Сколько следует отбирать школ, сколько классов, сколько учеников опрашивать в каждом классе? Решение этих вопросов в каждом исследовании — это поиск компромисса между требованиями теории (учет всех признаков, дифференцирующих единицы наблюдения и отбора, а также дисперсии этих признаков) и практическими возможностями организации исследования. Понятно, что, чем меньше школ, тем легче организовать опрос. Но в таком случае существенное влияние на выборочную совокупность могут оказать особенности школ, не учтенные исследователем.

2. Сочетание районированного (стратификационного) и кластерного («гнездового») подходов обычно используется и при построении территориальных выборок, по которым проводятся массовые опросы общественного мнения. Например, исследователь предполагает провести опрос населения по репрезентативной общенациональной выборке Украины.

Если бы он хотел составить случайную одноступенчатую выборку взрослого населения Украины (старше 18 лет), ему требовалось бы составить список всех избирателей (более 37 млн человек) и с заданной ошибкой репрезентативности отобрать из общего списка, пользуясь таблицей случайных чисел, требуемое количество единиц наблюдения (например, при средней ошибке — 3% требуется опросить 1800 человек). На такую процедуру не идет ни один исследователь. Во-первых, сбор и составление общего перечня единиц наблюдения — процедура очень трудоемкая, наталкивающаяся практически на непреодолимые препятствия. Но даже если ее осуществить, то в результате будет получено очень рассеянное в пространстве поле выборочной совокупности. Практически невозможно будет найти интервьюеров, которые согласятся проводить опрос 10-15 человек, локализованных на значительном расстоянии друг от друга. А если бы такие интервьюеры и нашлись, то существенно возрастают транспортные расходы. Учитывая, что на результате обязательно скажется «недостижимость» части отобранных единиц наблюдения, что в конечном итоге все равно нарушит принцип «равной вероятности попадания каждой единицы генеральной совокупности в выборку», приступать к такой

работе (подготовке одноступенчатой случайной выборки взрослого населения Украины) нецелесообразно.

Поэтому исследователь сразу решает строить многоступенчатую выборку. Теоретически здесь лучше применять гнездовой подход, и в качестве «гнезда» использовать избирательные участки (ЖЭКи, почтовые отделения). Это несколько бы упростило процедуру, по сравнению с одноступенчатой случайной выборкой, но в принципе те же организационные проблемы также стояли бы перед исследователем. Кроме того, он лишается возможности сопоставлять результаты исследования по различным регионам, так как их репрезентация не предусмотрена проектом выборки.

Для оптимизации процесса организации исследования при построении территориальной выборки обычно на первом этапе производят районирование: определяют типы регионов таким образом, чтобы можно было сказать, что население одного типа региона достаточно однородно по какому-либо признаку (признакам), а между населением разных типов существуют различия по этому признаку. Затем каждую административно-политическую единицу (область или район) относят к одному из выделенных типов. В основу типологии берется признак или комплекс признаков, которые, по мнению исследователя, могут существенно влиять на особенности изучаемого явления. Такими критериями типологии могут выступать социально-культурные особенности населения, сложившиеся исторически, особенности этнического состава населения административных подразделений, политические ориентации (выявленные в ходе выборов или референдумов). Таким образом, критерий типологии определяется в основном предметом исследования и компетентностью автора проекта выборки. Критерий типологии должен быть обоснован в проекте выборки. В дальнейшей работе из каждой страты (типа региона) исследователь будет отбирать число единиц наблюдения, пропорциональное численности населения регионов.

Второй шаг обычно также предусматривает дальнейшее районирование, так как исследователь знает, что обычно тип поселения может существенно сказываться на результатах исследования. Эта типология может быть произведена с различной степенью точности: например, можно выделить два типа (городское и сельское население); можно произвести более тонкую дифференциацию, учитывая численность населенных пунктов, их специализацию в народном хозяйстве и т.д. и т.п. Здесь опять исследователь ищет компромиссное решение: чем больше типов и признаков заложено в основу типологизации, тем меньше будет ошибка репрезентативности, но тем больший объем выборочной совокупности придется формировать, чтобы наполнить полученные типы. В рабочем плане выборки исследователь определяет численность выборочной подсовокупности каждого типа населенного пункта.

В дальнейшей процедуре все населенные пункты, распределенные по типам, в рамках каждого типа рассматриваются как равноценные «гнезда», из которых производится отбор необходимого числа населенных пунктов.

Следующая ступень может также осуществляться «гнездовым» способом, где в качестве «гнезда» выступают избирательные участки (жэки, почтовые отделения). Эти организации удобно рассматривать в качестве «гнезд», поскольку, во-первых, они располагают списками единиц наблюдения, а во-вторых, достаточно однородны по своей внутренней структуре. Если исследователь считает, что они не однородны по какому-то важному для исследования качеству, то на третьей ступени он может еще раз произвести районирование (стратификацию) в рамках каждого из населенных пунктов. Из окончательно стратифицированной совокупности кластеров («гнезд») производится их

отбор в выборочную совокупность. Все страты и кластеры на различных ступенях представляют собой единицы отбора.

На последней ступени осуществляется отбор единиц наблюдения. Эта процедура может осуществляться различными техниками случайного отбора. Выше описаны процедуры случайного отбора на основе списка (перечня) единиц наблюдения.

В тех случаях, когда исследователь не располагает списками с описанием необходимых для исследования характеристик, на одной из ступеней отбора (обычно на предпоследней) используется **маршрутный метод**. Процедура использования маршрутного метода достаточно подробно описана в методической литературе (в частности, в работе Э.Ноэль «Массовые опросы», с.154-157). Поэтому мы не будем останавливаться на технике подготовки маршрутного метода. Основной его особенностью является то, что этот метод основан на случайном подходе при отсутствии списка и сведений об основных параметрах генеральной совокупности. Основу выборки составляют адреса, а шаг отбора определяется особенностями домостроений (учитывается число квартир в доме, на лестничной площадке, этажность и т.п.).

Следует только обратить внимание исследователя, что лучше, когда отбор адресов для опроса осуществляет не сам интервьюер, а сотрудник, принимающий участие в организации опроса. Нередко исследователи ограничиваются подготовкой инструкции по осуществлению отбора квартир и не учитывают, что интервьюеру трудно держать в голове все повороты «налево во второй дом по правой стороне», считать в домах входные двери и учитывать, на каком этаже он опрашивал респондента в последний раз, когда его основной целью является контакт с респондентом и качественное проведение опроса. Следует заметить, что некоторые (достаточно солидные) фирмы составление списка адресов возлагают на интервьюера, предоставляя ему карту района. Однако необходимо учитывать, что это существенно повышает оплату не только работы интервьюеров, но и супервайзеров, которые должны контролировать не только качество интервью, но и правильность выполнения инструкций по составлению списка адресов. Чтобы избежать существенного повышения стоимости исследования, с одной стороны, и снижения качества работы, — с другой, лучше всю процедуру, предусмотренную инструкцией, осуществлять заранее, а интервьюеру предоставлять готовый список адресов и инструкцию по отбору респондента в семье.

На последней ступени — отбор единиц наблюдения — также можно использовать принцип случайного отбора респондента из членов семьи, проживающих по установленному адресу. Разработаны различные техники случайного отбора респондента из состава проживающих в квартире членов семьи. Основой большинства из них является список всех членов семьи, который должен составить интервьюер. Обычно интервьюер составляет список членов семьи в следующей последовательности: сначала всех лиц мужского пола в порядке увеличения возраста, начиная с самого младшего члена семьи, достигшего 18 лет, а затем всех лиц женского пола в таком же порядке. Из составленного списка случайно отбирается 1 человек. Для случайного отбора используются либо «кости» (кубик с числовыми значениями от 1 до 6), либо специальные таблицы, которые интервьюер получает от организатора исследования.

3. Примером многоступенчатой случайной выборки является выборка взрослого населения Соединенных Штатов, по которой проводит ежемесячные опросы-интервью (face-to-face) Институт Гэллага. На первой ступени проводится стратификация населенных пунктов: город-село; в свою очередь, городское население стратифицируется по типу города (города, в зависимости от численности населения разделяются на cities и

towns), а сельское население районируется по сегментам застройки. Из страт отбираются 350 «точек опроса» на основе принципа случайного отбора. Каждый интервьюер получает карту-схему «точки опроса» с отмеченной точкой начала «маршрута» и инструкции по установлению домовладений (домов, квартир) в соответствии с принципами маршрутного метода, а также инструкции по отбору респондента в семье (в соответствии с требованиями случайного отбора). Для работы в каждой точке приглашается интервьюер, не проживающий в данной местности. Всего опрашивается около 1000 человек.

Основу выборки составляют домовладельцы, т.е. из генеральной совокупности взрослого населения заведомо исключаются лица, находящиеся в тюрьме, больнице, проживающие в гостиницах или снимающие квартиру, проживающие на военных базах, а также в общежитиях учебных и религиозных заведений.

После завершения опроса, данные по основным демографическим характеристикам опрошенных сопоставляются с данными последней переписи; производится перевзвешивание результатов опроса с учетом данных переписи, что позволяет повысить «репрезентативность» лиц, которых трудно застать дома. Эти процедуры направлены на приведение выборки в соответствие с основными параметрами взрослого населения (старше 18 лет), проживающего в собственных жилищах.

Поскольку построение данной выборки Института Гэллага определяется принципами случайного отбора на каждой из ступеней, это дает возможность исследователям подсчитывать ошибку выборки и в ежемесячных публикациях в качестве приложения к результатам опроса предьявлять таблицу, позволяющую читателям «правильно читать» полученные данные, учитывая случайную ошибку выборки (например, «такого мнения придерживаются 23% «плюс-минус» 4%»).

Таким образом, многоступенчатый случайный отбор позволяет определить ошибку репрезентативности на каждой ступени и, соответственно, общую ошибку репрезентативности. Или наоборот, — на основании заранее заданной ошибки вычислить объем выборочной совокупности. Естественно, следует понимать условность правильности расчетов. Эта условность определяется теми критериями, показатели которых учитывались (а особенно — теми, которые не учитывались) при осуществлении типологии объектов репрезентации.

Если исследователь не вдается в точные математические расчеты при подготовке многоступенчатой случайной выборки, он, по крайней мере, может руководствоваться общим (условным) правилом: учитывать, что на каждой ступени число единиц отбора должно быть достаточно большим. В противном случае принцип случайности теряет свой смысл. Например, если из нескольких сот школ города в случайном порядке отбирать 4-6 школ, вполне понятно, что в их число могут попасть нетипичные объекты. Ориентировочное число единиц отбора должно составлять около 300-400 единиц. Кроме того, не следует забывать, что даже многоступенчатый способ, облегчающий процедуру случайного отбора, не решает проблему «недостижимости» респондентов, попавших в выборку в соответствии с проведенной процедурой случайного отбора. Процедура последующего перевзвешивания результатов обладает как определенными достоинствами, так и недостатками (коррекция информации определяется, во-первых, гипотетическими факторами исследователя, во-вторых, только теми, сведения о которых содержат статистические документы).

Если на определенных ступенях отбора автор проекта предполагает отбирать небольшое число единиц репрезентации, то во многих случаях целесообразнее использовать целенаправленный подход.

В практике эмпирических социологических исследований социологи достаточно часто обращаются к направленному отбору (либо на некоторых, либо на всех ступенях построения выборки). Нарушая определенные формальные требования, исследователь нередко добивается более достоверной информации, что впоследствии подтверждается результатами референдумов или выборов.

2. Целенаправленный (направленный) подход

В основу направленной выборки закладывается не перечень (список) единиц отбора и единиц наблюдения, а статистическое описание объекта исследования по наиболее значимым, с точки зрения их влияния на исследуемое явление, характеристикам генеральной совокупности. В качестве источников информации по описанию объекта исследования обычно используют данные государственной и ведомственной статистики, результаты предыдущих опросов и исследований.

Наиболее распространенными видами целенаправленного подхода являются: типический и квотный.

Принцип типичности

Типический подход заключается в том, что отбор единиц репрезентации осуществляется не случайным способом. Исследователь сознательно выбирает то или иное «гнездо» (населенный пункт, предприятие, школу, избирательный участок и т.п.). При выборе конкретных «гнезд» он руководствуется в основном тем, чтобы это «гнездо» было наиболее типичным (наиболее распространенным типом) среди всех единиц репрезентации. Обычно этими соображениями руководствуются в тех случаях, когда число единиц отбора не может быть достаточно большим, а «гнезда», как подсказывают исследователю знания (или интуиция), недостаточно однородны. Типичность объекта может устанавливаться как расчетным способом (учитываются такие параметры, как мода — в распределениях по номинальным шкалам и среднее арифметическое — в распределениях по метрическим параметрам), так и по экспертным оценкам компетентных лиц. Несколько экспертов просят назвать наиболее типичные «гнезда». Например, представителей властных структур района (области) просят назвать наиболее типичные города, учитывая предусмотренные исследователем характеристики. Обычно, для начала, опрашивают трех экспертов, если их мнение совпадает (все называют одни и те же города в количестве, запланированном проектом выборки), то этим числом экспертов можно и ограничиться; если не совпадает, то число экспертов расширяется. Так же можно отбирать и типичные предприятия отрасли; естественно, что эксперты определяются в каждом конкретном случае, исходя из специфики объектов репрезентации. Обычно прибегают к услугам руководителей соответствующего уровня.

В рамках типичного объекта можно проводить сплошной опрос, а можно выборочный. Решение этого вопроса зависит от целей исследования, объема и однородности контингента «гнезда».

Принцип квот

В практике социологических исследований часто используется также **квотная** (пропорциональная) выборка. Основу выборки в данном случае составляют сведения о наиболее важных пропорциях генеральной совокупности, например, распределение по возрастным, профессиональным группам, по полу, национальности, типу населенных пунктов и т.п. характеристикам генеральной совокупности. По данным пропорциям, исходя из общего объема выборочной совокупности вычисляются соответствующие квоты, которые распределяются среди интервьюеров. Те, в свою очередь, осуществляют выбор респондентов, руководствуясь указанными параметрами.

Информацию, которая закладывается в расчеты основы выборки, обычно получают из отчетных документов государственной или ведомственной статистики. Чаще всего используют данные по полу, возрасту, образованию, типу семьи, типу поселения и т.п. Например, если известно, что население Украины составляют 54% женщин и 46% мужчин, то и в выборке населения Украины мужчины и женщины должны быть представлены в той же пропорции; или, например, при опросе населения Украины выборка должна включать 32% сельского населения и 68% городского, поскольку именно такая пропорция характерна для населения Украины в целом. Учитываются и другие важные пропорции (по возрасту, уровню образования, региону и т.д.). Чем больше пропорций учитывается, тем более вероятно, что выборка будет репрезентативной и по неучитываемым параметрам.

Существуют два вида техники квотного отбора. Первая заключается в том, что пропорции записываются в виде одномерных распределений по каждому из параметров; после отбора каждой единицы, ее параметры исключаются из дальнейшего отбора. Чем больше объектов отобрано, тем жестче критерии отбора, так как ряд квот наполнен. Допустим, если из ста человек отобраны 80, но среди них 46 мужчин, то в дальнейшем по такому параметру, как пол, отбору в выборку подлежат только женщины, так как квота мужчин (46%) уже набрана и т.д.

Другая техника предусматривает подготовку исследователем многомерной таблицы (бланка квот), в которой все необходимые пропорции пересекаются. Ниже представлен образец бланка квотной выборки. Такие бланки могут быть подготовлены по различным областям, отдельно для городской и сельской выборок.

Пол	мужской			женский		
образование	высшее	среднее	н/среднее	высшее	среднее	н/среднее
возраст						
18-19						
20-29						
30-39						
40-54						

55-59						
60 и ст.						

Такой бланк сначала составляется для всей выборочной совокупности, затем расписывается для отдельных интервьюеров. В итоге интервьюер получает такой бланк-задание, расписанный на 10-15 человек, в каждой из клеточек которого указано число людей с соответствующими параметрами, подлежащих опросу.

Несомненно, интервьюеру значительно легче работать с бланком квот, чем со списком одномерных распределений. Но иногда исследователь вынужден прибегнуть к первому варианту, если он не может получить информацию о пропорциях внутри групп: как распределяются по образованию, допустим, мужчины и женщины отдельно. Расчеты таких кросс-табуляций обычно можно заказать в соответствующих статистических учреждениях.

Если исследователь располагает данными об основных пропорциях (распределениях) генеральной совокупности, то метод квот очень удобен и экономичен.

Правомерность использования квотного отбора в опросах общественного мнения, как правило, среди специалистов вызывает бурные дискуссии.

Основные аргументы противников использования квотного отбора сводятся к следующим положениям: 1) использование метода квот не позволяет оценивать случайную ошибку выборки; 2) интервьюеру предоставляется слишком большая свобода выбора, поэтому существует опасность, что при отборе конкретных респондентов он будет руководствоваться принципами удобства (опрашивать или знакомых, или себе подобных, или общительных, приятных людей), что может привести к систематической ошибке; 3) произвольность при определении величины выборки (объема выборочной совокупности).

Что касается первого аргумента, то теоретически его трудно опровергнуть. Однако практика социологических исследований позволяет доверять этому способу отбора. Во-первых, неоднократно подтверждались результаты опросов, проведенных на основе квотной выборки, результатами выборов и референдумов. Эти данные зафиксированы как в Германии (см.: Э.Ноэль. Массовые опросы, с.172), так и в Украине в исследованиях Центрально-Украинского отделения Всесоюзного центра изучения общественного мнения (выборы и референдумы 1989-1991 гг.). Во-вторых, методическими экспериментами, когда часть исследования проводилась по случайной выборке, а часть — по квотной. Не следует забывать, что случайная выборка может давать большие отклонения, чем показывают теоретические расчеты, поскольку в данном случае «единицы отбора» представлены не безликими предметами (черными и белыми шариками), которые находятся в барабане и «не сопротивляются» выбору, а живыми людьми, вносящими своими действиями «коррективы» в случайную выборку. Нередко за «принципиальной» позицией сторонников «только случайного» подхода стоит бюрократический формализм, в основе которого — желание «отчитаться» за корректность выборки.

Кроме того, репрезентативность выборочной совокупности, опрошенной на основании квотного отбора, можно оценивать по «контрольным параметрам» — характеристикам, о пропорциях которых у социолога есть сведения, но они не включались в параметры определения выборки на этапе конструирования (например, данные о распределениях по национальности, семейному положению, типу семьи и т.п.). Среднее отклонение по всем контрольным параметрам может быть условно принято за среднюю ошибку репрезентативности.

Что касается «свободы выбора и субъективности интервьюеров», то, в определенной мере, этот упрек несправедлив, так как параметры квот сами по себе вынуждают интервьюера искать людей, отвечающих заданным характеристикам, особенно к концу опроса. Но поскольку такая угроза систематической ошибки все же возможна, то, чтобы избежать этого, используют метод квот «с искусственным ограничением». Интервьюера можно ограничить территориально: например, ограничив «поле поиска респондентов» рамками определенной улицы. Список улиц можно составить на основе случайного отбора улиц из общего списка улиц данного населенного пункта, а можно на основе районирования. Принцип районирования улиц лучше использовать, когда выборочная совокупность относительно невелика, а улицы города достаточно неоднородны. Возможны и другие «искусственные ограничения», вводимые в квотную выборку с целью «принуждения» интервьюера к отказу от поиска респондентов среди наиболее доступных для опроса контингентов (введение дополнительных квот и т.п.).

И наконец, проблема объема выборочной совокупности, касается не только квотного подхода. И поскольку объем выборки определяется не только учетом ошибок и отклонений, но и другими особенностями исследования (о которых мы будем говорить позже), то в этом отношении принцип квот — не хуже других.

Метод квот используется на разных ступенях выборки. Например, достаточно часто при подготовке общенациональных выборок на первой ступени отбора из каждого административного подразделения (например, области) отбирается подвыборка, численность которой в выборочной совокупности составляет долю, соответствующую доле населения, проживающего в данной области, в общей численности населения страны.

Наиболее часто квотный отбор осуществляется на последней ступени — при отборе единиц наблюдения. Если на предыдущих ступенях были последовательно использованы принципы районирования (стратификации) и кластеризации («гнезд»), то вероятность систематического смещения (возможность того, что интервьюер будет отбирать респондентов в одиозном месте, например, опросит пациентов больницы, соблюдая все заданные квоты) сводится практически к нулю.

Основным недостатком целенаправленных выборок является необоснованность статистических процедур анализа. Это не значит, что статистические коэффициенты вообще нельзя считать при таком подходе. Аналитические вычисления могут быть представлены для любых числовых значений данных, безотносительно к способу отбора. Но величины этих коэффициентов не обладают абсолютным значением и той легитимностью, которую обеспечивает случайная выборка. При анализе данных, полученных на основании опроса, проведенного по выборке, использующей целенаправленный принцип, значение статистических показателей следует рассматривать как относительное.

Если все же принимать во внимание определенную относительность расчетов, проводимых при подготовке случайной выборки, обусловленную прежде всего ограниченностью числа характеристик, используемых в расчетах, то исследователю можно рекомендовать использовать следующий ориентир при определении типа выборки.

Если исследователь достаточно хорошо знаком с предметом и объектом исследования и собирается проводить собственно социологическое исследование, то ему лучше ориентироваться на целенаправленный отбор.

Если исследователь недостаточно хорошо знаком с предметом и объектом исследования и собирается проводить опрос общественного мнения, ему лучше ориентироваться на случайные принципы отбора.

3. Объем выборочной совокупности

Так сколько же человек необходимо опросить, чтобы можно было делать выводы обо всей исследуемой совокупности?

Конкретный ответ на этот вопрос нетерпеливо ожидает заказчик, который готов платить за необходимую ему достоверную информацию. Даже неискушенный в исследованиях человек понимает, что, с одной стороны, «чем больше опрошено, тем лучше», а с другой, — «чем больше людей требуется опросить для получения достоверной информации, тем дороже общая стоимость исследования».

Ответ на этот вопрос не менее нетерпеливо ждет и начинающий социолог, понимая, что без обоснованного ответа он не может приступить к обсуждению проекта конкретного эмпирического исследования в целом.

Самым нетерпеливым сразу предоставляем возможность ознакомления с формулой расчета величины выборочной совокупности, осуществляемой одноступенчатым случайным бесповторным отбором.

$$n = \frac{t^2 \cdot d \cdot N}{D^2 + t^2 \cdot d}$$

где N — объем генеральной совокупности,

n — объем выборочной совокупности,

t — коэффициент нормированного отклонения; в социологической практике в основном употребляют три значения нормированного отклонения: $t=1$, $t=2$, $t=3$; с вероятностями соответственно $P=0.68$, $P=0.95$ и $P=0.99$; эти значения интерпретируются следующим образом: если из генеральной совокупности извлечь 100 выборок одинаковым образом и с одинаковыми объемами, то можно гарантировать, что в 68 (95 или 99, соответственно) из них допустимая (граничная) ошибка не будет превышать среднюю ошибку;

D — допустимая ошибка; в социологических исследованиях максимальная величина допустимой ошибки принята равной (плюс-минус) 5%; уменьшение абсолютной величины допустимой ошибки достигается увеличением объема выборочной совокупности;

d — значение дисперсии признака, по которому рассчитывается репрезентативность в генеральной совокупности.

Если выборка многоступенчатая, то учитывается ошибка на каждой ступени. Разработаны формулы расчета объема выборочной совокупности для многоступенчатых комбинированных выборок. Социологам, предполагающим совершенствоваться в статистическом обосновании расчета объема выборки, прежде чем приступать к самостоятельным расчетам, следует внимательно ознакомиться с основами теории вероятности и математической статистики.

Особенности технологического подхода к освоению расчета объема выборки предполагают освоение основных правил и способов, позволяющих избежать грубых ошибок при формировании выборочной совокупности и интерпретации полученных на основе сформированной выборки данных.

Практика социологических исследований показывает, что обычно обоснованный объем выборочной совокупности составляет от 400 до 2000 единиц наблюдения.

Из этого следует, что в стандартном исследовании двух тысяч опрошенных вполне достаточно, чтобы получить репрезентативную, надежную и обоснованную информацию (конечно, при условии соблюдения всех правил процедуры отбора).

Но перед исследователем всегда стоит задача — по возможности сократить объем выборки для того, чтобы снизить организационную нагрузку и финансовую стоимость исследования.

Что должен учитывать автор проекта выборки, определяя **минимальный** объем выборочной совокупности, способной обеспечить достаточно обоснованный базис респондентов?

1. Однородность (неоднородность) исследуемого контингента. Чем более однороден контингент, с точки зрения влияния его характеристик на предмет исследования, тем меньший объем выборки может обеспечить качественную информацию. Однородность определяется, во-первых, числом характеристик, которые могут влиять на ответы респондентов, и, следовательно, должны учитываться при формировании выборки; во-вторых, разбросом значения каждой из этих характеристик в генеральной совокупности. Например, если исследователь предполагает задать вопрос: «Хотите ли Вы, чтобы Вам повысили зарплату?», можно предположить, что различные характеристики объекта исследования (пол, возраст, образование и даже доход) не будут существенно влиять на различия в ответах. Если же исследователь предполагает спросить: «Хотите ли Вы, чтобы предприятие, на котором Вы работаете, было приватизировано?», то прежде, чем приступить к формированию выборки, он должен задуматься, какие характеристики работников могут влиять на их отношение к приватизации. Конечным результатом такого предварительного анализа объекта должен стать перечень этих характеристик, а также сведения о том, как по данным характеристикам распределяются люди, представляющие собой генеральную совокупность. Чем меньше рассеивается признак по своему значению, тем меньший объем выборки требуется для обеспечения репрезентативности. В приведенной выше формуле расчета объема выборочной совокупности этот показатель называется дисперсией.

Принимая во внимание, что формула расчета объема выборки учитывает только один признак, исследователь среди всех гипотетических признаков должен сначала отобрать самые «влияющие», с его точки зрения; а затем из их числа выбрать один — наиболее разбросанный.

2. Исследователь, в зависимости от задач проекта, может повышать допустимую ошибку и варьировать доверительный интервал. Работая над проблемой объема выборки, он может задавать вопрос (как заказчику, так и самому себе): «Что если по данному пункту анкеты мы получим значение, на 1% отклоняющееся от истинного; это допустимо?», «А если значение будет отклоняться на 5%?» Обычно в этих пределах социолог ищет ответы на поставленные вопросы и решает проблему допустимой ошибки и доверительного интервала.

3. Рассчитывая минимальный объем выборки, не следует забывать, что он обеспечивает статистической достоверностью лишь одномерные распределения. Если социолог предполагает представлять данные в виде многомерных распределений и анализировать причинно-следственные зависимости, он должен принимать во внимание **аналитические** требования к выборочной совокупности: наполненность всех подгрупп, которые образуются в процессе анализа.

Дополнительная задача выборки в таком случае формулируется следующим образом: «каков должен быть объем выборки для того, чтобы произвести с нею все необходимые аналитические манипуляции с достаточно большой устойчивостью или надежностью итоговых показателей?» (Здравомыслов А. Методология и процедура социологических исследований, М., 1969. — с.99). С точки зрения аналитических возможностей, при определении объема выборки должен учитываться план анализа (составная часть программы исследования). Прежде чем приступить к проектированию выборки, исследователь должен построить макеты основных аналитических таблиц и посчитать, сколько подгрупп будет образовано в результате кросс-табуляции. Допустим, исследователь предполагает проверить гипотезу о влиянии уровня образования на удовлетворенность жизнью. Показатель предполагаемого фактора (уровня образования) включает четыре значения признака (неоконченное среднее, среднее общее, среднее специальное, высшее), а показатель изучаемого явления 5 значений (уровни удовлетворенности жизнью). Следовательно, макет аналитической таблицы (4x5) предполагает формирование 20 подгрупп. Минимальная численность каждой подгруппы должна составлять 25-100 человек (в зависимости от доверительного интервала, определяемого исследователем). Соответственно, общий объем выборки может варьировать от 500 до 2000 человек, в зависимости от заданной исследователем точности. Чем больше градаций анализируемых признаков и выше желаемая точность измерений, тем больше должен быть объем выборки.

Бывают случаи, когда та или иная социальная группа обладает, по мнению исследователя, достаточно выраженными особенностями, представляющими принципиальный интерес с политической и социальной точек зрения (например, предприниматели, фермеры, военнослужащие и т.д.) для анализа той или иной социальной проблемы. Результаты сравнительного анализа по различным параметрам стабильно подтверждают гипотезы исследователя об особенностях этой социальной группы. Но в репрезентативной региональной выборке представителей данной категории слишком мало, чтобы можно было сделать статистически обоснованные выводы. В таких случаях исследователь может обратиться к специальным выборкам, в частности, сформировать «выпуклую» выборку.

Выпуклая выборка включает основную (репрезентативную) совокупность, кроме того, составляется дополнительная подвыборка по интересующей исследователя категории. Минимальный объем дополнительной подвыборки должен составлять около 100 респондентов; оптимальный объем определяется особенностями предполагаемого анализа — количеством градаций в ответах на вопросы, по которым автор предполагает

сравнивать исследуемые категории населения: чем больше градаций, тем больше должна быть выборка — как основная, так и дополнительная.

В случаях проведения исследования по «выпуклой» выборке, в процессе общего анализа полученных результатов дополнительная выборка исключается из массива данных. Это позволяет автору делать выводы, касающиеся распределений мнения населения в целом (т.к. основная выборка репрезентативна). В процессе сравнительного анализа, исследователь включает «дополнительную подсовокупность», увеличивая численность анализируемой категории. В таком случае полученные результаты смогут пройти статистическую экспертизу. Конечно, подготовка выпуклой выборки — процедура, требующая дополнительных средств (как финансовых, так и организационных). Но в таких ситуациях требуется идти на какие-то жертвы: 1) отказаться от анализа особенностей социальных групп, доля которых слишком мала для возможностей их характеристики в рамках репрезентативного опроса населения; 2) отказаться от репрезентативности и набирать выборку по группам и в процессе анализа сравнивать их между собой, но не иметь возможности судить о совокупном мнении населения какого-либо региона; 3) производить дополнительные затраты на подготовку выпуклой выборки.

Таким образом, можно видеть, что проблема объема выборочной совокупности решается всегда путем поиска компромиссного решения между точностью измерений и затратами на обеспечение этой точности.

4. Заключение: технологическая схема подготовки выборки

Прежде всего необходимо четко определить, какой из принципов подхода к отбору респондентов является более адекватным основной цели исследования — *выборочный* или *аналитический*. (Напомним, что аналитический подход чаще целесообразнее применять в собственно *социологических исследованиях*; выборочный — в *опросах общественного мнения*).

В тех случаях, когда результаты социологического опроса части определенной категории населения предполагается прямо проецировать на всю исследуемую категорию, необходимо соблюдать следующие требования, предъявляемые к выборочному опросу.

1. Чтобы сформировать выборку, необходимо располагать сведениями об основных характеристиках генеральной совокупности.
2. Основными характеристиками считаются те, возможное влияние которых на изучаемое явление обосновывается исследователем.
3. Сведения о характеристике подразумевают: а) название характеристики (пол, возраст, образование, доход, тип поселения, национальность, владение языком и т.д., и т.п.); б) данные по распределению единиц совокупности по значениям данного признака.
4. Если исследователь не располагает сведениями об основных, с его точки зрения, характеристиках, он должен учитывать, по крайней мере, наиболее общие (как правило, значимые в большинстве исследований): пол, возраст, уровень образования, регион проживания, тип поселения. Знание этих параметров является минимально необходимой основой расчета и контроля выборки.

5. Если автор проекта не располагает никакими сведениями, позволяющими описать генеральную совокупность, то о выборке в истинном значении этого слова, говорить не приходится.

6. Если автор не может определить опрошенную совокупность как выборку, репрезентирующую определяемую часть населения, то, приводя данные результатов своего опроса, он должен предварять полученную информацию описанием этой совокупности по основным параметрам: привести данные полученных распределений по полу, возрасту, образованию, национальности, а также другим социально-демографическим характеристикам опрошенной совокупности.

В тех случаях, когда основной целью социологического опроса является анализ предмета — изучение причинно-следственных взаимосвязей, целесообразнее применять аналитический подход к отбору респондентов.

При аналитическом подходе следует учитывать следующие необходимые требования.

1. Объект исследования репрезентируется направленным отбором соответствующей группы людей — **основной** группы, а также — **контрольной** группы.
2. Группы (основная и контрольная) должны различаться по основному дифференцирующему признаку (гипотетическому фактору). По всем остальным параметрам (возрасту, полу, уровню образования, характеру деятельности и т.д.) группы должны быть максимально идентичны.
3. Численность групп может ограничиваться небольшим объемом (50-100 человек).
4. Если гипотетических факторов больше одного, то в рамках обеих групп (основной и контрольной) можно набирать равночисленные подгруппы. Введение каждого нового фактора требует соответствующего увеличения каждой из групп. В образованные подгруппы анализа необходимо включать 50-100 человек. Чем больше по численности будет каждая набранная подгруппа, тем выше вероятность того, что удастся статистически обосновать гипотезы исследования.

*В практике массовых опросов существует еще один способ отбора, не укладывающийся в рамки основных типов выборки. Это — «стихийная выборка», или «выборка первого встречного». Примерами таких выборок могут служить опросы, проводимые через средства массовой информации («прессовые опросы») или у дверей магазинов в тех случаях, когда интервьюеры опрашивают не людей, отобранных по специальной процедуре, а тех, кто подвернулся под руку и готов ответить на вопросы анкеты. Опросы, проведенные по такому принципу, позволяют получить определенную информацию. И если авторы исследования при обнародовании полученных данных указывают, что опрос проводился именно таким способом, то потребитель информации всегда может учесть, что выборка не является репрезентативной по отношению к соответствующим категориям населения в целом. Точнее, в таких исследованиях категория «выборка» используется очень условно, так как по существу опрошенная группа населения выборкой не является.

Организация сбора данных

1. Два способа сбора данных: преимущества и недостатки

При проведении эмпирического социологического исследования методами интервью или раздаточного анкетирования руководитель исследовательского проекта должен решить вопрос о способе сбора данных. Можно выделить два основных подхода к решению этого вопроса: заказать сбор данных организации, имеющей сеть интервьюеров, или организовать этот процесс своими силами.

Каждый из этих способов имеет свои преимущества и недостатки. Поэтому на этапе проектирования руководитель исследования должен тщательно взвесить все «за» и «против» того и другого способа в рамках осуществления своего проекта и после согласования с заказчиком или руководителем учреждения, отвечающим за финансирование проекта, принять окончательное решение еще до начала выполнения работы. Какими же преимуществами и недостатками обладает каждый из этих способов, которые должен учитывать исследователь, принимающий решение о способе сбора первичной социологической информации?

Организации (фирмы) по сбору данных

В настоящее время организациями, специализирующимися по сбору данных, обычно выступают подразделения при крупных социологических центрах (институты социологии, службы изучения общественного мнения, университетские кафедры социологии и т.п.), а также частные фирмы и агентства, основанные как различными предприятиями и учреждениями, так и отдельными лицами, имеющими опыт подобной работы.

Основным преимуществом таких организаций является то, что они обладают сетью подготовленных интервьюеров в различных регионах страны; и если выборка, предусмотренная исследованием, является достаточно разбросанной, то осуществить исследование своими силами в установленный срок зачастую является для исполнителей проекта непосильной задачей.

Другим преимуществом заказа этого этапа исследования организациям, специализирующимся на сборе данных, является возможность максимально сокращать сроки исследования, так как при необходимости они могут привлекать достаточно большое число интервьюеров.

Третьей особенностью такого способа организации является тот факт, что организации обладают списками подготовленных и достаточно опытных интервьюеров; и поэтому исследователям нет необходимости затрачивать время и средства на подбор и общую подготовку людей к данному виду работы. Это обстоятельство особенно следует принимать во внимание при сравнительно простом исследовании (например, опросах общественного мнения). Таким образом, заказ сбора данных специальным службам может быть выгоднее как с организационной, так и с экономической точки зрения в тех случаях,

когда нецелесообразно специально осуществлять процедуру отбора и подготовки интервьюеров.

Однако такой способ сбора данных имеет и существенные недостатки. В первую очередь это относится к возможностям контроля качества работы и, соответственно, качества полученной информации; особенно в тех случаях, когда сбор информации осуществляется в местах, достаточно отдаленных от исследовательской группы. Фактически очень трудно (а зачастую и невозможно) уберечься от плохого качества работы, если у исследователя ограничены возможности контролирования процесса сбора данных. Нередко у исследователя может не быть полной уверенности в качестве первичной социологической информации не только после окончания сбора данных, но и даже после того как данные обработаны и проанализированы. В таком случае, сталкиваясь в процессе анализа с некоторыми неожиданными данными (артефактами или неожиданными коэффициентами связи), исследователь затрудняется в выборе определения причины выявленного опросом явления. Перед ним неизбежно возникает вопрос: надо ли заниматься аналитическим поиском объективных причин или причиной труднообъяснимых, с первого взгляда, данных является недобросовестная работа интервьюеров.

Таким образом, даже если сбор данных реально был осуществлен очень добросовестно, отсутствие полной уверенности исследователя в качестве проведенной работы существенно мешает ему при анализе информации, когда он сталкивается с неожиданными или не поддающимися достаточно убедительному объяснению данными.

Сбор данных собственными силами

Когда исследователь сам организует работу интервьюеров, у него значительно больше возможностей проконтролировать весь процесс сбора данных. Кроме того, поскольку исследователь лично участвует в отборе и подготовке интервьюеров, то и интервьюеры в значительно большей мере испытывают чувство личной ответственности перед авторами проекта.

В случае самостоятельной организации процесса сбора информации, выше не только качество контроля, но и уверенность исследователя в достоверности полученных данных. Кроме того, даже если в процессе сбора информации были какие-либо отклонения, накладки, погрешности, то исследователь, будучи о них осведомленным, при анализе информации знает, в каком случае некоторые неожиданные данные могут являться результатом снижения качества при сборе информации и в какой степени это могло повлиять на результаты исследования, а в каком случае объяснение полученным данным следует искать в объективных закономерностях социальных реалий и углублять содержательный анализ.

Однако при самостоятельной организации сбора информации есть и существенные недостатки. В первую очередь следует учитывать значительную дополнительную нагрузку, связанную с бременем ответственности за подбор и подготовку интервьюеров, организацию их работы, полный контроль качества выполнения всех предусмотренных проектом требований, замену интервьюеров, финансовые расчеты с достаточно многочисленным составом внештатных сотрудников и т.п.

Таким образом, на этапе проектирования исследовательский коллектив должен взвесить все достоинства и недостатки каждого из этих способов, учитывая особенности

конкретного проекта и программы исследования, собственные возможности, потребности и ресурсы, и принять решение о том, какой способ сбора информации предпочтительнее для этого проекта.

В каждом исследовании есть свои особенности, которые могут повлиять на этот выбор. Но в самом общем виде рекомендации в пользу выбора того или иного способа организации сбора данных могут быть следующими.

Сбор данных лучше заказать организациям, специализирующимся на сборе первичной социологической информации в следующих случаях:

- 1) исследование проводится по сравнительно простой анкете, легко доступной для понимания и заполнения;
- 2) сбор данных необходимо осуществить в достаточно сжатые сроки;
- 3) выборка является типичной и при этом довольно разбросанной.

Самостоятельную организацию сбора данных лучше осуществлять при следующих особенностях проекта:

- 1) проект является научно-исследовательским, с заранее запланированными и не лимитированными сроками — т.е. нет необходимости фиксировать мнение (чувства, состояния) большого числа людей примерно в одно и то же социальное время;
- 2) исследовательский коллектив имеет достаточно большой штат постоянных сотрудников из числа вспомогательного персонала, чьи усилия можно будет в течение достаточно длительного времени (от полугода до одного года) сосредоточить на выполнении данного проекта;
- 3) выборка либо локализована достаточно компактно (одно или несколько крупных предприятий, учреждений; один населенный пункт и т.п.) или, наоборот, некоторые единицы выборки находятся в слишком отдаленных или труднодоступных местах*.

После того, как принято окончательное решение о способе сбора информации, необходим план действий по подготовке к осуществлению выбранного пути.

Каждый из способов организации сбора информации обладает своими особенностями и правилами, учет и соблюдение которых позволяет максимально приближаться к основной цели эмпирического исследования — получение качественной достоверной информации об изучаемых социальных фактах, явлениях, процессах, закономерностях.

2. Технология работы с организациями (фирмами) по сбору данных

Выбор фирмы

В том случае, когда принято решение осуществлять сбор данных с помощью организаций, специализирующихся в области сбора социологической информации, возникает проблема выбора организации (фирмы) для заключения контракта**.

Эта проблема является сложной не только для социологов Украины и других стран бывшего Советского Союза, для которых организация крупномасштабных социологических исследований с привлечением вспомогательных организаций на коммерческой основе является довольно молодой отраслью научной и научно-практической деятельности, но и для исследователей, работающих в странах, в которых на протяжении десятилетий накоплен достаточный опыт использования специальных опросных служб.

Например, в США существует специальная промышленная ассоциация, называемая Market Research Association (M.R.A.), которая объединяет почти все агентства по сбору данных. Тем не менее выбор конкретного агентства представляет, по мнению американских исследователей, существенную проблему, так как практически все агентства на словах готовы взяться за выполнение работы, уверяя, что они могут обеспечить все необходимые условия; хотя их возможности и оснащенность совершенно различны, начиная от агентств, представленных одним человеком — владельцем фирмы, осуществляющим всю координационную деятельность из своей квартиры*, до достаточно солидных фирм, имеющих филиалы в различных регионах, которые оснащены не только необходимой техникой и специальными залами с видео- и аудиотехникой для сбора и подготовки интервьюеров, но даже магазинами и кухнями, способствующими привлечению людей к работе.

Степень оснащенности организации, несмотря на то что этот аспект, конечно, принимается во внимание при выборе фирмы для осуществления сбора данных, сама по себе еще не является гарантией качественного выполнения заказанной работы: добросовестного соблюдения всех требований проекта, полного контроля работы интервьюеров и т.п. условий, гарантирующих достоверную информацию. Более того, даже одна и та же фирма может осуществлять сбор данных с разной степенью качества в зависимости от многих условий.

В первую очередь качество их работы зависит от степени подготовленности и четкости требований самого исследователя, заказывающего сбор данных, а также от степени контроля с его стороны за работой фирмы. Но качество работы фирмы по сбору данных может зависеть и от конкретных обстоятельств, например, степени загруженности работой (заказами) в данный момент времени. Выполняя одновременно несколько заказов, фирма может привлечь интервьюеров из резервного списка или набрать новых, не имеющих достаточного опыта. При этом не затрачивается необходимое время на их подготовку (специфика общей подготовки интервьюеров и их подготовки к конкретному исследованию будет рассмотрена в разделе «подготовка интервьюеров»).

Этапы работы

При планировании работы с фирмой по сбору данных следует выделить шесть основных этапов:

- 1) выбор фирмы;

- 2) предварительный контакт, подписание контракта (договора);
- 3) передача материалов;
- 4) подготовка и инструктаж организаторов сбора данных (полевого директора, кураторов, супервайзеров);
- 5) мониторинг, надзор и контроль;
- 6) получение данных, проверка и оплата.

1. **Выбор фирмы.** Выбор фирмы осуществляется на основе получения предварительной информации о наличии таких служб. Пока еще, к сожалению, в Украине нет специализированных бюллетеней с характеристикой соответствующих фирм, но, тем не менее, в среде социологов функционирует общая информация о деятельности таких организаций и качестве работ, которые они выполняли. Кроме того, в прессе публикуются данные с ссылкой на организации, проводящие сбор информации и т.п. Исследователю, планирующему заказ по сбору данных, следует лично посетить такие фирмы, побеседовать с руководителем, ознакомиться с действующими тарифами и техническим оснащением фирмы. Желательно побеседовать с некоторыми сотрудниками, непосредственно занимающимися организацией работы интервьюеров. Следует отметить, что общее впечатление никогда не бывает полной гарантией высокого качества работы.

Если исследователь видит заинтересованность руководителей нескольких фирм в получении заказа на его исследование и затрудняется с окончательным выбором фирмы, он может попросить руководителя дать адреса и телефоны предыдущих заказчиков с тем, чтобы можно было ознакомиться с их впечатлением от результатов работы фирмы.

2. **Предварительный контакт.** Предварительный контакт с выбранной фирмой лучше осуществлять по следующей схеме: передать в фирму письменное уведомление с описанием требований проекта (выборка, величина анкеты, сроки выполнения), предварительно сообщив по телефону, что такие материалы переданы, и договориться о личном визите. Обсуждение взаимных возможностей и обязанностей лучше проводить по телефону, а в личной беседе.

В предварительной беседе обсуждаются сроки проведения опроса, тип опроса (раздаточное анкетирование или интервьюирование), конкретные точки опроса (населенные пункты, в которых будет осуществляться сбор информации), объем и тип выборки, методы контроля качества работы интервьюеров, стоимость одного интервью и общая стоимость работы. (Подготовленную анкету предварительно передавать не следует, но ее надо принести с собой, чтобы руководитель фирмы мог ознакомиться с ее содержанием и оформлением, так как от сложности анкеты зависит стоимость опроса). Если переговоры с фирмой проводятся задолго до начала исследования (например, на этапе подготовки проекта), то тогда заключается лишь устное соглашение о принципиальной возможности фирмы взять на себя выполнение такой работы. В этом случае стороны договариваются о том, что, когда подойдет срок сбора информации, исследователь уведомит фирму. Такое уведомление следует делать не менее, чем за неделю в случае небольшого опроса, и не менее, чем за месяц, в случае крупномасштабного исследования.

Во время предварительной беседы целесообразно договориться о названии проекта. Это название не обязательно должно точно отражать суть исследования, оно призвано

выполнять функцию ярлыка, который позволяет фирме ориентироваться при переговорах — о каком исследовании идет речь. Такое название должно содержать одно-два ключевых слова.

3. Передача материалов. Перед передачей материалов исследования следует уведомить фирму — как они будут доставлены и когда ожидать их получение. Каждый пакет переданных материалов должен содержать надпись (номер пакета и название исследования) и опись вложенных в него материалов; копии этих описей хранятся и у исследователя. Эта мера позволяет легко ориентироваться в материалах как работникам фирмы (ведь фирма часто одновременно выполняет не один заказ, а несколько), так и исследователям вести точный учет переданных материалов.

4. Инструктаж работников фирмы. В числе передаваемых в фирму по сбору данных материалов должны быть письменные инструкции по проведению опроса, адресованные супервайзерам, бригадирам и интервьюерам. Если исследователи располагают такой возможностью, то лучше в назначенный день лично прийти проинструктировать сотрудников фирмы, которые будут курировать данный опрос. Такой инструктаж обычно включает общие сведения о проекте (его цель и основное содержание), напоминаются основные правила проведения опроса, в случае нарушения которых будут применены материальные санкции, освещаются особенности и специальные детали, на которые следует обратить внимание именно в данном опросе, а затем переходят к тщательному ознакомлению с материалами исследования: анкетой, письменными инструкциями и карточками ответов (если они предусмотрены проектом)*.

Самая важная часть инструктажа — «прохождение» анкеты: внимательное ознакомление с каждым вопросом анкеты и инструкциями к нему. Здесь исследователь должен не только познакомить супервайзеров с каждым вопросом анкеты, но и с возможными реакциями респондентов, прокомментировав, как должен реагировать интервьюер на те или иные возможные реакции. (Как могут реагировать респонденты, исследователь обычно знает как из своего общего опыта и компетентности, так и из результатов претестирования — пилотажного опроса). Затем в присутствии исследователя сотрудники фирмы должны провести пробный опрос по анкете друг на друге.

Основная цель такой подготовки — полное ознакомление со всеми деталями и подробностями опроса. Все сотрудники, участвующие в организации опроса, должны иметь одинаковое (максимально унифицированное) представление о том, как должны вести себя интервьюеры.

Всю эту процедуру, которая длится обычно три-четыре часа, супервайзеры должны воспроизвести с интервьюерами, которые будут принимать непосредственное участие в данном опросе.

Инструктаж — наиболее важный элемент организации опроса, способствующий созданию условий получения качественной информации. Основная цель организации сбора информации — стараться создать идентичные условия для респондентов, что позволяет обеспечить сбор сопоставимой информации.

Еще раз напомним два основных правила, позволяющие приблизиться к этой цели: 1) наличие письменных инструкций для супервайзеров и интервьюеров; 2) доскональное знание интервьюером анкеты, по которой он проводит опрос.

Если исследователи не могут лично участвовать в инструктаже (в силу удаленности от места организации сбора информации или другим причинам), предполагается, что процедура инструктажа обеспечивается руководителем фирмы или назначенным им «полевым директором» проекта. Однако в этом случае исследователь должен осознавать, что он перекладывает ответственность за качество работы на фирму, с которой заключен контракт.

5. Мониторинг и контроль. Для повышения уверенности в качестве работы целесообразно осуществлять «инспекционный надзор» на протяжении всего процесса сбора данных. Если приступить к контролю только после получения массива заполненных анкет, то можно столкнуться с многими неприятными сюрпризами, которые гораздо легче предотвратить в процессе сбора информации, чем исправить положение после окончания основной работы.

Осуществление постоянного «инспекционного» участия в процессе сбора данных включает в себя отслеживание процесса (мониторинг) и контроль.

Мониторинг предполагает, что одному из соисполнителей исследовательского коллектива, заказавшего сбор данных, отводится функция инспектора-куратора этого процесса. Он держит постоянную телефонную связь с полевым директором, назначенным фирмой, ежедневно (или, по крайней мере, несколько раз в неделю) звонит и спрашивает, какая часть работы сделана, не возникли ли какие-либо вопросы или проблемы, осложняющие процесс сбора информации и т.п.

Такое постоянное «участие» мобилизует работников фирмы, способствует предотвращению нарушения инструкций, графика работы, пренебрежительного отношения сотрудников фирмы к требованиям заказчика и т.п.

Необходимо также контролировать выборку. Если исследование проводится по квотной выборке, то от интервьюеров должны быть получены сведения о респондентах (имя, адрес и телефон респондента), чтобы можно было осуществить прямой контроль выполнения их работы. Такая мера повышает ответственность интервьюера: интервьюеры значительно реже идут на возможные нарушения, когда знают, что их работа может быть проконтролирована, и в случае выявленных нарушений они могут быть лишены оплаты или даже (в случаях прямых фальсификаций) полностью отстранены от работы интервьюера.

Если исследование проводится по адресной выборке, с указанием конкретных респондентов, подлежащих опросу, то исследователи для проверки качества работы интервьюеров могут использовать «подставные адреса» — адреса инспекторов. В качестве таких инспекторов могут быть привлечены как штатные сотрудники исследовательского коллектива или спонсора исследования, так и специально приглашенные для этой работы внештатные сотрудники. Такой прием позволяет не только определить, достаточно ли точно соблюдается выборка и инструкции по отбору респондентов, но и оценить качество работы интервьюера в процессе проведения опроса, его следование инструкциям.

В подобном способе контроля нет необходимости при достаточно простых исследованиях, он скорее используется в тех случаях, когда поведение интервьюера может существенно повлиять на качество получаемой информации.

6. Прием работы, проверка и оплата. Когда опрос завершен и фирма передала исследователю заполненные бланки анкет, в первую очередь необходимо проверить полноту и качество заполнения. Некачественно заполненные анкеты должны быть отобраны и отложены. (На анкетах, которые заполнены правильно, но по какой-либо причине вызывают сомнение, следует сделать соответствующие пометки). На каждой из отбракованных анкет следует четко написать причину, по которой эта анкета изъята из массива. Их нельзя уничтожать и они должны быть сохранены до момента полного расчета с фирмой за выполненную работу. Качественно заполненные анкеты должны быть пронумерованы и подготовлены к последующей докодировке и вводу.

Часть принятых анкет готовится к контролю. Здесь могут быть применены разные способы отбора анкет для контроля. Могут быть отобраны бланки анкет «по подозрению». (Во время просмотра анкет после их получения у исследователей могли возникнуть сомнения в отношении добросовестности некоторых интервьюеров или некоторых конкретных интервью, тогда они должны были сделать соответствующую отметку на обложке анкеты, например, знак вопроса). Отбор анкет для контроля может быть проведен по случайной выборке, например, каждая десятая или каждая двадцатая (доля контролируемых анкет зависит от финансовых и организационных возможностей исследовательского коллектива). Но какая-то часть анкет должна быть проверена обязательно.

Качество заполнения при визуальном контроле может не вызвать сомнений, но проверка может обнаружить недостоверность содержащейся в них информации. Поэтому наряду с контролем «по подозрению» необходимо готовить к прямому контролю и какую-то часть анкет, не вызывающих подозрение при визуальном контроле. Редко, но все же бывают случаи прямой фальсификации данных, когда интервьюер вместо опроса сам заполняет бланки анкет. Гораздо чаще распространены случаи частичной фальсификации: интервьюер не полностью зачитывает вопросы, упрощая их формулировки, опускает некоторые (как правило, наиболее сложные или щепетильные) вопросы, а потом сам дозаполняет их дома. Иногда сотрудники фирмы, отвечающие за сбор данных, сами «корректируют» некачественно заполненные анкеты и т.п.

Проверка чаще всего осуществляется по телефону. Респондента просят ответить, приходилось ли ему участвовать в опросе, когда это было, сколько примерно времени ушло на интервью; затем ему задают несколько вопросов из анкеты, сопоставляя с заполненным бланком*. Если у респондента нет телефона, то такая же проверка осуществляется посещением по адресу.

В случае обнаружения расхождения между данными, содержащимися в заполненной анкете, и результатами проверки, из принятого массива данных могут быть исключены как отдельные анкеты, так и все анкеты отдельных интервьюеров (если такие погрешности обнаружены у одного интервьюера в нескольких проверенных анкетах). Если подобные случаи обнаружены в работе ряда интервьюеров, исследователь может забраковать весь массив данных, собранных фирмой. Такие условия следует оговорить в контракте, заключаемом с фирмой, при его подписании.

Изложенные выше основные правила организации работы по сбору информации в случаях, когда эта работа заказывается организациям специализирующимся на данном виде работы, помогут исследователю избежать основных недостатков такого способа организации сбора информации: низкого качества полученной информации и собственной неуверенности в качестве работы на этом этапе.

3. Технология самостоятельной организации сбора данных

В тех случаях, когда исследователи по тем или иным причинам предпочитают сами организовать процесс сбора информации, то они принимают на себя ответственность за подбор, общую подготовку и непосредственный контроль работы интервьюеров, которые в большинстве случаев работают далеко от центра, проводящего исследование.

Если проект исследования предусматривает небольшую выборку (от 30 до 100 человек) и достаточно длительный срок проведения опроса (3-12 месяцев), то исследование может быть проведено силами 1-3 человек (исследователь и 1-2 помощника). Но таких исследований немного; как правило, это экспертные опросы, поисковые исследования, исследования конкретной компактно локализованной выборки, учебные работы (дипломные и курсовые), исследования, предусматривающие не репрезентативную выборку, а опрос основной и контрольной групп с последующим сравнением результатов (принцип эксперимента) и т.п.

Значительно чаще проекты исследования предусматривают репрезентативные выборки определенных групп населения, объем которых (даже при достаточно длительных сроках опроса) требует дополнительного привлечения интервьюеров, число которых зависит от характера и объема выборки, а также сложности и объема инструментария. В случаях организации сбора информации своими силами для осуществления этого этапа работы создается группа сбора информации. Количественный состав участников такой группы зависит от вышеуказанных особенностей проекта исследования (выборки и инструментария): однако общая структура группы и принципы организации ее работы весьма типичны.

Структура рабочей группы

Типичная структура группы сбора данных может быть представлена следующей схемой: полевой директор — супервайзеры — бригадиры — интервьюеры.

Руководитель проекта обычно является и основным руководителем группы сбора данных. Он назначает соруководителя («полевого директора»), который несет полную ответственность за соблюдение всех требований проекта на этапе сбора информации. В качестве полевого директора может выступать и сам руководитель проекта, но это целесообразно только лишь в том случае, если исследование не очень большое, и руководитель проекта в это время свободен от выполнения других функций. Полевого директора лучше назначать из штатного числа исполнителей проекта или сотрудников учреждения, в котором работает руководитель проекта, поскольку последнему хорошо известны деловые и человеческие качества кандидата на эту роль, его пригодность для осуществления необходимых функций.

Функции полевого директора. В основные обязанности полевого директора входит набор, подготовка и инструктаж интервьюеров и супервайзеров, координация их деятельности, организация контроля процесса работы и качества результатов.

Требования к кандидатам на эту должность отличаются необходимостью сочетания взаимоисключающих, на первый взгляд, личностных и деловых качеств: гибкость в общении с людьми и в то же время требовательность к ним; умение самостоятельно

принимать решение в нестандартной ситуации и при этом исполнительская дисциплинированность в соблюдении основных требований. Другими словами, полевому директору требуются качества, необходимые любому руководителю. Однако в данном случае нет времени отрабатывать необходимые организационные способности, стиль и методы руководства, что обычно достигается в процессе длительного взаимодействия с подчиненными. Поэтому желательно, чтобы у кандидата на должность полевого директора уже был определенный опыт организационной работы.

При формировании группы сбора информации необходимо учесть, что у полевого директора обязательно должен быть технический помощник. Требования к выбору кандидатуры на эту роль определяются спецификой основных обязанностей: поскольку сбор социологической информации предполагает отправку и получение большого количества пакетов с рабочими материалами, технический помощник должен обладать достаточной физической силой и не иметь противопоказаний к физической работе в связи с состоянием здоровья. Также необходимо учитывать, что при передаче материалов используются не только средства почтовой связи, но и различные другие способы — передача через проводников поезда, водителей междугородних автобусов и т.п. (мы здесь не будем останавливаться на причинах подобной транспортировки материалов, а лишь заметим, что опыт проведения массовых опросов показывает, что таких ситуаций не удастся избежать практически ни в одном эмпирическом исследовании), то иногда возникает необходимость «встречать» материалы в «неудобное» время (ночью, рано утром или поздно вечером). Из специфики такого рода деятельности вытекают и требования к качествам кандидатуры на эту должность: физическая выносливость, определенного рода непритязательность, личные возможности к мобильности в любое время суток. В тех случаях, когда организатор сбора информации заранее не предусматривает вакансию технического помощника полевого директора, в группе сбора информации нередко возникают дискуссии и даже конфликты в случаях необходимости передавать или получать материалы в «неудобное» время суток. Кандидатуру на эту должность можно выбрать как из штата сотрудников организации, проводящей исследование, так и пригласить на работу по контракту, обязательно проинформировав о характере предстоящей работы и заручившись согласием выполнять данные функции.

Супервайзеры, бригадиры, интервьюеры. Количество необходимых для сбора данных супервайзеров, бригадиров и интервьюеров определяется: 1) сроками сбора информации; 2) особенностями выборки (объемом и характером); 3) объемом и сложностью инструментария.

Единицей расчета необходимого числа интервьюеров может служить норма нагрузки — не более 10-15 интервью в неделю на одного интервьюера. Если проект исследования предусматривает опрос достаточно большого числа респондентов в сжатые сроки, то сокращение срока необходимо осуществлять за счет увеличения числа привлекаемых к работе интервьюеров.

Кого можно привлекать к работе в качестве интервьюера? Этот вопрос нередко обсуждается в методической литературе. Проблема пола, возраста, уровня образования и социального положения кандидатур нередко становится предметом дискуссий. Как правило, эти дискуссии носят умозрительный (теоретико-методический) характер. Например, обсуждаются вопросы о том, можно ли привлекать к данной работе студентов или пенсионеров, могут ли качественно выполнять данные функции лица с низким уровнем образования и т.п. Но опыт работы с интервьюерами показывает, что способности человека к данной работе часто не зависят ни от его формально-демографических характеристик, ни даже от внешнего впечатления о таких качествах, как

общительность, добросовестность, находчивость, выдержанность, интерес к данной работе (здесь перечислены характеристики, которые в методической литературе чаще всего называются в числе необходимых качеств интервьюера). Опыт показывает, что продемонстрировав эти качества организаторам исследования, человек часто не проявляет их в течение всего времени проведения опроса. Интерес к данной работе у некоторых интервьюеров, которые руководствовались преимущественно этим мотивом, может быстро иссякнуть в силу достаточной стереотипности этой работы; добросовестность может также оказаться недолговременной, особенно при отсутствии постоянного контроля; определять грань между общительностью и болтливостью также приходится «на глазок». Некоторые авторы считают необходимым элементом набора интервьюеров специальное тестирование. Здесь следует заметить, что это все же не столь ответственная и значимая работа (к тому же она, как правило, довольно кратковременная), чтобы подвергать кандидатуры громоздкому и дорогостоящему тестированию; тем более, что еще не созданы универсальные тесты профпригодности интервьюеров (возможно и потому, что в этом нет реальной потребности).

Способность к работе интервьюера лучше всего проверяется в процессе самой работы. Поэтому проблема организации работы интервьюеров в основном заключается не в критериях отбора, а в необходимости подготовки, подробного инструктажа, и, самое главное, контроле качества работы. Конкретные инструкции и требования, а также сам факт постоянного контроля побуждает большинство интервьюеров вполне добросовестно выполнять «выполнимые» требования. Интервьюеров, некачественно выполняющих работу, после одного-двух предупреждений следует отстранять от работы, не говоря уже о тех, в работе которых обнаружена фальсификация. Поэтому полевым директором обязательно должен располагать резервным списком интервьюеров.

Основные функции супервайзера заключаются: 1) в подготовке материалов исследования и списков интервьюеров по точкам опроса, подробном инструктаже (на подготовительном этапе); 2) мониторинге, инспекционном надзоре и контроле работы (на этапе сбора данных); 3) получении заполненных анкет, проверке качества заполнения и подготовки финансовых документов к оплате работы интервьюеров. Супервайзеры несут основную ответственность за качество работы интервьюеров. Обычно в исследовании работает от трех до десяти супервайзеров. Как правило, это штатные сотрудники учреждения, выполняющего исследование. Качество их работы прямо зависит от требовательности и организованности полевого директора, а уверенность определяется опытом подобной работы. Количество супервайзеров, необходимых для проведения сбора информации по конкретному проекту, определяется установленными сроками исследования и особенностями выборки*.

В тех случаях, когда выборка слишком разбросана (точки опроса находятся на достаточно далеком расстоянии друг от друга), и в каждой точке опросу подлежат менее 200 респондентов, организаторам исследования нецелесообразно набирать (или назначать) столько супервайзеров, чтобы каждый из них имел физическую возможность непосредственно контролировать процесс работы интервьюеров. В этих случаях проектом предусматривается оплата работы бригадиров. Бригадира обычно назначают супервайзеры из числа набранных интервьюеров. (Обычно функции бригадира поручают человеку, уже имеющему опыт работы интервьюера и положительно себя зарекомендовавшему). В случаях участия в сборе данных бригадиров, они фактически выполняют функции супервайзеров в компактно локализованных местах сбора информации.

Этапы работы

Можно выделить пять основных этапов процесса сбора данных:

- 1) подготовка;
- 2) инструктаж;
- 3) мониторинг и инспекционный надзор;
- 4) получение анкет;
- 5) подготовка массива анкет к обработке.

Подготовка

Подготовка к сбору данных заключается в назначении супервайзеров, подборе интервьюеров и бригадиров, определении конкретных функций и нагрузки супервайзеров. Подготовительный этап считается завершенным, когда каждый супервайзер составил полные списки интервьюеров, подготовил задание каждому интервьюеру по выборке (адреса или квоты), упаковал комплекты материалов (бланки анкет и письменных инструкций*), предварительно оповестил всех бригадиров и интервьюеров о дате начала исследования, дате и времени проведения инструктажа.

Инструктаж: принципы работы интервьюера

Идеальным вариантом инструктажа является общее собрание всех участников сбора данных: руководителя проекта, авторов инструментария, полевого директора, супервайзеров, бригадиров, интервьюеров. Если исследование является научным (основная цель — изучение определенных явлений, процессов, закономерностей и т.п.), не лимитировано жесткими сроками этапа сбора информации (и в силу этого в нем участвует не очень большое количество интервьюеров), то общее собрание всех участников исследования перед началом сбора информации является необходимым условием. Если же исследование проводится преимущественно с конкретно-практическими целями (опрос общественного мнения по какой-либо проблеме, например, электоральный выбор, рейтинг политических лидеров, отношение к средствам массовой информации, реакция населения на те или иные политические решения и т.п.), и сроки опроса достаточно сжаты, то инструктаж осуществляется в несколько этапов: руководитель проекта и полевой директор инструктируют супервайзеров, супервайзеры — бригадиров, бригадиры — интервьюеров. Несколько этапов инструктирования осуществляются быстрее и дешевле, чем подготовка к организации общего инструктажа. Однако в этом случае информация, касающаяся требований и особенностей опроса, всегда в той или иной мере претерпевает определенные искажения. Задача организаторов исследования свести эти искажения к минимуму. Для этого необходимо, чтобы инструктаж на всех этапах максимально соответствовал первоначальному сценарию. Обычно сценарий инструктажа состоит из трех частей: 1) ознакомление с общими требованиями к интервьюеру; 2) подробное разъяснение особенностей заполнения конкретного инструментария и; 3) тренинг интервью по данной анкете.

Общие требования к интервьюеру построены в соответствии с основными принципами сбора социологической информации.

1. Принцип «организованности» или «дисциплинированности»

Для обеспечения идентичности условий опроса, проводимого разными интервьюерами, каждый интервьюер должен строго соблюдать все инструкции, полученные перед проведением каждого конкретного опроса. В нестандартных ситуациях (в ситуациях, не предусмотренных инструкцией) или в случаях вынужденного нарушения инструкции интервьюер должен поставить в известность своего бригадира, супервайзера или руководителя исследования, фамилии, имена и телефоны которых ему должны быть известны.

2. Принцип «репрезентативности»

Интервьюер должен опрашивать только тех лиц, которые указаны в задании. Все замены, не предусмотренные инструкцией, производятся только по решению супервайзера или руководителя. Самостоятельные замены лиц, подлежащих опросу категорически запрещены. Если интервьюер оставляет анкету для самозаполнения, то при получении анкеты он должен уточнить — кто ее заполнял.

3. Принцип «конфиденциальности»

Интервьюер всегда должен помнить, что ему категорически запрещено разглашать свои впечатления от беседы с конкретными людьми. Свои впечатления, характеристику особенностей опроса организаторам исследования он представляет в отчете интервьюера только в обобщенном виде (без указания адресов и фамилий). В случае нарушения этого правила интервьюер лишается оплаты и исключается из списка интервьюеров.

4. Принцип «добровольности»

На протяжении всего опроса интервьюер должен быть вежливым, корректным и не забывать, что участие респондента в опросе всегда носит добровольный характер — респондент в любой момент имеет право отказаться отвечать как на какой-либо конкретный вопрос, так и от продолжения участвовать в опросе в целом.

5. Принцип «унификации информации»

При получении ответов на вопросы анкеты интервьюер (если нет дополнительной инструкции) не должен подсказывать или разъяснять респондентам смысл вопросов, значение слов или терминов, непонятных респонденту. Если респондент не понимает

вопрос, интервьюер должен повторить его еще раз, помедленнее, не изменяя формы вопроса.

Установление того факта, что респондент не может ответить на сформулированный таким образом вопрос, уже является достижением интервьюера, и представляет интерес для авторов исследования. (В одних случаях — это информация о том, что вопрос сформулирован неудачно и не дает качественной информации, в других — авторам исследования необходимо получить реакцию именно на данную формулировку вопроса). Главная задача интервьюера — зафиксировать реакцию каждого респондента именно на ту формулировку вопроса, которая представлена в бланке анкеты.

Все свои личные впечатления и комментарии по поводу формулировок вопросов интервьюер излагает после проведения опроса в отчете интервьюера. Как правило, эти замечания, комментарии, «живые» впечатления представляют ценный дополнительный материал для авторов исследования.

6. Принцип «своевременности»

Интервьюер должен своевременно (в срок, указанный супервайзером или бригадиром) сдать работу. В случаях возникающей возможности срыва сроков работы, интервьюер должен своевременно сообщить об этом супервайзеру или бригадиру исследования. Несоблюдение срока сдачи работы (без уважительных причин или без предупреждения бригадира) влечет за собой снижение оплаты за выполненную работу.

7. Принцип «качества»

Качество информации определяется качеством содержания информации и качеством заполнения анкет (качеством технического оформления).

Качество содержания. Качественной является информация, реально отражающая мнение тех людей, которые отвечали на вопросы анкеты. Одной из основных задач социологического сбора информации является проблема снятия возможного влияния интервьюера на характер ответов респондентов. Для того чтобы максимально снять эффект влияния интервьюера, необходимо соблюдать следующие правила.

1. Интервьюер не должен высказывать своего отношения к проблемам, содержащимся в анкете ни до, ни во время, ни после интервью. В ходе интервью не должно быть ни обсуждений, ни полемики между интервьюером и респондентом. Полностью безликое поведение интервьюера, естественно, невозможно, рекомендуется — приветливый взгляд, молчаливое одобрение любого высказывания, суждения и т.п.

2. Интервьюер не должен спешить с записью варианта ответа. Значительное число респондентов дает сходные ответы. Под влиянием стереотипа, сложившегося в ходе опроса ряда респондентов, интервьюер часто, не дожидаясь конца ответа, записывает тот, который ожидает; респондент, увидев, что ответ записан, решает, что так правильнее и не уточняет свою позицию.

Качество заполнения. Вся документация, сдаваемая интервьюером, (заполненные бланки анкет, бланки заданий, отчет интервьюера) должна быть оформлена в соответствии с требованиями, содержащимися в инструкциях интервьюеру. В первую очередь это относится к заполнению бланков анкет, так как исправить погрешности в этой работе довольно сложно. Качественно заполненной является анкета, в которой на каждый вопрос есть отмеченный (в соответствии с инструкцией) ответ или соответствующая пометка интервьюера о причине отсутствия ответа.

Оплата работы интервьюера производится после контроля качества работы по всему массиву опрошенных (примерно через 2-4 недели после сдачи работы). Если анкета заполнена не полностью или неправильно (не в соответствии с требованиями инструкции) оплата за данную анкету снижается пропорционально проценту брака.

Мониторинг и инспекционный надзор

Бригадир исследования ведет ежедневный учет проделанной работы: связывается с интервьюерами и уточняет, сколько человек уже опрошено, кто именно опрошен. Время от времени осуществляется выборочный контроль: в процессе сбора данных бригадир должен по каждому интервьюеру осуществить, по крайней мере, одну проверку.

Способы проверки принципиально не отличаются от проверки работы фирмы, в том случае, когда сбор данных заказывается специальным фирмам (см. соответствующий раздел).

Супервайзер, в свою очередь, поддерживает связь с бригадами и ведет журнал, в котором фиксирует результаты промежуточных устных отчетов бригадиров о выполненной работе. Отчет о работе каждого бригадира должен фиксироваться не менее двух раз в неделю. Если в сборе данных принимают участие довольно много бригадиров и работа проводится в сжатые сроки (1 неделя), то контроль осуществляется один раз — примерно в середине срока. В случаях обнаружения фальсификации (полной или частичной) бригадир должен заменить данного интервьюера, воспользовавшись резервным списком.

Получение анкет

Получение анкет также осуществляется по правилам, описанным выше. Бригадир проверяет полноту и качество заполненных анкет, в соответствии с которыми рассчитывает оплату работы интервьюера, предупреждая, что оплата будет произведена после проверки работы организаторами исследования. Бригадир получает от интервьюера отчет о работе, где указаны как общие впечатления интервьюера, так и конкретные пункты вопросов анкеты, вызывавшие особые трудности или непонимание респондентами смысла данных вопросов, отказы отвечать на данный вопрос и т.п. Бригадир окончательно оформляет трудовой соглашение (контракт). Если интервьюер производил отбор респондентов по квоте, он должен предоставить список опрошенных респондентов с указанием их адреса, телефона, имени и возраста. Если у бригадира возникают сомнения в добросовестности интервьюера, он может выборочно проверить качество его работы, так как с момента принятия материалов у интервьюера он принимает на себя полную ответственность за качество его работы. Работа бригадира завершается комплектацией

следующих материалов: 1) заполненные бланки анкет; 2) резервные бланки анкет; 3) отчеты интервьюеров; 4) списки респондентов; 5) трудовые соглашения; 6) отчет бригадира. Отчет бригадира включает описание точки опроса, выборки, характеристики работы интервьюеров, описание погрешностей, накладок, сбоев в работе. Весь комплект бригадир передает супервайзеру, уведомив предварительно о способе передачи и ожидаемом сроке получения.

Супервайзер должен визуально проверить все полученные материалы, отбраковать некачественно заполненные анкеты, отметить и отложить анкеты, вызывающие подозрение, и передать материалы полевому директору.

Полевой директор организует контроль работы интервьюеров. Контроль работы интервьюеров проводят супервайзеры. Лучше, если супервайзер контролирует не тех интервьюеров, работу которых он курировал, а других. Осуществляется контроль как анкет, вызвавших подозрение, так и выборочно 10-20% от общего массива анкет.

Подготовка массива анкет к обработке

После осуществления контроля и дополнительного опроса (необходимого в тех случаях, когда забраковано более 5% анкет от общего массива) все анкеты пронумеровываются. Подготовка массива анкет к обработке заключается в докодировке (например, кодировке открытых вопросов, точек опроса, кодов интервьюеров и т.п.) и перекодировке в случаях необходимости (если не были учтены возможности программных средств обработки или у авторов возникли новые гипотезы уже после того, как начался процесс сбора данных и т.д.). Процедуру подготовки анкет к вводу данных может осуществлять как группа организации сбора данных, так и группа обработки информации. В каждом конкретном случае этот вопрос решает руководитель проекта. Таким образом, этап сбора данных может считаться законченным после того, как массив качественно заполненных анкет пронумерован. После этого наступает этап обработки данных.

*Следует отметить, что характерной особенностью выборки, по которой, как правило, осуществляют сбор информации организации, является следующая схема: областной центр, города с прилегающими к ним сельскими поселениями и поселками городского типа с наиболее удобным транспортным сообщением. Конечно, это не исключает возможностей того, что организация возьмется за сбор информации в труднодоступных точках опроса, не типичных для ее обычного ареала деятельности, но тогда проблема целесообразности обращения в такую организацию должна обсуждаться отдельно.

**Поскольку в нашей практике слово «организация» имеет несколько значений, а «организации, специализирующиеся на сборе данных» слишком громоздко выглядит в тексте, для упрощения дальнейшего чтения будем называть эти организации «фирмами», хотя в нашей стране они имеют разные статусы.

*Многие из таких агентств по сбору социологической информации в США основаны людьми, которые начинали свою деятельность интервьюерами и обладают опытом и определенными практическими знаниями в этой области. Не имея специальных знаний в

области социологии и никогда не работая над другими этапами исследовательских проектов, они, тем не менее, хорошо усвоили основные правила и требования к такому этапу социологического исследования, как сбор первичной социологической информации. На основе своего опыта и знаний они организуют частные фирмы, занимающиеся организацией работы интервьюеров: подбором, подготовкой, контролем работы, заменами интервьюеров, оформлением сопутствующих документов — т. е. всеми необходимыми организационными аспектами. И если они добросовестно относятся ко всем требованиям авторов проекта, то в результате их работы исследователь может получить достаточно качественную информацию, не затрачивая усилий на работу с интервьюерами.

* Обеспечение интервьюеров ручками, бланками для дополнительных записей и т.п. обычно производит фирма, но если у исследователя есть такие финансовые возможности, то предоставление фирме таких аксессуаров, которые могут быть оставлены респондентам в качестве сувениров «на память об опросе», повышает эмоциональную заинтересованность респондентов и снижает процент отказывающихся участвовать в опросе.

* Респонденту обычно объясняют, что проводится контроль работы интервьюеров; если он будет проявлять беспокойство, что известны его адрес и телефон, ему можно объяснить, что эта информация содержится в его анкете только до момента окончания контроля работы интервьюеров; в компьютер его личные данные вводиться не будут, а после окончания ввода все бланки анкет вообще уничтожаются.

* В нашей стране супервайзеры обычно работают по географическому принципу — каждый из супервайзеров исследования курирует работу интервьюеров в определенном регионе. В США, например, где выше оснащенность различными средствами связи (телефон, телефакс и т.п.) и передвижения (личные автомашины), руководители исследования, наоборот, стараются разрушить «региональный» подход к организации связки «супервайзер — интервьюер», чтобы возможные личные контакты не препятствовали тщательному контролю работы интервьюера.

* Следует учитывать, что комплект должен предусматривать определенный запас бланков анкет и инструкций. Обычно целесообразно иметь 20% запасных бланков; при этом 10% передавать интервьюерам (или бригадирам), а 10% остается у полевого директора. Интервьюеры по окончании работы должны не только сдать заполненные бланки анкет, но и оставшийся запас, а в случаях его использования, отчитаться за утерянные или бракованные бланки.

Лекция 6.

Обработка данных

1. Подготовка к обработке данных

Планирование обработки

Обработка данных, как и другие стадии эмпирического социологического исследования, требует тщательной подготовки. Подготовку к обработке данных целесообразно осуществлять в два этапа: планирование и доработка программы обработки и анализа данных.

Хотя обработка данных начинается после поступления первого комплекта заполненных анкет, планирование и подготовка к этой процедуре должны быть проведены на более ранних стадиях исследования при разработке программы исследования.

Окончательную доработку программы ввода, обработки и анализа данных по конкретному исследованию целесообразно осуществлять в период полевого этапа сбора первичной социологической информации.

Планирование обработки включает в себя: 1) выбор компьютерной программы обработки; 2) точное определение методов статистического анализа, которые будут использованы применительно к каждому из изучаемых признаков (вопросов анкеты).

Выбор компьютерной программы обработки и анализа данных

В настоящее время в Украине используются два основных пакета программ обработки и анализа первичной социологической информации: SPSS и OCA. Программа SPSS является наиболее распространенной программой анализа первичной социологической информации на международном рынке социологических исследований. Данная программа включает в себя основные методы статистического анализа, используемые в работе с первичной социологической информацией. Кроме того, именно этот формат является наиболее распространенной формой обмена социологической информацией с международными партнерами.

Программа OCA (Обработка Социологических Анкет), разработанная киевским ученым А.П.Горбачиком, является основной компьютерной программой анализа первичной социологической информации, используемой практически всеми ведущими социологическими центрами Украины. Эта программа, так же как и SPSS, включает в себя основные методы статистического анализа первичной социологической информации. Она компактна, удобна в работе, легко осваивается; в силу чего получает все большее распространение не только в Украине, но и в ряде стран «ближнего зарубежья» (России, Литве, Белоруссии и т.д.). Существенным достоинством этой программы является ее совместимость с программой SPSS: данные, введенные в компьютер с помощью программы OCA, при необходимости легко переводятся в формат SPSS.

Профессиональным социологам, прежде чем приступать к эмпирическому социологическому исследованию, необходимо, на наш взгляд, освоить оба или, по крайней мере, один из этих пакетов программ обработки социологических данных.

В тех случаях, когда заказчик эмпирического исследования, не являясь профессиональным социологом, выступает лишь в роли потребителя социологической информации, а обработку и анализ заказывает специалистам, то при подготовке к исследованию, на этапе обсуждения проекта, он должен выяснить — какая программа обработки данных будет использована. Эта информация будет ему полезна, во-первых, для оценки степени профессионализма исполнителей обработки и анализа информации (бывают случаи, когда к анализу первичной социологической информации привлекаются люди, не имеющие достаточной профессиональной подготовки; и в таких случаях подобный вопрос является в определенной степени тестом на уровень профессиональной

компетентности). Во-вторых, информация об используемой программе обработки данных может заказчику пригодиться в дальнейшем использовании полученного в результате социологического опроса массива данных; поэтому лучше, если исполнители используют достаточно распространенные (отработанные и апробированные) программы обработки социологических данных.

Планирование методов статистического анализа

Определение методов социологического анализа на этапе планирования исследования является, пожалуй, основной характеристикой уровня профессионализма социолога, который, проектируя эмпирическое исследование, не забывает, что последующий **анализ является не более чем математико-статистической проверкой гипотез**, заложенных на этапе подготовки программы и инструментария (анкеты) конкретного социологического исследования. К сожалению, слишком часто встречаются опубликованные данные, свидетельствующие о том, что автор при планировании не только не учитывал последующую обработку и методы анализа, но даже и не догадывается о том, что он приводит данные, недостоверность которых слишком очевидна.

Поскольку, на наш взгляд, именно по этой причине (несоответствие методов анализа возможностям, заложенным в программе исследования) возникают основные ошибки и погрешности выводов, интерпретации и широкого публичного освещения полученного материала, остановимся подробнее на проблеме планирования методов анализа и разборе ошибок, которые наиболее часто бросаются в глаза в публикуемых результатах социологических опросов.

В тех случаях, когда исследователь приступает к опросу, не имея программы анализа полученных данных, обычно в процессе спонтанного анализа возникают трудности (у квалифицированного социолога) или ошибки (у неквалифицированного социолога) по двум основным причинам: 1) социально-демографические группы, которые интересуют исследователя, слишком малочисленны для возможности статистической проверки различий, которые они показывают в своих ответах на поставленные в анкете вопросы; 2) шкала (перечень ответов на вопрос), предложенная в анкете, не позволяет применять эффективные методы анализа, так как в период подготовки анкеты автор не задумывался о том, каким именно образом он будет анализировать полученные результаты.

Наиболее часто в средствах массовой информации встречаются ошибки, допускаемые непрофессиональными социологами по первой причине авторы социологических опросов приводят результаты сравнительного анализа социальных групп, численность которых в выборке явно недостаточна для уверенности в том, что полученные различия статистически значимы.

Например, автор исследования, посвященного отношению населения к экономическим реформам, пишет, что «молодежь в возрасте от 18 до 20 лет значительно чаще поддерживает радикальные рыночные реформы, чем другие возрастные группы населения, поскольку, если среди населения в среднем 36% выразили поддержку радикальным рыночным преобразованиям в экономике страны, то среди молодежи 18-20 лет 53% высказались за подобные преобразования», при этом в преамбуле к приведенным данным и соответствующему выводу сообщается, что исследование проводилось по репрезентативной выборке взрослого населения г. Киева, и в общей сложности было

опрошено 400 человек. В данном случае сам по себе вывод не вызывает ни удивления, ни, тем более, принципиальных возражений, поскольку не противоречит ни жизненному опыту читателя, ни распространенным стереотипам общественного мнения. Однако в какой мере, высказывая подобный тезис, правомерно сослаться на результаты данного конкретного исследования? Даже не владея точной статистической информацией о демографическом составе населения г.Киева, внимательный читатель может предположить, что лица в возрасте 18-20 лет составляют 2-3% взрослого населения. Поскольку автор указал, что выборка была репрезентативной, то, следовательно, в числе опрошенных им людей возрастная группа 18-20 лет включает 8-12 человек. Даже неискушенному в социологии читателю ясно, что 8-12 человек вряд ли могут представлять молодежь Киева данной возрастной группы в целом.

Мы не будем здесь останавливаться на подробном изложении того, как надо было планировать и строить выборку, если бы автора исследования действительно на этапе проектирования интересовал сравнительный анализ именно данной возрастной группы населения (этот вопрос подробно изложен в разделе, посвященном построению выборки). Однако с большой долей вероятности можно предположить, что подобное приведение данных является результатом не ошибки расчетов на этапе планирования исследования, а отсутствия планирования анализа вообще. К сожалению, в эмпирической социологии все большее распространение получает практика анализа социологических данных, когда заказываются двумерные таблицы, содержащие демографические данные — на все вопросы анкеты, а «социолог» просматривает их все подряд, выбирая таблицы, в которых «бросается в глаза» разброс данных. После этого начинается объяснительный «анализ» и логическая интерпретация. Не случайно в тех случаях, когда автор публикаций результатов опросов общественного мнения касается проблемы различий в ответах социально-демографических групп, чаще всего он обращает внимание на такие социальные группы, как «предприниматели», «военнослужащие», «домохозяйки». Как правило, именно эти группы по своим ответам «выбиваются» из общего ряда распределения. Но собранный материал, обычно, не позволяет сделать выводы о том, являются ли отклонения в их ответах особенностями сознания данных социальных групп населения, или это результат малочисленности этих групп в репрезентативной выборке населения.

Даже не владея знаниями в области математической статистики, но проявив необходимый для научного анализа здоровый скептицизм, можно догадаться, что если на поставленный в анкете вопрос (имеющий 3-5 вариантов ответа) отвечали 20-30 человек, то определенный вариант могли отметить 4-10 человек. В таком случае даже случайная ошибка только одного респондента (если он невнимательно отвечал на вопрос), или интервьюера, который его опрашивал, или оператора, который вводил данные в ЭВМ, изменяет результаты на 10-25%. Именно поэтому немногочисленные группы населения резко выбиваются при сравнительном сопоставлении полученных данных, и именно поэтому они привлекают внимание неискушенного социолога при том методе анализа, который упоминался выше, когда исследователь, не имея ни гипотез, ни плана анализа, заказывает сотни таблиц и «смотрит» — какие именно результаты представляют интерес. Фактически из этого «творческого процесса» автоматически исключается главный компонент анализа, ради которого проводился выборочный опрос — статистическая проверка гипотез. В приведенном примере, как минимум, требовалось проверить — является ли различие между 53% и 36% статистически значимым с учетом численности опрошенных.

Мы умышленно в качестве примера некомпетентного анализа привели иллюстрацию с достаточно «безобидными» в социально-политическом и научном отношении выводами (в

данном случае, если бы исследование и анализ были проведены достаточно грамотно, автор скорее всего — с теми или иными различиями в числовых данных — пришел бы к тем же выводам). К сожалению, во многих случаях необоснованные выводы не столь безобидны с социально-политической точки зрения (например, один из авторов политологического анализа ситуации в Чечне во время проведения военных действий Российским правительством на территории Чечни, ссылаясь на данные социологического опроса ВЦИОМ, характеризует бедственное положение русских в Чечне в период, предшествующий развязыванию войны). В публикации, несомненно имеющей острое политическое и гражданское значение, не приводится необходимая описательная характеристика исследования, на данные которого ссылается автор анализа. Однако описание исследований, приводимое этим социологическим центром в других публикациях, позволяет предположить, что группа «русские, проживающие в Чечне» слишком малочисленна во всероссийской выборке, чтобы можно было делать столь многозначительные выводы, в определенной степени направленные на оправдание силовых акций. В данном случае спекуляция статистически не обоснованными цифрами имеет слишком неприглядный конъюнктурно-политический характер.

Два примера, приведенные выше, различаются по своему политическому значению, но они сходны в одном — авторы, приводя данные социологических исследований, руководствовались статистически не обоснованными данными. Необоснованность в данном случае определяется прежде всего тем, что на этапе подготовки к исследованию приводимый анализ (сравнение малочисленных в общем объеме генеральной совокупности групп населения с другими) не планировался автором.

Ошибки другого рода, когда при подготовке анкеты, формулируя ответы на вопрос, автор не задумывается, какие методы анализа он будет применять к полученным результатам опроса, встречаются реже, так как непрофессиональные социологи редко используют удобные и эффективные методы анализа, ограничиваясь представлением процентных распределений. Однако в последнее время появились публикации с грубыми ошибками, вызванными неправомерными и, естественно, непредусмотренными манипуляциями со шкалами ответов. Чаще всего это связано с острыми и вызывающими всеобщий интерес данными, касающимися рейтингов политических деятелей. Достаточно распространенным является следующий подход к определению рейтинга. Сначала автор, не задумываясь о том, каким образом он будет не только анализировать, но и представлять данные, предлагает в анкете вопрос со следующим веером вариантов ответов: «В какой степени Вы доверяете политическому лидеру N?»

- 1 — совершенно не доверяю
- 2 — в определенной степени не доверяю
- 3 — в определенной степени доверяю
- 4 — полностью доверяю
- 5 — затрудняюсь ответить

В дальнейшем, при описании данных, он, в лучшем случае, публикует результаты опроса именно в таком виде. Но исходная форма приведения данных представляет наименьший интерес, поскольку рейтинги интересны в сравнении (или это мониторинговое сравнение — изменение степени доверия конкретному лидеру в течение какого-либо промежутка времени, или сравнение рейтингов различных политических фигур). Данные

же, представленные в таком виде, неудобно сравнивать, потому что визуально сравниваются лишь крайние группы («полностью доверяю» или «совершенно не доверяю»). Исходные данные надо перерассчитывать: (складывать процент «полностью доверяющих» с «доверяющими в определенной степени» и ту же процедуру проводить с недоверяющими. Но если автор исследования не планировал ничего иного, кроме представления разделения опрошенных на доверяющих и не доверяющих, то зачем было предлагать пять вариантов ответа, когда уже в анкете их можно было преобразовать в три группы и сразу получить материал, подготовленный к такого рода анализу. В данном случае исследователь (или читатель) сталкивается лишь с лишней работой.

Значительно хуже, когда исследователь, недостаточно подготовленный к социологическому анализу, пользуясь возможностями программы обработки и анализа социологических данных, но не понимая сути и содержания операций с числовым материалом, заказывает по сформулированному именно таким образом вееру ответов подсчет индекса — среднего арифметического полученных данных. Он не осознает, что в этом случае порядковый номер варианта ответа является баллом (в приведенном примере — выражающим степень доверия), и, по его шкале, вариант ответа «затрудняюсь ответить» при подсчете среднего значения засчитывается как 5 баллов. Эта ошибка (с различными вариантами) является достаточно распространенной, когда анализ результатов социологического исследования проводят лица, недостаточно знакомые с азами статистического и социологического анализа.

Своевременное планирование обработки и анализа данных позволяет профессиональному социологу более эффективно подготовить анкету, а социологу с недостаточным опытом — своевременно осознать проблемы и трудности, с которыми он столкнется при обработке и анализе данных. Планируя анализ, он может проконсультироваться со специалистом, или упростить анкету, осознавая, что его основная задача — получение общих распределений результатов опроса общественного мнения; поэтому, получив данные, не будет углубляться в анализ, с математическими основами которого он недостаточно хорошо знаком. В этом случае ему следует помнить — чем меньше вариантов ответов на альтернативный вопрос (напомним, что альтернативным называется вопрос, на который можно дать только один вариант ответа), тем более наполненными будут полученные группы, и, соответственно, больше вероятность, что обнаруженные различия в ответах различных групп населения будут статистически значимыми. Следует, однако, заметить, что сокращать число ответов следует не механически, а обобщая и укрупняя близкие по значению варианты ответов таким образом, чтобы полученные в конечном результате варианты в достаточно обобщенном виде охватывали все множество возможных ответов. Например, в приведенной выше иллюстрации нельзя просто изъять какой-либо (или какие-либо) варианты ответов, а следует обобщить близкие по значению, преобразовав веер ответов в следующий вид:

- 1 — в общем не доверяю
- 2 — трудно сказать, доверяю или нет
- 3 — в общем доверяю

В любом случае планирование анализа на этапе проектирования исследования поможет социологу осознать свои возможности и достаточно корректно ограничить свои задачи.

Планирование анализа включает в себя:

- 1) формулирование гипотез, которые будут проверяться на основании эмпирического исследования;
- 2) оценку каждого пункта (вопроса) анкеты под углом зрения возможности математико-статистической проверки гипотезы адекватными методами.

Планирование анализа на этапе проектирования исследования позволяет еще до начала опроса, во-первых, откорректировать выборку; а во-вторых, отредактировать анкету таким образом, чтобы максимально эффективно и адекватно возможностям исследователя использовать полученные данные.

Своевременно сформулированные гипотезы позволяют исследователю откорректировать выборку таким образом, чтобы при анализе демографические группы, характеристика которых, по мнению автора, является фактором, влияющим на изучаемое явление, были достаточно наполнены, чтобы можно было проводить статистический анализ (по общепринятым в эмпирической социологии стандартам, численность таких групп должна быть 50-100 человек — будь то «молодежь в возрасте 18-20 лет», или «предприниматели», или «военнослужащие», или другие социально-демографические группы, численность которых в обычной репрезентативной выборке, как правило, недостаточна для того, чтобы можно было проводить корректный, со статистической точки зрения, сравнительный анализ) с другими социальными группами по их отношению к различным социальным явлениям.

Если при подготовке анкеты исследователь основное внимание уделяет вопросам (их содержанию, формулировкам, структуре, последовательности), то при планировании анализа в поле его зрения главным компонентом становится шкала (веер ответов). Здесь необходимо, во-первых, определить тип шкалы (номинальная, порядковая, метрическая), а затем критически оценить, соответствуют ли варианты ответов типу шкалы. Другими словами, при планировании анализа каждый пункт предварительно подготовленной анкеты необходимо оценить с точки зрения определения методов его анализа и представления в заключительном научном документе, учитывая как собственные возможности (уровень подготовки), так и возможности программы обработки данных, которая будет использована.

Окончательную подготовку к обработке и анализу данных целесообразно проводить в период полевого этапа исследования (сбора интервьюерами первичной социологической информации).

После получения первого комплекта заполненных анкет можно начинать пробный ввод данных по подготовленной программе. Это лучше начинать до получения всего массива анкет, так как довольно часто даже у опытных социологов после начала ввода анкет обнаруживаются некоторые недоработки программы ввода. На этом этапе (пока продолжается полевой этап по сбору данных) еще есть время и возможности откорректировать паспорт ввода или разработать дополнительные инструкции по перекодировке данных.

Мониторинг получения заполненных анкет

Качество первичной социологической информации на полевом этапе сбора данных прямо зависит от тщательности контроля этого процесса исследователем. Вся получаемая

информация должна фиксироваться в журнале руководителя-организатора опроса. Естественно, что ежедневный контроль существенно повышает нагрузку и финансовые расходы (в основном на телефонные переговоры); однако эта мера позволяет гарантировать наиболее качественную работу на самом ответственном этапе работы, ради которого, собственно, и готовилось эмпирическое исследование. Если после окончания эмпирического исследования обнаруживаются недостатки в выборке или в инструментарии, исследователь может изменить интерпретацию полученного материала в соответствии с реальными особенностями выборки или инструментария, но если погрешности были допущены во время сбора данных, и они остались вне поля зрения исследователя, он получит искаженный материал, даже не подозревая об этом. Поэтому на усилиях по контролю качества работы интервьюеров не следует экономить ни трудовые, ни финансовые ресурсы. Опыт показывает, что ответственность интервьюера прямо пропорциональна контролю, во-первых потому что большинство людей не идут на нарушение инструкций, зная, что это станет известно, а во-вторых, каждый интервьюер работает более добросовестно, если он знает, что и другие работники выполняют свою работу качественно, в таком случае мысль о том, что его личные усилия не напрасны, поддерживает его терпение в те моменты, когда добросовестное выполнение инструкций сопряжено с существенным увеличением психоэмоциональной нагрузки.

Кроме повышения гарантии качества работы, постоянный мониторинг позволяет принять своевременные меры в случаях существенных сбоев или срывов работы, которые могут происходить в некоторых точках опроса. Журнал организатора исследования представляет собой по сути протокол проведения опроса, позволяющий отслеживать, сколько анкет находится «в поле», сколько получено заполненных анкет супервайзерами и где именно происходят задержки с получением анкет. Такой «текущий протокол» сбора данных позволяет исследователю не только своевременно обнаружить, но и предвидеть «выпадающие» точки опроса. В таком случае исследователь имеет время и возможность принять решение (оказание дополнительной помощи, замена интервьюера, супервайзера, изменение точки опроса и т.п.).

2. Процедура обработки анкет

Ручная обработка

Когда получен весь массив заполненных анкет, к передаче его для ввода в ЭВМ, необходимо осуществить ряд процедур, требующих ручной обработки. Прежде всего, это сквозная нумерация массива анкет.

Каждая (из заполненных и полученных организатором опроса) анкет представляет собой документ, который должен быть пронумерован. Следует напомнить, что анкеты нумеровались и супервайзером, по мере их получения от интервьюеров, но это была промежуточная нумерация отчетных документов супервайзера. На этапе получения заполненных бланков анкет руководителем-организатором исследования каждая анкета получает свой номер базового источника первичной социологической информации; этот номер обязательно вводится в ЭВМ вместе со всей информацией, содержащейся в анкете.

Второй операцией, требующей, как правило, ручной обработки, является кодировка (докодировка, перекодировка). Ручная кодировка требуется в следующих случаях:

а) в анкете содержатся открытые или полуоткрытые вопросы;

б) в процессе сбора данных обнаружена необходимость в перекодировке (например, в анкете указан населенный пункт, в котором проводился опрос, а в массив данных вводится общее название региона проживания и т.п.);

в) в паспорт необходимо ввести дополнительную информацию (фамилию интервьюера, заполнявшего анкету; фамилию супервайзера и т.п.).

Руководитель должен заранее подготовить инструкцию по кодировке данных. Следует заметить, что даже в тех случаях, когда перед кодировщиком не ставится специальная задача проверки качества заполнения, часто именно на этом этапе работы, лица, занимающиеся ручной кодировкой данных, обнаруживают брак в заполнении анкет. Это происходит потому, что супервайзеры, которые обязаны проверять работу интервьюеров, часто принимают основной массив анкет в сжатые сроки (нередко в момент приема анкет супервайзером от одного интервьюера в помещении присутствуют другие интервьюеры, которые ждут своей очереди); в подобной ситуации даже добросовестные супервайзеры, принимая от интервьюеров анкеты, в основном обращают внимание на наличие ответов на все пункты опросного листа; другие же погрешности (противоречивые сведения, ошибки заполнения — отметка двух вариантов ответов там, где должен быть отмечен только один вариант и т.п.) остаются незамеченными.

Кодировщик работает в более спокойной обстановке, поэтому он нередко обращает внимание на логические погрешности в заполненной анкете. Поэтому руководителю целесообразно просить кодировщиков отмечать номера тех анкет, которые вызвали подозрение. При выборочном контроле работы интервьюеров эти анкеты включаются в массив, подлежащий проверке.

Поэтому в рамках организации подготовки анкет к вводу в ЭВМ целесообразно функции кодировщика и функции редактора анкет поручить одним и тем же исполнителям.

Редактирование заполненных анкет

В задачу редактора входит:

а) проверить заполненность анкет;

б) правильность заполненности;

в) приведение в соответствие различных форм бланков анкеты в тех случаях, когда опрос проводится по нескольким формам анкеты.

Оценка заполненности анкет

Нередко анкеты, полученные в результате опроса заполнены не по всем пунктам. Если напротив вопроса, который остался без ответа, есть пометка или комментарий респондента или интервьюера, то тогда пропуски в заполнении анкет следует отнести за счет особенностей инструментария. Эти анкеты могут быть отмечены и после ввода информации в ЭВМ (отложены отдельно для специального анализа автором программы исследования). Имеющиеся в них пометки и комментарии позволяют глубже понять

чувства и эмоции, проанализировать мысли и сомнения, которые испытывали респонденты, отвечая на вопросы анкеты.

Но даже те анкеты, которые при беглом осмотре показались полностью заполненными, должны быть проверены тщательно. Редактор проверяет каждую страницу и каждый блок вопросов, чтобы быть уверенным, что респондент выполнял все инструкции и кодировал ответы на вопросы в соответствующем месте и соответствующим способом. Автор исследования устанавливает квоту заполненности (процент допускаемых пропусков) и сообщает кодировщику, выше какой квоты неотвечен анкета бракуется и изымается из массива данных. Для этого автор программы исследования подготавливает письменные инструкции по кодировке и редактированию анкет. Это желательно делать в любом случае, но совершенно необходимо, когда к кодировке и редактированию подключается несколько помощников. В этом случае они должны руководствоваться совершенно идентичными указаниями.

В итоге редактирования, массив проверяемых анкет делится на три группы: 1) принятые в обработку; 2) отбракованные; 3) вызывающие сомнения. Эта информация передается автору исследования, и он решает, что делать дальше. Возможно, что даже анкеты, не выдерживающие квоту заполненности, по решению ответственного исполнителя проекта, могут быть приняты в обработку с соответствующей кодировкой пропусков (отметкой — нет ответа). Полностью следует отбраковывать недозаполненные анкеты в том случае, если в начале респондент заполнял ответы на вопросы, а потом прекратил, другими словами, — анкеты, заполненные в начале бланка и не заполненные к концу.

Редактирование анкет с вопросами-фильтрами

Если в анкете использовались вопросы-фильтры, редактор должен особенно внимательно проверить, чтобы на эксклюзивные вопросы отвечали только те респонденты, которые соответствуют фильтру. Например, вопрос-фильтр дифференцирует тех, кто работает в настоящее время, и тех, кто в настоящее время не работает. Затем следует инструкция: «На вопросы NN... отвечают те, кто в настоящее время работает; неработающие должны перейти к вопросу N...». Редактор должен убедиться, что респонденты, которые в вопросе-фильтре отметили, что они не работают в настоящее время, не заполняли блок вопросов, касающийся работающих. Такая ситуация вполне реальна, так как респонденты и интервьюеры (что случается значительно реже) не обратили достаточного внимания на инструкцию, и вполне могут фиксировать ответы на вопросы, касающиеся отношения к работе или характеристики рабочего места, имея ввиду предыдущую работу, которой респонденты занимались до ее прекращения. Задача редактора в подобном случае вычеркнуть их ответы в бланке анкеты, чтобы операторы ЭВМ не вводили эту информацию в компьютер. В этих случаях редактор в своем протоколе отмечает подобные анкеты, и эта информация будет полезна для руководителя-организатора полевого этапа работы, как свидетельствующая о недостаточно высоком качестве работы супервайзеров, плохо проинструктировавших интервьюеров.

Перед началом редактирования следует обратить внимание редактора, что вопросы-фильтры могут быть не прямыми, а косвенными, например: «Собираетесь ли Вы голосовать на предстоящих выборах?»; «Есть ли у Вас родственники, проживающие вне пределов Украины?» и т.п. В тех случаях, когда следующий вопрос начинается со слова «если» (например, «Если Вы собираетесь голосовать, то за кого именно?») или «Если у Вас есть родственники, проживающие за пределами Украины, то где именно?»), то

предыдущий вопрос является своего рода фильтром, и на последующий должны ответить только те респонденты, которые дали на него утвердительный ответ. Редактор должен на такие вопросы обратить особое внимание, так как их следует подготовить к адекватному вводу в ЭВМ, а также предоставить информацию о качестве работы интервьюеров. Обычно не следует вводить данные в таком виде в файл информации, надеясь на то, что подобные ошибки будут устранены позднее (на этапе анализа информации).

Опыт показывает, что наиболее эффективным путем редактирования является предоставление каждому из редакторов бланка анкеты, со всеми необходимыми инструкциями и отметкой (обводящей линией), тех блоков вопросов, которые следуют после фильтра. Этот бланк анкеты является своего рода «ключом», помогающим проверять качество каждой заполненной анкеты. Редактор может сравнивать страницы проверяемой анкеты, сопоставляя ее с «ключевой»; которая особенно помогает редактору в начале работы, когда структура анкеты не является для него достаточно знакомой и легко узнаваемой. Обычно после первой порции проверенных анкет у редактора отпадает необходимость сопоставлять проверяемую анкету с ключевой, так как эталон проверки уже достаточно прочно закрепляется в его памяти.

Редактирование заполненных анкет — очень ответственный этап работы, так как структура ответов, полученных на полевом этапе, претерпевает определенную трансформацию. Некоторые исследователи предпочитают отказаться от этого этапа, надеясь получить в файле информации полный «слепок» того, что было сделано «в поле» во время непосредственного опроса, а лишь потом самостоятельно разобраться в тех погрешностях, которые были допущены интервьюерами и респондентами. Однако опыт показывает, что в таких случаях корректировка нередко осуществляется операторами, которые вводят информацию в ЭВМ: в одних случаях ее просто невозможно ввести в том виде, в каком она представлена в анкете (например, когда требуется лишь один вариант ответа, а респондент отметил два или более вариантов); в других случаях операторы относятся слишком инициативно к своей работе, и если видят противоречивую или неправильную, с их точки зрения, информацию, то сами корректируют ее по своему усмотрению. В этих случаях автор исследования утрачивает контроль над теми трансформациями, которые претерпевают полученные в результате опроса данные, так как каждый из операторов по ходу своей работы принимает индивидуальные решения. Рациональнее осуществить редактирование централизованно, в соответствии с унифицированными инструкциями под руководством главного научного исполнителя проекта; тогда внимание оператора будет сосредоточено лишь на точности переноса информации с бланка анкеты в файл данных.

В заключение следует подчеркнуть, что в последнее время развиваются компьютерные технологии ввода данных, при которых результаты опроса считываются компьютером непосредственно с бланков анкет. И если предварительное редактирование при вводе информации в файл с помощью операторов позволяет повысить качество и эффективность работы в целом, то при автоматизированном считывании информации процедура редактирования является обязательным этапом, без которого обработка данных невозможна.

Ввод данных

Когда массив заполненных анкет обработан по всем правилам (проверен, зарегистрирован, произведены перекодировка и редактирование), ввод данных

представляет собой процедуру считывания оператором кода каждого из пунктов анкеты и ввода этого кода в файл информации в соответствии с форматом паспорта, подготовленного на основе требований используемой программы ввода и обработки данных социологической информации.

Обычно эту процедуру осуществляют лица, специализирующиеся на подобных операциях. Сама по себе, эта процедура несложная и требует лишь внимания и добросовестного отношения: чем больше опыт работы оператора, тем меньше времени затрачивается на этот этап работы. При достаточном числе опытных операторов процедура ввода данных может быть осуществлена в минимальные сроки.

Поэтому процедуру ввода информации исследовательскому коллективу лучше не осуществлять собственными силами (если в его распоряжении нет соответствующего штата опытных операторов), а заказывать службам, специализирующимся на подобной работе.

В качестве наиболее рационального подхода к процедуре обработки данных может быть рекомендован следующий баланс между затратами труда собственно исследовательской группы и привлеченных лиц (или организаций): ручная подготовка к вводу данных должна осуществляться собственными силами под руководством и при непосредственном участии руководителя исследовательского коллектива; ввод данных в ЭВМ лучше заказывать организациям, специализирующимся на подобного рода работе.

Когда исследователь заказывает ввод данных специальной организации, его участие в дальнейшей работе должно состоять в следующем:

- 1) договариваясь с руководителем группы ввода информации, он должен передать ему «ключевую» анкету, с помощью которой работали редакторы, проверяющие качество заполнения; в эту «ключевую» анкету исследователь около каждого пункта-вопроса должен внести пометку о типе шкалы (номинальная, порядковая, метрическая) в соответствии с теми методами статистического анализа, которые он собирается использовать;
- 2) после подготовки паспорта ввода данных исследователь должен внимательно ознакомиться с распечаткой паспорта; внести необходимые уточнения и корректировку в тех случаях, если какие-либо детали его смущают или вызывают несогласие;
- 3) после осуществления ввода данных необходимо произвести контроль качества ввода.

Проверка качества ввода осуществляется в основном методом повторного ввода.

В тех случаях, когда исследователь располагает достаточными временными и финансовыми ресурсами, он обычно заказывает повторный ввод всего массива данных. В этом случае все анкеты, в которых выявлено несоответствие между результатами первого и повторного ввода, вводятся еще раз (пока не будет достигнута полная согласованность). Средства, затрачиваемые на повторные вводы можно при этом сэкономить на штрафных санкциях по отношению к операторам, допустившим ошибки. Однако при ограниченных сроках на проведение исследования подобная тщательность, к сожалению, трудноосуществима. Поэтому на практике часто используются сокращенные варианты контроля качества ввода: повторный ввод части анкет (10%, 20% и т.д.); проверка ввода нескольких пунктов анкеты (особо важных, с точки зрения исследователя) по всему массиву и т.д. По результатам такой выборочной проверки подсчитывается процент брака.

Как должен поступать исследователь, обнаружив определенный процент брака на этапе ввода информации? Это зависит от задач, целей и условий проводимого исследования. В случаях, когда сроки исследования позволяют, можно при определенном превышении процента брака (например 1%) потребовать произвести повторный ввод. Если исследователь предполагает подобную меру, он должен оговорить это условие в трудовом соглашении (контракте), заключаемом с руководителем ввода информации перед началом работы. Если же сроки исследования сжаты (например, в опросах общественного мнения по вопросам, имеющим крайне актуальное значение), исследователь может при анализе просто учитывать погрешность ввода, и, определяя качество информации, перерасчитывать совокупную ошибку. В таких случаях, проводя анализ информации, исследователь должен в своих выводах быть более корректным, менее категоричным, делать дополнительную поправку на процент брака, допущенного при вводе информации.

Существуют и другие методы контроля качества ввода информации в ЭВМ; в большинстве случаев они связаны с особенностью программы, используемой для обработки социологической информации.

Статистическая обработка данных

Результаты опроса населения, введенные в компьютер, сами по себе являются «сырыми» данными, представляющими формализованные сведения, как тот или иной респондент ответил на вопросы анкеты. Подобного рода информация может представлять лишь индивидуальный интерес, например, для отдела кадров или руководителя организации, принимающего на работу конкретного индивида. Но для понимания социальных процессов или явлений сырые данные не являются информацией в полном смысле этого слова. Различие между «данными» и «информацией» заключается в том, что сами по себе данные лишены интерпретационного смысла. Процесс их систематизации и представления в виде, имеющем смысловое значение, называется обработкой данных.

Таким образом, суть обработки данных заключается в переходе от «сырых данных» к «социальной информации». Задача исследователя — определить пути такого преобразования полученного материала и такой формы его представления, которые в конечном итоге позволят отвлечься от конкретных ответов отдельных людей и установить наиболее важные и значимые закономерности и связи, обнаруживаемые в собранном материале.

В настоящее время, когда социологи для обработки данных используют персональные компьютеры и соответствующие программы обработки и анализа социологических данных, нередко возникает ситуация, когда исследователь, получив файл данных, даже не задумывается — каким образом ему следовало бы систематизировать и структурировать собранный материал. Он автоматически заказывает предусмотренную программой обработку и получает результат в том виде, который заложен разработчиками компьютерной программы. Как правило, такой «исследователь» ограничивается одномерными и двумерными рядами распределений (иногда он даже не знает, что форма систематизированных таким образом данных называется рядом распределения).

Когда уровень технического развития не позволял автоматически использовать готовые программы обработки и анализа социологических данных, исследователь даже не приступал бы к сбору информации, пока не подготовил полный алгоритм их обработки (в основном ручной, с некоторыми элементами возможной механизации), потому что в

противном случае эти данные являлись бы лишь архивом анкет, на сбор которых не целесообразно было бы затрачивать средства и силы. За последнее десятилетие обработка данных претерпела принципиальную трансформацию в связи со все расширяющимися возможностями электронной техники. У исполнителя отпала прежняя необходимость в подготовке полного алгоритма обработки полученных данных. Широкий доступ к персональным компьютерам и современное программное обеспечение позволяют даже лицам, не имеющим практически никакой специальной подготовки, преобразовывать сырой материал в определенный тип информации.

Существуют различные точки зрения на проблему необходимой степени подготовленности социолога к обработке и анализу первичных социологических данных, уровня его квалификации в области математической статистики, лежащей в основе социологического анализа эмпирических данных.

Одна точка зрения определяется парадигмой профессиональной корпоративности, направленной на утверждение того, что человек, плохо разбирающийся в значении и содержании математических формул, по которым компьютерные программы производят статистические расчеты, не должен заниматься первичным социологическим анализом.

Другая точка зрения заключается в том, что необязательно досконально знать математико-статистический аппарат, так же как и принципы работы компьютера, чтобы пользоваться разработанными программами.

Если аргументы сторонников первой точки зрения не нуждаются в особом раскрытии, так как представляется достаточно понятной позиция, что любое занятие требует определенной квалификации, и большинство людей предпочтет обратиться к профессионалу, когда возникает потребность в выполнении работы, требующей определенной профессиональной квалификации.

Но и вторая позиция не остается без аргументов, имеющих определенный практический смысл. Обосновывая ее, Памела Алрек и Роберт Сеттл (сотрудники одной из американских консультационных служб) обращаются к аналогии с ситуацией авиаперелета: «Никто ведь не будет утверждать, что человек, собирающийся лететь самолетом в отдаленный город, должен быть специалистом в области аэродинамики и самолетостроения. Большинство авиапассажиров имеет весьма смутное представление о том, как устроен самолет, и уж, определенно, не смогут поднять его в воздух, рассчитать курс и т.д. Тем не менее они способны осуществить перелет и прибыть в место назначения достаточно быстро и безопасно. Это возможно потому, что все квалифицированные технические работы уже предварительно осуществлены соответствующими специалистами... Пассажиры, однако, все же несомненно, нуждаются в определенных знаниях и информации, например, куда они собираются лететь, какие авиалинии и рейсы могут быть использованы, как добраться до аэропорта, и когда они могут вылететь; они, конечно, нуждаются также в средствах на билет; но они не обязаны знать, как построен самолет и даже как он летает». Исследователь, по мнению сторонников такой позиции, находится в аналогичном положении, и в его задачу не входит вдаваться в детали особенностей статистического аппарата, так же как и в устройство компьютера. Формулы, которые лежат в основе анализа, в свое время были разработаны и использованы в компьютерных программах опытными специалистами в области математики и статистики. Поэтому любой, кто испытывает потребность и интерес к изучению социальных фактов, закономерностей, принципов, может использовать компьютерные программы в готовом виде, не особенно вдаваясь в суть математико-статистического аппарата, заложенного в основу социологического анализа.

Первая точка зрения, по нашему глубокому убеждению, является более правильной. Тем более, что аргументы сторонников второй позиции представляются сомнительными хотя бы потому, что аналогия с пассажиром неправомерна: пассажир, пользуясь результатами деятельности других специалистов, платит за билет; социолог же, принимаясь за анализ первичного материала, выполняет заказ и обычно сам взимает плату за выполняемую работу. Тогда как пассажир выступает потребителем результатов чужого квалифицированного труда, социолог выступает в роли квалифицированного исполнителя. Именно на его знания и квалификацию полагается заказчик или потребитель конечного продукта социологической информации, который, по нашему мнению, скорее выступает в роли пассажира (если уж воспользоваться приведенной выше аналогией).

Следующий раздел «Анализ первичной социологической информации» направлен на раскрытие основных правил и требований при проведении анализа первичной социологической информации. Данное учебное пособие призвано, на наш взгляд, компенсировать пробелы в подготовке среднего звена в социологической профессии. Поэтому здесь не будут раскрываться основы теории вероятности и математической статистики, позволяющие высококвалифицированному специалисту с полным пониманием всего материала в совокупности и осознанием всех возможностей математического аппарата вскрыть тенденции и закономерности, которые могут быть получены из сырого первичного материала. В этом разделе будут изложены только основные способы и приемы анализа, наиболее часто используемые при освещении данных социологических опросов; подчеркнуты основные правила и требования, нарушение которых приводит к искажению фактов; особое внимание будет сосредоточено на тех ошибках, которые чаще всего допускаются при анализе первичного материала специалистами, не имеющими достаточной базы знаний. Другими словами, основное внимание будет сосредоточено на освоении технологического процесса. Данное пособие, на наш взгляд, позволит в определенной степени ликвидировать пробел в подготовке специалистов-технологов в области анализа первичного социологического материала, основные функции которых — компетентное выполнение стандартного анализа результатов массовых опросов.

Лекция 7.

Анализ и проблема измерения

1. Уровни социологического анализа эмпирических данных

После того, как результаты опроса введены в компьютер и сформированы в соответствующий файл, называемый массивом данных, дальнейшая обработка осуществляется в зависимости от выбранного уровня, вида и конкретных процедур анализа. Можно выделить два основных уровня анализа: описание результатов исследования и объяснение полученных фактов.

Описание результатов

Первый уровень — *описание* полученных фактов в основном сводится к оформлению и интерпретации одномерных и двумерных распределений респондентов по их ответам на поставленные в анкете вопросы. Задачи исследователя в этом случае сводятся к последовательному осуществлению следующих процедур:

- 1) сформировать ряды распределений;
- 2) произвести статистическую экспертизу полученных рядов распределений;
- 3) оформить полученные и статистически оцененные ряды распределений в таблицы или графики таким образом, чтобы они при максимальной информационной нагрузке были достаточно доступны для восприятия и понимания рядовым потребителем.

Объяснение полученных фактов

Второй уровень анализа — *объяснение* полученных фактов. В контексте анализа первичных социологических данных под объяснением понимается: установление связей между типами ответов на различные вопросы анкеты, оценка характера и направленности этих связей (попытка поиска причинно-следственных отношений); оценка и описание вероятности проявления этих связей на основании собранного материала. В данном случае объяснение носит не столько характер логического обоснования выдвигаемого автором объяснения, сколько статистической обоснованности результатов проверки гипотез, выдвинутых в программе исследования. Попытки объяснения полученных фактов на этом уровне анализа основываются на различных статистических процедурах множественного анализа, среди которых простейшим (но очень эффективным) является сравнительный анализ трехмерных распределений, направленный на поиск причинного фактора.

Прежде чем перейти к рассмотрению технологии конкретных процедур анализа — описания и объяснения, следует внимательнее присмотреться к одной из самых распространенных ошибок, которую допускают неискушенные социологи на этом этапе работы. Их стратегическая ошибка, нередко порождающая искажение фактов, заключается в том, что, применяя процедуры описания данных, исследователь стремится объяснить их, интуитивно чувствуя, что более высокий уровень анализа требует объяснения. Но его объяснения основаны не на результатах статистической проверки выдвинутых гипотез, как того требует анализ эмпирического материала, а совсем на другом базисе. Когда исследователь является специалистом в области изучаемой проблемы, но плохо знаком с азами статистического анализа, он анализирует проблему в выдвинутой им парадигме, привлекая результаты в качестве иллюстративного материала; это не самый худший случай, но здесь надо отдавать себе отчет, что автор такого анализа в море полученных фактов будет тенденциозно подбирать иллюстрации своих положений, вытекающих скорее из его логики и позиции, нежели из результатов исследования. Но чаще, социолог, выполняя заказное эмпирическое исследование, не является профессиональным специалистом в области исследуемой проблемы; его основная задача — профессионально провести опрос и компетентно проанализировать собранный материал. В таком случае, если он начинает «объяснять» полученные факты без осуществления соответствующего статистического анализа, то его объяснения, в лучшем случае, основываются на общей осведомленности, логике и здравом смысле; но, обычно, наряду с этим, основой объяснения полученных фактов становятся также житейские предрассудки и стереотипы, а иной раз и просто вольная фантазия. Такой «анализ» представляет собой личную точку зрения исследователя как человека и гражданина, с привлечением некоторых данных, полученных в результате опроса. При этом потребитель информации вводится в заблуждение, поскольку считает, что все приводимые объяснения действительно вытекают из результатов исследования.

Первым условием корректности анализа результатов массового опроса является даже не столько квалификация исследователя, сколько адекватное его представление об уровне своей квалификации и применении методов анализа и интерпретации, адекватных уровню своей подготовленности. Если социолог освоит все необходимые правила описания и представления полученных фактов, то грамотно оформленный материал практически всегда несет информацию, находящую своего заинтересованного потребителя. Сказанное вовсе не означает, что автор не может давать объяснение описанным фактам. Но он должен прежде всего осознавать сам и обязательно довести до сведения читателя, что это объяснение проистекает не из полученного материала, а из общей позиции автора.

В социологии следует различать первичный и вторичный анализ результатов массовых опросов. Первичный анализ предполагает статистическую обработку данных опроса и требует определенной подготовленности в области математической статистики. Под вторичным анализом подразумевается общий логический анализ проблемы; и в данном случае результаты опросов привлекаются лишь как установленные факты для иллюстрации (подтверждения или опровержения) определенных высказываний автора вторичного анализа. При этом практически не имеет значения, привлекает ли автор данные собственного исследования или данные других источников (публикаций, документов и т.п.), поскольку свои выводы он обосновывает логически.

Чем должен определяться выбор уровня анализа? Во-первых, как отмечалось выше, квалификацией исполнителя, и прежде всего его подготовленностью в области математической статистики. Во-вторых, задачами исследования; так, например, для опросов общественного мнения в подавляющем большинстве случаев вполне достаточно описательного уровня.

Если выбор уровня анализа определяется квалификацией исполнителя и задачами исследования, то выбор конкретных процедур и статистических методов зависит от типа шкал.

Какие статистические процедуры могут быть использованы при анализе характера ответов на данный вопрос анкеты, а какие нет, в первую очередь определяется видом и типом построенной шкалы.

Все шкалы в практике социологического анализа можно свести к трем основным видам: номинальная, порядковая и метрическая.

Тип шкалы определяется особенностями такой ее характеристики, как «дискретность-непрерывность».

Шкала характеризуется как «дискретная» в том случае, когда варианты ответов в общем виде не связаны друг с другом (между двумя «соседними» вариантами не предполагается возможность промежуточно-переходного варианта). Например, ответы на вопрос о национальности (1 — украинец, 2 — русский, 3 — другая национальность) представляют собой типичный образец дискретной шкалы. Даже если респондент имеет родителей разной национальности (например, мать — украинка, отец — русский), все равно данная шкала не позволяет этого установить (при необходимости измерений этих различий следует применять другой инструментарий). Сама шкала не предполагает «промежуточных» ответов.

Непрерывные шкалы предполагают, что между двумя «соседними» ответами могут быть расположены промежуточные ответы. Например, ответы на вопросы о возрасте

предполагают, что между возрастом 20 лет и возрастом 21 год могут быть люди с возрастом 20,1 года, 20,2 года и т.д. Даже если шкала не предусматривает такие градации, потенциально она их включает.

Именно степень «непрерывности» шкалы определяет возможности применения различных статистических процедур.

Именно под этим углом зрения шкалы разделяют на номинальные, порядковые и метрические. Все номинальные шкалы являются по своему типу дискретными; все метрические — непрерывными; порядковые же шкалы занимают промежуточное положение. Одной из аналитических задач исследователя является определение степени дискретности-непрерывности порядковой шкалы. Ниже мы уточним особенности этой процедуры.

Напомним, что обозначение шкал проводится до начала обработки данных, и результаты заносятся в «ключевую» анкету (напротив каждого пункта-вопроса анкеты отмечен вид шкалы: номинальная, порядковая, метрическая).

Таким образом, прежде, чем приступать к статистическому анализу, необходимо научиться достаточно свободно определять вид и тип шкалы.

2. Обозначение шкал

Номинальная шкала

Номинальными называются шкалы, в которых код ответа является не более чем наименованием категории. Что это означает? Большинство компьютерных программ по обработке данных оперируют не словами и буквами, а цифрами; поэтому информация для обработки в большинстве случаев не словами, а числами (соответствующим кодом). Например, вопрос анкеты «Ваш пол?» сопровождается двумя вариантами ответов:

1 — мужчины

2 — женщины

Респондент, ориентируясь на содержание слов («мужчины», «женщины»), обводит соответствующий код («1» или «2»). Во всех дальнейших операциях в файле данных под кодом «1» подразумеваются мужчины, под кодом «2» — женщины. В тех случаях, когда шкала номинальная, при всех проводимых операциях эти цифры будут означать лишь *наименование* соответствующей категории и не будут означать число, которым обозначен код: введенная цифра не имеет своего числового значения. Мы вполне могли бы использовать любые другие цифры или даже числа (например, «7» и «2» или «10» и «20»). Это не имеет никакого значения для дальнейших операций, если шкала обозначена номинальной. Но для удобства обработки используют обычный порядок чисел. При этом не предполагается, что мужчины в каком-то отношении первичны, а женщины вторичны. Автор без всякого ущерба для анализа мог бы поменять их местами.

Почему это важно понимать в процессе анализа данных? Обозначая эту шкалу как «номинальную», автор отказывается от числовых манипуляций с ней. Обычно при числовых манипуляциях, когда цифры обозначают числа, подразумевается, что «1»

меньше «2» в два раза; или, если от «2» отнять «1», то получится «1»; если к «1» прибавить «2», то получится «3» и т.п. В случае же обозначения шкалы как номинальной, цифры «1» и «2» обозначают только наименование соответствующих категорий.

При работе с номинальными шкалами следует учитывать два правила. Во-первых, надо точно знать, какая категория каким числовым кодом обозначена. Во-вторых, при анализе данных с числовыми знаками кодов нельзя осуществлять никаких арифметических действий — сложения, вычитания, умножения, деления; в частности, нельзя подсчитывать среднее арифметическое. Это вполне понятно исследователю, когда он анализирует демографические данные; но достаточно часто допускаются ошибки, когда данные касаются отношений. Примером такой ошибки часто является анализ ответов на вопрос: «В какой степени Вы доверяете N?». Допустим, что веер ответов сформирован следующим образом:

- 1 — совершенно не доверяю
- 2 — в определенной степени доверяю
- 3 — полностью доверяю
- 4 — трудно сказать, доверяю или нет
- 5 — я его не знаю

Эта шкала может быть обозначена только как номинальная, так как код «4» по своему значению не соответствует 4 баллам доверия, а код «5» не означает высшего балла доверия. Неопытные авторы, анализируя данные, заказывают подсчет индекса (рейтинга доверия), и тогда соответствующая кодовая позиция в последующих математических процедурах используется как число с соответствующим значением. В этом случае наибольшие баллы получают те лица, которых респонденты не знают, и те, к которым у них не сформировалось определенное отношение. Чтобы избежать подобных казусов, исследователь должен хорошо разобраться в характере шкал и точно их определить до начала анализа. Лучше, конечно, когда это делается при подготовке анкеты. Но если на этом этапе исследователь упустил возможность сознательно сформировать шкалы, или после начала анализа ему захотелось применить не предусмотренные исследовательской программой статистические приемы анализа, он должен определить тип шкалы и, при необходимости, перекодировать информацию.

Примерами номинальных шкал в анкетах являются ответы на вопрос о национальности, профессии, отрасли, в которой работает респондент, населенном пункте или регионе, в котором он проживает, мотивах того или иного поступка и т.д.

Номинальная шкала может быть подготовлена исследователем как при составлении анкеты, так и в процессе ручной обработки анкет (до в ввода информации в компьютер). Например, на вопрос «Ваша национальность?» автор закладывает веер ответов: 1 — украинец, 2 — русский, 3 — еврей, 4 — белорус, 5 — поляк, 6 — другая; или 1 — украинец, 2 — русский, 3 — другая (конкретное содержание веера ответов определяется задачами исследователя, особенностями выборки или представлением автора инструментария о том, какие национальности могут попасть в выборку). В ситуации, когда исследователь точно не знает, представители каких национальностей могут попасть в число респондентов, и не хочет терять эту информацию, он может поставить вопрос о национальности в открытой форме. Так часто поступают при определении профессии

респондентов, мотивов поведения и т.п. информации, когда исследователь не уверен, что он осведомлен о всех возможных вариантах ответа.

Но в любом случае, закрытый «веер» ответов необходимо подготовить, по крайней мере, в виде кодировочного листа до перекодировки и редактирования массива анкет. Если исследователь не был уверен в полноте своей осведомленности в период конструирования анкеты и данный вопрос поставил в открытой форме, он может после получения заполненных анкет поручить одному из помощников выписать ответы на этот вопрос (это может быть список по всем анкетам или по части анкет) и на основании выписанных данных подготовить кодировочный лист — список национальностей, в котором напротив каждой национальности будет указан ее кодировочный номер. Как видим, для ввода в компьютер в любом случае каждый вариант ответа должен иметь соответствующий числовой код, который, напомним, в случае номинальной шкалы является лишь аналогом соответствующего наименования.

Номинальные шкалы бывают *альтернативные* и *вариантные* (многовариантные). Альтернативные шкалы допускают возможность лишь одного ответа на поставленный вопрос (пол, национальность и т.п.). Вариантные шкалы допускают возможность нескольких ответов на поставленный вопрос. Например, в вопросе о причинах, по которым респондент хочет поменять место работы или место жительства, могут допускаться несколько вариантов ответов на поставленный вопрос.

Определение разновидности номинальной шкалы имеет принципиальное значение для дальнейшей обработки и анализа данных. Поэтому до начала обработки исследователь должен определить не только вид, но и разновидность шкалы.

Обычно со шкалами, касающимися демографических характеристик, исследователь не испытывает затруднений — казалось бы, в этом случае все шкалы должны быть альтернативными. Но и здесь есть определенные «подводные камни». Например, формулируя ответы на вопрос о том, какое учебное заведение окончил респондент, исследователь подразумевает альтернативную шкалу. Но он часто не учитывает, что респондент мог окончить не одно учебное заведение. Или, например, вопрос о том, какая профессия у респондента, также может упускать из вида, что респондент не знает, идет ли речь о профессии по базовому образованию, или о той, по которой он проработал большую часть жизни, или о той, по которой он работает в настоящее время. Поэтому в формулировку вопроса автор анкеты должен либо внести соответствующие (уточняющие) характеристики (если он хочет добиться альтернативной номинальной шкалы), либо обозначить шкалу как вариантную.

Несколько сложнее обстоят дела с вопросами, касающимися мнений и, особенно, отношений людей к различным социальным процессам и явлениям. Чаще всего определение разновидности номинальной шкалы осуществляется исследователем волонтарно (в форме дополнительной инструкции к вопросу). Например, к вопросу (с соответствующим перечнем ответов) «Что, по Вашему мнению, препятствует эффективному проведению экономических реформ?» можно дать разные инструкции: 1) «Отметьте только один, самый важный, вариант ответа»; 2) «Отметьте не больше 2-х (3-х, 5-ти и т.д.) вариантов ответа»; 3) «Отметьте все, подходящие, с Вашей точки зрения, варианты ответа».

Конкретный вариант инструкции определяет исследователь, в зависимости от той информации, которую он намерен получить, интерпретируя результаты. Важно только

помнить, чтобы такая инструкция обязательно присутствовала в формулировке вопроса, если сама постановка вопроса подразумевает возможность многовариантного ответа.

При определении вида шкалы необходимо учитывать, что если такой вариант ответа, как «затрудняюсь ответить», стоит в конце перечня вариантов ответов (и соответственно имеет самый большой, по числовому значению, код), то такую шкалу всегда следует принимать за номинальную.

Порядковая шкала

Порядковой шкалой является такая последовательность ответов на вопрос, при которой каждый очередной вариант ответа характеризует последующее ранговое место в континууме значений. При использовании порядковой шкалы всегда можно сказать, что каждый последующий вариант ответа наделяется значением «выше» или «ниже», «больше» или «меньше» и т.п. Примером порядковой шкалы могут быть ответы на вопрос об уровне образования:

1 — начальное

2 — среднее

3 — высшее

Другим примером порядковой шкалы является определенный веер ответов на вопрос: «В какой степени Вы удовлетворены ...?»:

1 — совершенно не удовлетворен

2 — скорее не удовлетворен

3 — трудно сказать, удовлетворен или нет

4 — скорее удовлетворен

5 — полностью удовлетворен

В этих случаях кодовый номер варианта ответа указывает на его место в определенно заданном порядке. В таких случаях обязательно есть критерий континуума (своеобразная линейка), вдоль которой расположены варианты ответов: в первом примере — это уровень образования; во втором — степень удовлетворенности. Мы всегда можем сказать, что высшее образование по своему уровню выше среднего, а среднее выше начального.

Когда используются порядковые шкалы, кодовое значение варианта ответа может более продуктивно (по сравнению с номинальными) использоваться при статистическом анализе данных, так как в этом случае цифровое значение кода наполняется определенным числовым смыслом.

В зависимости от степени непрерывности одни порядковые шкалы могут рассматриваться как *ранговые*, другие — как *интервальные*. Такое разделение осуществляется на основе определения «дистанции» между вариантами ответов.

Если исследователь считает, что «дистанцию» между вариантами ответов определить невозможно, или — между разными вариантами она различна, то порядковую шкалу, следует оценивать как ранговую. Например, в приведенном выше измерении уровня образования нельзя сказать, что среднее образование настолько же «больше» начального, насколько высшее — «больше» среднего. В качестве другого примера может быть приведена шкала ответов на вопрос о том, каким по счету в семье ребенком является респондент: два человека могут указать вторую позицию, но при этом у одного респондента между ним и первым ребенком в семье может быть разница в один год, а у другого — 10 лет. Можно видеть, что при использовании ранговой шкалы теряется информация о количественных различиях в рамках отдельных градаций.

При использовании такого типа шкал кодовое значение ответа принимается за определенный балл, характеризующий порядковый ранг исследуемого качества. В силу того, что ранговые шкалы, определяя порядок (последовательность) ответов, не учитывают дистанцию (интервал) между обозначенными категориями, они имеют существенную ограниченность, с точки зрения применения к ним различных математических процедур. В частности, по отношению к ранговым шкалам неприменим расчет среднего арифметического, а также все коэффициенты, включающие в свои расчетные формулы значение среднего арифметического.

Если исследователь предполагает и обосновывает, что интервал между вариантами ответов приближается к равному значению, он принимает порядковую шкалу за условно *интервальную**.

Истинно интервальную шкалу исследователь может получить в результате кропотливой предварительной работы (так строятся шкалы Терстоуна, многие тестовые процедуры и т.д.). При подготовке оперативного эмпирического исследования отработка интервальности не всегда целесообразна, так как это достаточно долговременная процедура (обычно этот процесс длится 1-2 года). Но, работая над формулировками вариантов ответов, даже в оперативном исследовании автор старается так сформулировать ответы, чтобы «дистанции» между вариантами воспринимались респондентом как равные интервалы. К такому типу шкал обычно относятся варианты ответов на вопросы, касающиеся удовлетворенности, уровня доверия и т.п. При изучении подобных признаков исследователь нередко прибегает к тому, чтобы уйти от вербальных формулировок, и сразу предлагает респонденту оценить свое отношение в баллах. Например, после вопроса «В какой степени вы доверяете ...?» предлагается инструкция: «Отметьте свой вариант ответа на шкале, где 1 балл означает полное недоверие, а 5 баллов — полное доверие». Такая форма предъявления вера ответов подразумевает условное обозначение интервальности — расстояния между интервалами ответов определяются в один балл.

Конечно, следует отдавать отчетность в условности таких «интервалов», их значительной субъективности. Но в той мере, в какой такое шкалирование имплицитно обосновано, шкала может быть обозначена как интервальная. Это позволяет расширить арсенал статистических методов анализа данного признака («уровня доверия», «степени удовлетворенности» и т.п.). Это относится, в первую очередь, к тем случаям, когда исследователь планирует при анализе использовать индексы (рейтинги доверия, индексы удовлетворенности и т.п.).

Таким образом, определив первоначально шкалу как порядковую, исследователь, пользуясь критерием «дискретности-непрерывности», устанавливает, будет он ее интерпретировать как ранговую, или как интервальную с той или иной мерой обоснованности. В зависимости от этого решения строится план анализа — применение

тех или иных статистических процедур. Правомерность их применения полностью обусловлена степенью обоснованности определения типа шкалы.

Метрическая (интервальная) шкала

Шкала с безусловно равными значениями интервалов между позициями ответов называется интервальной, или метрической. Метрическая шкала всегда имеет единицу измерения, которой меряется расстояние между позициями. Примерами метрической шкалы являются данные по возрасту (единица измерения — год); по доходу (единица измерения — денежная единица) и т.д. Метрическая шкала практически не имеет ограничений в применении статистических процедур; в этом отношении метрические шкалы являются наиболее эффективными для аналитика. Стремление добиваться интервальности порядковых шкал определяется потребностью исследователя приблизить порядковую шкалу к аналитическим возможностям метрической.

Но и метрические шкалы, с точки зрения определенных потребностей анализа и, особенно, представления данных, имеют свои недостатки. Например, при описании и представлении результатов, данные, полученные с помощью метрической шкалы, выглядят довольно громоздко: например, шкала возраста обычно варьирует от 18 до 75 лет; еще большую протяженность имеет шкала дохода. Если автор хочет представить данные, касающиеся того, как люди разного возраста относятся, например, к экономическим реформам, то таблица, в которой первую группу составят лица 18 лет, вторую — 19, третью — 20 и т.д., будет иметь очень много строк, и пользователю трудно будет проследить тенденцию, фиксируемую в различиях ответов лиц разных возрастных категорий. Поэтому в процессе анализа, переходя от одних статистических процедур к другим, исследователь может преобразовывать тип шкалы; так например, группируя данные по возрасту, социолог фактически преобразовывает метрическую шкалу в порядковую. В ряде случаев это не только возможно, но и необходимо.

Но прежде чем перейти к проблеме преобразования шкал, следует лишний раз напомнить, что самая главная проблема, о которой не должен забывать социолог — это своевременное и правильное определение типа шкалы. И прежде чем применять ту или иную статистическую процедуру анализа данных, полученных по конкретному вопросу анкеты, исследователь должен точно определить тип шкалы, которую образует совокупность ответов на данный вопрос. При определении типа шкалы, неопытного исследователя иногда может ввести в заблуждение форма ответов, особенно, если ответы представлены в цифровом оформлении.

Например, на одном из методических семинаров мы просили студентов-социологов определить тип шкалы, полученной из ответов на вопрос: «Отметьте, пожалуйста, время, в которое Вы обычно смотрите телепередачи». Веер ответов был представлен линейкой, размеченной на часы, составляющие в совокупности определенное время суток:

6_7_8_9_10_11_12_13_14_15_16_17_18_19_20_21_22_23_24

УТРО

ДЕНЬ

ВЕЧЕР

Только одна студентка в группе смогла правильно определить тип данной шкалы — номинальная. Подавляющее большинство определили ее как метрическую. Это произошло потому, что варианты ответов были представлены, во-первых, числами, а во-вторых, имели равный интервал. В данном случае студенты (потенциальные исследователи) ориентировались на форму, а не на содержание шкалы. Метрические шкалы используются тогда, когда в центр анализа ставятся средние значения анализируемых величин (средний возраст, средний душевой доход и т.п.), и эти средние величины несут определенную содержательную и смысловую нагрузку. К примеру, если бы вопрос стоял о том, сколько часов в день обычно респондент смотрит телевизор, то из полученных данных о том, что половина респондентов указала 2 часа, а другая половина — 6 часов, можно было бы сделать вывод: в среднем, один человек смотрит телевизор 4 часа (разумеется, данные взяты условно). Но если шкала используется в таком вопросе, который был приведен выше, то при вычислении среднего арифметического можно получить даже не искаженную, а абсурдную информацию: например, если половина респондентов смотрит телевизор в час дня, а другая половина в шесть часов вечера, вовсе не следует, что в среднем они смотрят телевизор в четыре часа дня.

Оценивая тип шкалы, исследователь должен обращать внимание не столько на форму, сколько на содержание ответов.

Преобразование шкал

Поскольку каждая из шкал имеет свои преимущества и свои недостатки для использования различных процедур статистического анализа данных, то исследователь при подготовке программы обычно старается сконструировать тот тип шкалы, который в наибольшей степени соответствует целям анализа. Но возможны ситуации, когда автор хочет применить методы статистического анализа, не соответствующие данной шкале. Такая ситуация может возникнуть вследствие разных причин, первое место среди которых обычно занимает отсутствие исследовательской программы и плана анализа. Другой причиной необходимости преобразования шкалы является новая гипотеза, для проверки которой необходимы статистические процедуры, не предусмотренные первоначальным планом. В третьих, нередко исследователь попадает в ситуацию, когда ему приходится приступать к анализу данных, собранных по инструментарию, который разрабатывали другие авторы. В подобных случаях исследователь оказывается как бы в вынужденной ситуации необходимости преобразования шкал.

Есть и объективные причины, по которым социолог уже на этапе планирования анализа осознает, что в процессе работы он должен будет преобразовывать определенные шкалы. Планирование преобразования шкал необходимо в ряде случаев.

Один из таких случаев — это рассогласование между возможностями сбора информации и потребностями анализа. Исследователь, составляя анкету, отдает себе отчет в том, что тип шкалы не соответствует потребностям анализа, но сконструировать шкалу по-другому принципиально невозможно, так как он понимает, что есть значительное число респондентов, которые не смогут найти «свой» ответ на шкале, необходимой для анализа. Поэтому на этапе сбора данных будет утрачена принципиальная информация. Например, при измерении степени доверия к различным политическим фигурам оптимальной является интервальная шкала (допустим, респондента просят оценить степень своего доверия по пятибалльной шкале, где 1 балл означает крайнюю степень недоверия, а 5 баллов — высшую степень доверия). Такая шкала позволяет подсчитывать

количественный индекс доверия и сравнивать рейтинги доверия по отношению к различным политическим деятелям. Но исследователь понимает, что часть респондентов может просто не знать того или иного политического деятеля. Если такой вариант исключить из веера ответов на вопрос, касающийся доверия, то полученная информация будет искажена. Включая же этот вариант ответа, исследователь получает номинальную шкалу. При этом он может заранее планировать ее преобразование в процессе анализа для получения расширенных возможностей применения статистического аппарата.

Возможна и другая ситуация, порождающая необходимость планирования преобразования шкал. Исследователь на этапе планирования анализа предполагает использовать различные статистические процедуры, одни из которых адекватны номинальным шкалам, другие порядковым или метрическим. Например, задавая вопрос об отношении населения к частной собственности на землю, социолог в одних случаях хочет оперировать индексом отношения к приватизации земли (допустим, сравнения между собой отношение к приватизации различных социальных групп); а в других случаях анализа он хочет сформировать две основные группы населения, различающиеся между собой по данному показателю («за приватизацию земли» и «против приватизации земли») и анализировать различие между этими двумя категориями по их ценностным ориентациям, социальному статусу, психологическому состоянию и т.д. В первом типе анализа исследователю необходимо пользоваться статистическими процедурами, применимыми к интервальным шкалам, а во втором — к номинальным.

Во многих случаях существует принципиальная возможность преобразования шкал.

Любая метрическая шкала может быть преобразована как в номинальную, так и в порядковую. Простейший способ преобразования метрической шкалы в номинальную — дихотомия (группировка на две группы по определенному критерию). Например, по возрастной характеристике всех респондентов можно разделить на две группы: до 30 лет («молодежь») и старше 30 лет. Здесь возможна группировка с любой степенью дробности. Какая группировка является оптимальной?

Во-первых, это зависит от гипотез. Если автор имеет достаточно обоснованную гипотезу о границах «социального возраста» (например, возраст окончания вуза или возрастной критерий выхода на пенсию и т.п.), которые могут влиять на изучаемые им установки, он формирует группы с определенными «социально-возрастными» границами. Если нет достаточно обоснованных «возрастных» гипотез, можно использовать возрастную стратификацию, принятую в государственной статистике. Но здесь неопытного социолога подстерегают определенные «подводные камни»: чем более дробная группировка, тем меньшей численности будут полученные им группы; а чем меньше численность анализируемых групп, тем труднее установить, являются ли *значимыми* зафиксированные при анализе данных различия в отношении этих групп к тем или иным социальным явлениям. Исследователь не должен забывать, что если в анализируемой группе меньше 100 человек, то когда он говорит, сколько процентов из них относятся к чему-либо положительно или отрицательно, существует определенного рода натяжка*.

В тех случаях, когда исследователь не имеет собственной гипотезы в отношении влияния возраста на тот или иной изучаемый показатель, а хочет лишь сориентироваться сам, влияет ли возрастной фактор на изучаемые явления, ему можно порекомендовать условную трехуровневую возрастную стратификацию типичной региональной репрезентативной выборки взрослого населения (18-29, 30-55, 56-75 лет). Такой подход при рабочем анализе позволит в первом приближении увидеть тенденции влияния возрастного фактора на изучаемые процессы и явления.

Что касается номинальных и порядковых шкал, то довольно часто их преобразование связано с потребностью определить место такого варианта ответа, как «трудно сказать» («затрудняюсь ответить» и т.п.). Как уже отмечалось, если этот вариант ответа стоит в конце общего веера ответов, то такая шкала всегда должна рассматриваться как номинальная (сужающая возможности анализа: исследователь не должен считать индексы и ряд коэффициентов корреляций с другими переменными). В каких случаях она может быть преобразована в порядковую? В тех случаях, когда измеряется *эмоциональное* отношение людей, вариант ответа «трудно сказать» может рассматриваться как «нулевая точка между положительной и отрицательной осями шкалы; поэтому при преобразовании шкалы его можно ставить посередине симметричной шкалы и присваивать среднее (медианное) значение. Варианты ответов типа «Не знаю» в этом случае интерпретируются как «неответ»; и анкеты с этим вариантом исключаются из анализа по данному признаку. Примерами шкал, измеряющих эмоциональное отношение, могут служить как прямые ответы на вопрос: «Как Вы относитесь (положительно или отрицательно) к ...?», так и вопросы, касающиеся удовлетворенности, доверия и т.п.

Как преобразовать номинальную шкалу, допустим, удовлетворенности, в порядковую? Например, в анкете был поставлен вопрос с предложенным веером ответов: «В какой степени Вы удовлетворены...?»

- 1 — полностью удовлетворен
- 2 — в определенной степени удовлетворен
- 3 — в определенной степени не удовлетворен
- 4 — совершенно не удовлетворен
- 5 — трудно сказать

В таком виде веер ответов представляет собой номинальную шкалу, позволяющую строить распределение по отдельным градациям процентов (или частот) респондентов, которые выражают ту или иную степень удовлетворенности или неудовлетворенности. Сравнительный анализ таких данных проводить достаточно неудобно (например, при ежемесячном мониторинге, если происходит изменение в показателе процентов каждой из пяти групп, трудно сказать, насколько же изменился индекс удовлетворенности в целом).

Но в данном случае вариант ответа «трудно сказать» можно рассматривать как нейтральную позицию на шкале удовлетворенности, так как не имеет значения, по каким именно причинам респондент затрудняется ответить; важно что он не высказывает ни удовлетворенности, ни неудовлетворенности (по этому критерию он занимает промежуточную позицию). Поэтому при преобразовании шкалы он может быть переставлен в данном случае на третье место с присвоением соответствующего рангового значения. Преобразованная шкала выглядит следующим образом:

- 1 — полностью удовлетворен
- 2 — в определенной степени удовлетворен
- 3 — трудно сказать (трудно сказать, удовлетворен или нет)
- 4 — в определенной степени не удовлетворен

5 — совершенно не удовлетворен

В таком виде шкалу можно считать порядковой, в которой кодовый номер отражает, по крайней мере, ранговое место в континууме степени удовлетворенности. Условно считая интервалы равными (а эту условность социологи допускают настолько часто, что, по негласному конвенциональному согласию, в настоящее время ее можно признать легитимной), можно, определив шкалу интервальной, считать средний индекс*.

В данном случае самое высокое ранговое место — у ответа «совершенно не удовлетворен». Следовательно, полученный в баллах индекс должен быть назван «индексом неудовлетворенности». Если же исследователь хочет полученный количественный индикатор назвать «индексом удовлетворенности», он должен поменять порядок следования вариантов ответов с соответствующими кодами:

- 1 — совершенно не удовлетворен;
- 2 — в определенной степени не удовлетворен;
- 3 — трудно сказать, удовлетворен или нет;
- 4 — в определенной степени удовлетворен;
- 5 — полностью удовлетворен.

Если при планировании анализа исследователь предполагал использовать такую переменную, как удовлетворенность (например, удовлетворенность жизнью, работой и т.п.), в качестве критерия группировки населения на тех, кто удовлетворен, и на тех, кто не удовлетворен, он может преобразовать имеющуюся в анкете номинальную шкалу в номинальную же, но с меньшим перечнем ответов, укрупняя группы:

- 1 — удовлетворены;
- 2 — не удовлетворены;
- 3 — затрудняются определить свое отношение.

Данные, представленные в таком виде, более удобны для сравнительного анализа. И если автор не планирует расчет индексов и коэффициентов корреляций, то именно в таком виде (с укрупненной группировкой) лучше конструировать вопрос в анкете. В таком случае и респондентам легче отвечать, и при обработке не надо будет проводить дополнительного преобразования. Но если исследователь планирует применять различные виды анализа, требующие преобразования шкал, то в анкете в таком случае лучше предусмотреть более дробную оценку измерения респондентов. Степень дробности измерения называется *точностью шкалы*. Чем выше точность шкалы, тем тоньше она реагирует на изменение в динамике отношения и на различия в отношении к разным объектам. Следовательно, чем более ясное представление на этапе планирования у исследователя — каким именно образом он будет анализировать материал, тем более адекватный инструментарий он подготовит. Другими словами, если исследователь не планирует сложный анализ, нет смысла повышать точность шкал, перегружая анкету лишними вариантами ответов на предлагаемые вопросы. При желании сохранить возможность углубленного анализа, он должен очень внимательно отнестись к процедуре

преобразования шкалы, прежде чем использовать разнообразный арсенал статистических методов анализа.

3. Статистический анализ

Выбор методов статистического анализа

Эмпирическая социология располагает достаточно широким набором статистических процедур, позволяющих сырые данные, содержащиеся в отдельных анкетах, преобразовать таким образом, чтобы можно было установить определенные факты массового сознания и проанализировать связи между установленными фактами.

Выбор конкретных методов статистического анализа обычно определяется двумя основными предпосылками: 1) характером данных (форма материала, обусловленная сформированными при подготовке анкеты шкалами); 2) предполагаемой формой конечного «продукта» работы (отчета, научной статьи, публикации в средствах массовой информации и т.д.). Другими словами, выбирая конкретный способ статистического анализа, исследователь должен быть уверен, что данный его вид является наиболее подходящим не только для этого типа данных, но и для конечного продукта исследования. Иногда исследователь использует неадекватные средства анализа не столько потому, что он недостаточно хорошо знаком с тонкостями статистических процедур, а потому, что не учитывает конечный вид оформления полученной информации — продукта исследования. В таких случаях, применяя достаточно сложные и трудоемкие способы исчисления, автор получает довольно частные результаты, хотя и корректные с математической точки зрения (может быть, даже «красивые», с точки зрения специалиста в области математической статистики), но малоинтересные для конкретного потребителя (заказчика) социальной информации, испытывающего потребность в практическом использовании полученных результатов.

Как отмечалось выше, существуют два основных типа методов статистического анализа, предназначенных для решения двух основных типов задач: 1) представление (описание) данных; 2) измерение связей и зависимостей между данными, полученными в результате ответов на различные пункты (вопросы) анкеты.

Приступая к выбору статистических методов, социолог должен:

- 1) определить уровень проводимого анализа (описательный, объяснительный), руководствуясь формой конечного продукта (итогового документа по результатам исследования);
- 2) выбрать конкретные статистические процедуры по отношению к каждому из пунктов анкеты, руководствуясь типом шкалы, которую представляет структура веера ответов на поставленный в анкете вопрос.

Обычно не используют одни и те же статистические средства по отношению ко всем вопросам анкеты, поскольку чаще всего различные вопросы описываются различными шкалами.

В данном учебном пособии не будет приводиться подробное описание математических основ статистических процедур, а также математических выкладок получения основных

формул расчетов соответствующих коэффициентов. Задача нашего учебника ограничивается введением читателя в круг основных понятий и категорий, без знания которых манипуляции с результатами социологических опросов значительно чаще приводят к ошибкам и искажению информации, нежели к описанию, анализу и интерпретации полученных данных.

Ряд распределения

Практически любой вид анализа (независимо от типа шкалы) начинается с формирования ряда распределения, что является первым шагом в обобщении данных полученных от каждого конкретного респондента. Ряд распределения — это оформленная соответствующим образом запись результатов подсчета количества респондентов, отметивших каждый из вариантов ответов на поставленный в анкете вопрос.

Обобщенные данные по каждому пункту анкеты называются *переменными* (поскольку ответы одних респондентов могут отличаться от ответов других) в контексте статистического анализа и *признаками* в контексте социологического анализа.

Исследователь может проводить анализ по каждой отдельно взятой переменной; может проводить анализ по двум переменным и показывать связь между ними; может проводить анализ по нескольким переменным и показывать различные типы связей между ними. Когда исследователь анализирует результаты ответов на один, отдельно взятый вопрос анкеты, такой вид анализа называется *одномерным*, когда исследователь включает в анализ две переменные и устанавливает характер связи между ними, такой вид анализа называется *двумерным*, когда исследователь устанавливает характер связей между несколькими переменными, такой вид анализа называется *многомерным*. Одномерный анализ используется практически в каждом исследовании; большинство исследователей проводит также и двумерный анализ. Многомерный анализ (факторный, кластерный и др.) требует достаточно высокого уровня подготовки в области математической статистики, необходимой, если не для того, чтобы уметь «вручную» выполнять те процедуры, которые в настоящее время «помогает» выполнить компьютер, то, по крайней мере, для того, чтобы подобрать адекватный метод, «чувствовать» смысл и содержание значений получаемых числовых результатов. Поэтому тем исследователям, которые хотят овладеть процедурами многомерного анализа, можно посоветовать обратиться к специальной литературе. В данном пособии мы подробнее остановимся на процедурах, проводящихся в рамках одномерного и двумерного видов анализа, которые чаще всего используются в технологии стандартного анализа результатов массовых опросов. Мы приведем основные правила описания, представления и интерпретации полученных результатов и остановимся на наиболее распространенных ошибках, допускаемых исследователями на этом этапе работы.

* В социологии и психологии существуют методы взвешивания, позволяющие не условно, а достаточно корректно преобразовывать порядковые шкалы в интервальные. С методами корректного преобразования порядковых шкал в интервальные можно ознакомиться в специальной литературе.

*Подробнее этот момент будет рассмотрен при изложении анализа методом группировки. Здесь же хотелось бы лишний раз привлечь внимание исследователей и пользователей социологической информации к осознанию условности различий в процентах, когда *численность анализируемой группы меньше 100 человек*.

* При таком преобразовании социолог — должен обратить особое внимание на согласованность названия индекса и порядка следования вариантов ответов с соответствующим ранговым кодом.

Лекция 8.

Анализ одномерных распределений

1. Дискретные шкалы: номинальная, порядковая (ранговая)

Частотные и процентные распределения

Приступая к обобщению полученного материала, исследователь, как правило, прежде всего формирует одномерные распределения по каждому пункту анкеты, т.е. подсчитывает, сколько респондентов отметили тот или иной вариант ответа на каждый из поставленных в анкете вопросов. Этот простейший вид распределения называется *частотным*. Частотным называется такое распределение, в котором напротив каждого варианта ответа указывается число респондентов в его *абсолютном* значении. Ниже приведен пример частотных распределений.

Вопрос: «*Укажите Ваш пол*»

	Частота
	(N)
Мужской	797
Женский	1010
Всего	1807 человек

Вторым шагом одномерного анализа является подсчет процентов, так как полученные распределения чаще представляют в процентах.

Вопрос: «*Укажите Ваш пол*»

(N = 1 807)

Мужской	44,1
Женский	55,9
Всего	100,0%

Такое представление данных называется *процентным* распределением.

Таблицы одномерных распределений по всем пунктам анкеты обеспечивают достаточно полную картину описания полученных данных в тех случаях, когда основная цель исследования — определение общественного мнения по тем или иным социальным проблемам.

Подготовка одномерных распределений («одномерок») — достаточно простая процедура. Но в практике социологического анализа исследователь может столкнуться с рядом вопросов, над которыми он не задумывался до начала анализа данных.

- 1) Как лучше представлять данные — в частотах или процентах?
- 2) По отношению к какому числу следует рассчитывать процент?
- 3) С какой степенью точности необходимо указывать данные?

Частоты или проценты?

Данные, как правило, представляют в процентах. Прежде всего потому, что в таком виде полученный результат можно сравнивать с аналогичными данными других опросов, или же сравнивать распределение мнений у различных групп населения и т.п. Часто полученный результат сам по себе еще ни о чем не говорит; лишь сравнение одних данных с другими позволяет получить определенную информацию. Например, если получены данные о том, что «угрозу возникновения межнациональных конфликтов» считают реальной 25% населения Украины, то эти результаты, хотя и представляют собой определенную информацию, но не позволяют сделать окончательный вывод о ситуации в области межнациональных отношений. Для соответствующего вывода необходимо сравнение этих данных с другими; если, допустим, в предыдущем году подобный ответ дали всего 8%, то можно фиксировать весьма неблагоприятную тенденцию; а если в предыдущем году так считало 40% населения, то можно сделать вывод о существенных положительных сдвигах в данной области. Можно также сравнивать долю населения, выражающую такие опасения в разных регионах и т.д. Проценты в большинстве случаев позволяют определить различия в пропорциях и проводить сравнительный анализ на основе количественных величин. Кроме того, данные, выраженные в процентах, значительно упрощают восприятие, что позволяет читателю без дополнительных вычислений «увидеть» полученную картину. Таким образом, возможность сравнения и простота восприятия обуславливают предпочтение процентных распределений.

Но в некоторых случаях, когда данные приводятся только в процентах, могут быть сделаны неверные выводы. Один из таких примеров приводится в книге Э.Ноэль «Массовые опросы»: «В одной из газет была помещена заметка о соотношении мужчин и женщин среди слушателей гессенской и баварской радиопередач, делающих объявления о найме на работу: «При сравнении с результатами по сфере действия гессенской радиопередачи можно утверждать, что доля женщин среди «слушателей за день» в Баварии значительно меньше, чем в Гессене — 61% против 68% в Гессене». После этого сообщения Баварская радиостанция потребовала исправлений или пояснений в абсолютных числах. На основании этой заметки можно было сделать вывод, что гессенскую радиопередачу слушает больше женщин. Таблица, дополненная абсолютными числами прояснила картину:

	Слушали рекламу о найме рабочей силы	
	в Баварии	в Гессене
Женщины	61% = 1 430 000	68% = 735 000
Мужчины	39% = 920 000	32% = 345 000
	100% = 2 350 000	100% = 1 080 000

В тех случаях, когда **информация касается изучения потребительского спроса**, то потенциального рекламодателя в первую очередь интересует распределение потенциальных потребителей в абсолютном, а не в относительном значении. Поэтому в определенных случаях необходимо приводить не столько полученные проценты, сколько абсолютные значения, а также их экстраполяцию на население в целом того региона, по отношению к которому была составлена выборочная совокупность. Это положение относится в основном к **маркетинговым исследованиям**, анализу **сбыта и потребления**, изучению **рекламы**, а также изучению некоторых политических вопросов (например, электората и электоральных прогнозов).

В любом случае, когда исследователь приводит результаты опроса в процентном распределении, он в каждой таблице должен привести данные, касающиеся общего числа людей, по отношению к которым приводимые проценты рассчитывались. Обычно абсолютные значения (общее число опрошенных, по отношению к которым считался процент) в таблицах обозначаются литерой «N». Можно также давать полный текст, характеризующий эту цифру.

Чтобы избежать искажения или возможностей неправильной интерпретации информации, полученной в результате опроса общественного мнения, следует руководствоваться следующим правилом: **в каждой таблице процентных распределений необходимо указывать абсолютное число людей, по отношению к которым вычислялся процент.**

По отношению к какому числу следует вычислять процент в одномерных распределениях?

Казалось бы, этот вопрос не должен вызывать сомнений; само собой разумеется, что по отношению к числу опрошенных. Но этот вопрос и не возникает, если на поставленный в анкете вопрос ответили все респонденты. Но довольно часто общее число опрошенных и число людей, давших содержательный ответ на поставленный вопрос, не совпадают. К этому приводят разные причины.

Во-первых, интервьюеры, несмотря на достаточно жесткие инструкции и контроль, все же могут пропускать кодировку некоторых ответов (не говоря уже об анкетах, полученных методом самозаполнения). Не следует забывать, что интервьюеры — живые люди, а процесс интервью — достаточно сложная эмоциональная ситуация (респонденты в процессе интервью нередко отвлекаются, дают слишком пространственные ответы; интервьюер ищет пути возвращения к теме интервью, и в какие-то моменты, даже получив ответ на поставленный вопрос, может не отметить его в бланке анкеты. Иногда, проверяя заполненный бланк, он может вспомнить, как именно ответил респондент. Но если он точно не помнит или пропустил сам вопрос в процессе интервью, то лучше в анкете оставить пропуск, чем вынуждать интервьюера отметить ответ, который не был получен от респондента).

Во-вторых, существуют и объективные причины, по которым на поставленный в анкете вопрос отвечают далеко не все респонденты. Среди них две являются самыми распространенными.

1. Бывают такие вопросы, которые не имеют отношение к данному респонденту. Например, вопрос, который приводился выше: «Удовлетворены ли вы своей работой?» не имеет отношение к тем, кто в настоящее время не работает. В данном случае можно отсеять неработающих, но в таблице следует указать абсолютное значение числа работающих. Может быть и более неопределенная ситуация. Например, вопрос касается оценки работы детских учреждений (детских садов, школ и т.п.): «Улучшились или ухудшились условия пребывания детей в детских садах?» Кто должен давать ответ на этот вопрос? На первый взгляд, вроде бы понятно — родители, имеющие детей соответствующего возраста. Однако, респондента может волновать этот вопрос, даже если у него самого нет таких детей (респонденты могут иметь внуков, братьев, сестер, племянников, детей знакомых, могут иметь профессиональное отношение к детским учреждениям; в связи с этим их очень беспокоит этот вопрос, и они имеют выраженное отношение к нему). Более того, родители, имеющие детей, могут не интересоваться и не иметь собственного мнения в отношении работы детских учреждений. В этих случаях исследователь получает несоответствие между общим числом респондентов и числом лиц, давших ответ на вопрос анкеты, а значит, не может достаточно точно определить контингент ответивших.

2. В другом случае некоторая часть респондентов (иногда достаточно большая), не хочет давать ответ на предложенный вопрос в силу тех или иных моральных соображений или не понимает формулировку вопроса, составленного автором анкеты. Довольно часто значительная часть респондентов не отвечает на вопросы интимного характера или на конкретные вопросы, касающиеся материального положения (охотно давая общую оценку своего материального положения, многие люди не хотят отвечать на вопросы, конкретизирующие условия их жизни). Например, на вопрос о том, что респонденты приобрели за последний год, ответ дали только чуть больше трети опрошенных, остальные респонденты отказались отвечать на этот вопрос. В результате, процентное распределение к числу ответивших существенно отличается от процентного распределения к числу опрошенных.

	% ко всем N=1807	% к ответившим N=683
Приобрели за последний год		
дача	0.9	2.3
садовый участок	4.0	10.7
легковая машина	1.6	4.2
цветной телевизор	3.6	9.5
новая мебель	3.1	8.2
библиотека	15.9	42.2
стерео- или видеоаппаратура	1.5	4.1
магнитофон, радиоприемник	4.7	12.4
спортивное, туристское снаряжение	1.8	4.8
холодильник	2.7	7.0
стиральная машина	2.5	6.6

швейная машина	0.7	1.8
модная одежда	16.2	42.8
нет ответа	62.2	

Оказавшись в такой ситуации (если она не была предусмотрена программой исследования), автор анализа нередко испытывает затруднение — результаты какого из распределений следует привести в аналитической таблице.

В подобной ситуации нередко оказывается исследователь не только при изучении проблем личного характера, когда не все люди хотят давать откровенные ответы на предлагаемые вопросы, но и при изучении политических пристрастий и ориентаций (поддержка населением различных политических партий или лидеров, электоральные ориентации и т.п.). На вопросы, касающиеся политических предпочтений, многие респонденты часто не отвечают не потому, что не хотят высказать свое мнение, а потому, что оно у них еще не сформировалось. Если на поставленный в анкете вопрос, в силу тех или иных причин, не ответила значительная часть респондентов, то исследователь испытывает большое затруднение — какому из процентных распределений следует отдать предпочтение при анализе и интерпретации результатов опроса.

Например, на вопрос: «Доверяете ли Вы N?» были получены следующие процентные распределения.

	% ко всем N=1807	% к ответившим N=1091
скорее доверяю	20.8	34.6
трудно сказать, доверяю или нет	16.0	26.6
скорее не доверяю	23.5	38.8
нет ответа	39.7	
Всего	100.0	100.0

В данном случае около сорока процентов населения по тем или иным причинам фактически отказались отвечать на предложенный вопрос. Какой вывод может сделать исследователь? Должен ли он ориентироваться только на ту часть населения, которая отвечала на данный вопрос, и делать вывод о том, что данному политику доверяют около 35% населения; или же можно говорить лишь о том, что данному политику доверяет 21%?

Нами приведены лишь некоторые примеры многочисленных проблемных ситуаций, возникающих перед исследователем уже на первом этапе анализа на самом простом уровне — описание одномерных распределений.

В каждом конкретном случае автору анализа необходимо находить свое **обоснованное** решение. Если же автор не может принять обоснованное решение, ему можно рекомендовать воспользоваться общими принципами представления данных одномерных распределений.

1. В таблице следует приводить полную формулировку вопроса (в том виде, в котором он был представлен в анкете).

2. Если доля неответов на данный вопрос менее 5%, то в большинстве случаев можно приводить процентное распределение к общему числу опрошенных (респондентов).
3. Если доля неответов превышает 5%, то лучше приводить оба процентных распределения (ко всем опрошенным и к числу ответивших).
4. В каждой таблице необходимо указывать абсолютное число людей, по отношению к которому вычислялся приводимый процент. В том случае, когда приводятся два процентных распределения, в таблице следует указывать как абсолютное число лиц, принимавших участие в опросе, так и абсолютное число людей, давших ответ на данный вопрос.
5. В тех случаях, когда шкала номинальная, в таблице не приводятся кодовые значения вариантов ответов; в тех случаях, когда шкала порядковая, в таблице следует приводить кодовые значения вариантов ответа.

***С какой степенью точности следует
приводить данные в процентах?***

При оформлении результатов анализа данные процентных распределений приводят, как правило, с точностью максимум до одной десятой. При этом следует учитывать, что даже один десятичный знак затрудняет восприятие данных, так как потребитель информации, сравнивая цифры, мысленно производит округление. Поэтому в тех случаях, когда можно отказаться даже от десятых, лучше производить округление до целых чисел. Следовательно, решая проблему округления, автор фактически должен выбрать один из двух вариантов: приводить данные в целых числах или указывать десятичные значения. Чем должен руководствоваться исследователь, решая эту проблему?

Общий принцип сводится к тому, что чем больше выборка, тем более вероятно, что десятые доли процента могут представлять значимое информационное значение.

При этом не следует забывать, что при представлении данных в таблице при небольших выборках (в несколько сотен единиц наблюдения) десятичные знаки не только затрудняют восприятие таблицы, но и могут вводить читателя в заблуждение. Например, если автор анализа, приводя сравнительные данные о доверии населения различным политическим деятелям, указывает, что одному доверяют 11.5%, другому 11.9%, а третьему — 12.4% (при том, что выборку составляют 400 человек), то может сложиться впечатление, что существует определенное различие в уровне доверия этим политическим фигурам, тогда как при такой выборке эти различия статистически не значимы. Совсем другой вывод можно было бы сделать, если бы при этих же данных было опрошено 10 000 человек.

В каждом конкретном случае можно подсчитать — какая степень точности статистически обоснована, а какая — нет. Но если решать вопрос о точности округления в целом, то можно руководствоваться следующими общими правилами.

1. Совершенно недопустимо, чтобы в одной таблице одни данные были представлены в целых числах, а другие — с точностью до десятых. Правило, которому необходимо следовать в любом случае, — все данные должны приводиться с одинаковой степенью точности.

2. Никогда не следует забывать, что если процент берется от числа, значение которого меньше ста, то исследователь оперирует с условными процентами. Когда мы говорим «двадцать процентов населения поддерживают...», то подразумевается, что «поддерживают...» двадцать человек из каждых ста с определенной вероятностью. (Степень этой вероятности определяется численностью опрошенных и ее репрезентативностью по отношению к генеральной совокупности). Переходя к анализу различий в полученных данных, выраженных в процентах, и отвлекаясь от абсолютной численности анализируемой группы, автор анализа нередко забывает, что говорить о «двадцати процентах» от 10-20 человек просто смешно, потому что речь идет о мнении 2-4 человек, и любое случайное изменение (в опрос попал другой респондент данной группы; человек ответил, не подумав; в конце концов, оператор ошибся при вводе информации и т.п.), в итоге обусловит совершенно иную картину в процентном распределении.

3. При объеме выборочной совокупности от 100 до 1000 человек данные процентных распределений обычно лучше приводить в целых числах. В этом случае различия в данных, приводимых в целых числах, или с указанием соответствующих десятых, не принципиальны, но воспринимать данные значительно удобнее в целых числах.

Исключение составляет освещение результатов мониторинговых исследований. Например, в тех случаях, когда проводится ежемесячный (ежегодный) опрос по одним и тем же переменным, различия в результатах опроса от месяца к месяцу могут носить не значимый характер; однако фиксируемая устойчивая тенденция сама по себе представляет определенную информацию.

4. При объеме выборочной совокупности более 1000 человек, данные в большинстве случаев лучше приводить с десятыми, так как в тех случаях, когда выборка составляет несколько тысяч человек, различие в десятых долях процента может быть не только показательным, но и статистически значимым.

Эти правила носят достаточно общий характер и предназначены в качестве рекомендации исследователю, испытывающему затруднения, связанные с представлением результатов опроса. В каждом конкретном случае автор анализа может самостоятельно решать — с какой степенью точности ему следует представлять материал. Главное, чтобы это решение было достаточно осознанным и обоснованным.

2. Непрерывные шкалы: интервальная, метрическая

Меры центральной тенденции

Если для шкал дискретного типа в качестве описательных статистик используются в основном частотные и процентные распределения, то для шкал непрерывного типа — статистики, позволяющие охарактеризовать среднее значение полученных данных. Напомним, что к непрерывным относятся метрические и порядковые шкалы (в тех случаях, когда у автора есть основания рассматривать порядковую шкалу как интервальную).

Основными обобщающими характеристиками данных, полученных по непрерывным шкалам, являются *средние величины* или *меры центральной тенденции*: среднее

арифметическое, мода и медиана. Меры центральной тенденции позволяют определить и описать наиболее типичную величину полученных данных.

Среднее арифметическое — величина, которую наиболее часто используют при характеристике изучаемого контингента людей по различным метрическим шкалам. Например, если в анкете задавался вопрос о заработной плате, то, характеризуя полученные данные, указывают среднюю заработную плату и т.п.

Для подсчета среднего арифметического значения выписываются ВСЕ значения полученной переменной (обычно в порядке возрастания) и напротив каждого значения указывается количество лиц, отметивших данный ответ в анкете. Приведем пример вычисления среднего арифметического значения возраста опрошенных старшеклассников (для удобства восприятия и расчетов ниже приводятся упрощенные вымышленные примеры). Результаты опроса выписываются следующим образом:

Возраст	N
14	40
15	60
16	200
17	72
18	28
Всего	400(человек)

Затем значение каждого признака умножается на количество лиц, обладающих данным признаком, и полученные произведения складываются:

14x	40=	560
15x	60=	900
16x	200=	3200
17x	72=	1224
18x	28=	504
Всего	400	6388

Среднее арифметическое вычисляется путем деления суммы произведений каждого значения признака на число лиц, обладающих данным признаком, на общее количество опрошенных:

$$6388 : 400 = 15.97$$

В данном случае средний возраст опрошенных старшеклассников составляет 15.97 лет. Существуют и другие средние показатели: *мода* и *медиана*.

Под модой подразумевается наиболее часто встречающееся в ряду распределения значение признака. В приведенном примере модой является значение 16 лет, так как такой возраст характерен для 200 человек (самое высокое значение).

Под медианой подразумевается значение признака, которое находится в середине ранжированного ряда. В данном случае это значение тоже равно 16 годам.

Среднее арифметическое, мода и медиана — основные характеристики, позволяющие описать основное (типичное) свойство данных, полученных по массиву респондентов. В

процессе описания данных перед исследователем стоит задача — какой из этих показателей выбрать. Это зависит от типа шкалы изучаемого признака.

Чем выше измерительный уровень шкалы, тем больше характеристик можно использовать для ее описания. Если для дискретных шкал можно пользоваться только такими характеристиками, как мода и медиана, то для непрерывных шкал можно считать среднее арифметическое.

Среднее арифметическое — величина, позволяющая получить *определенную* характеристику изучаемого объекта (например, по ней можно количественно сравнивать сотрудников разных предприятий по среднему возрасту, средней зарплате и т.п.). Однако этот показатель сам по себе достаточно односторонний, что иногда возмущает людей, не только не имеющих специальных знаний в области математической статистики, но и даже с низким уровнем образования: «Как можно говорить о средней заработной плате у нас на предприятии (в стране), если я получаю так мало, а «начальство» имеет такую высокую зарплату!». Чувство «социальной справедливости» даже на житейском уровне подсказывает, что нельзя говорить о том, что «в среднем» у населения высокий жизненный уровень, когда определенная часть получает «очень много», достаточно большая часть — «очень мало», и неизвестно сколько получает ту самую «среднюю величину». Здесь на интуитивном уровне определяется, что одного только среднего значения недостаточно, чтобы охарактеризовать картину по определенному показателю. Но интуитивный житейский скептицизм срабатывает, к сожалению, тогда, когда речь идет о вещах, затрагивающих непосредственные интересы людей. При анализе достаточно «абстрактных» показателей начинающие социологи нередко не учитывают «достижения» статистики, в частности, необходимости дополнять данные по средним значениям полученных показателей такими характеристиками, как рассеяние (изменчивость, колеблемость) признака и форма распределения.

Используя вышеприведенный пример, проиллюстрируем значение показателей изменчивости признака. Мы рассчитывали данные по среднему возрасту обследованных старшеклассников. Среднее арифметическое значение указывает, что их средний возраст около 16 лет. Если читателю представить весь ряд распределения, то можно увидеть, что самый младший из опрошенных был 14 лет, а самый старший — 18, т.е. возраст изучаемого контингента отклонялся от средней величины, максимум, на два года. Но по метрическим шкалам, включающим значительное число значений признака, общее распределение, как правило, не приводится, а одно лишь значение среднего арифметического может ввести в заблуждение, так как если бы мы в той же школе опросили 350 детей 9-10 лет и 50 учителей в возрасте 50-60 лет, то среднее арифметическое всех опрошенных тоже равнялось бы примерно 16 годам. В данном случае совершенно разные контингенты опрошенных имели бы одинаковое значение этого параметра, так как во втором случае опрашивался неоднородный контингент, в частности, по такому показателю, как возраст.

Таким образом, для того, чтобы получить адекватное представление об изучаемом признаке, одного среднего значения явно недостаточно.

В общем виде для отбора статистических процедур, используемых для описания одномерных распределений, в зависимости от типа шкалы можно руководствоваться следующей схемой.

Тип шкалы	Средние величины	Разброс (меры
-----------	------------------	---------------

	(меры центральной тенденции)	изменчивости)
Дискретный		
номинальные	мода	
порядковые/ранговые	медиана, мода	ранг, минимум, максимум
Непрерывный		
метрические, порядковые/интервальные	среднее арифметическое, медиана, мода	стандартное отклонение, коэффициент вариации, ранг, минимум, максимум

Квалифицированный статистический анализ средних значений заключается в том, что, приводя значение среднего арифметического, исследователь должен, привести и данные по значению изменчивости этого признака (вариации, рассеянию, колеблемости) — величины отклонений от среднего значения.

Меры изменчивости

Наряду с мерами центральной тенденции, для описания данных необходимо приводить и параметры, характеризующие степень изменчивости (вариации, рассеяния) признака. Существует целый ряд величин, позволяющих оценить меры рассеяния признака. В статистике обычно приводится ряд параметров, позволяющих оценить изучаемую совокупность в целом. Это указание *минимума* и *максимума*. Минимум — это наименьшее значение признака, полученное в результате измерения (в приведенном примере — 14 лет в первом массиве опрошенных и 9 лет — во втором). Максимум, соответственно, — наибольшее значение (в приведенном примере 8 лет в первом массиве и 60 — во втором). Если исследователь при описании полученных данных, дополнительно к указанию среднего возраста привел бы только эти величины, то уже было бы понятно, что описываются разные возрастные контингенты. Указание границ значения признака называется *вариационный размах*. Но и этого параметра недостаточно для адекватного представления о характеристике возраста опрошенных.

Наиболее информативным параметром рассеяния являются такие показатели, как *стандартное отклонение* и *дисперсия*.

Стандартное отклонение и дисперсия — взаимосвязанные величины, показывающие среднюю величину отклонения (всех опрошенных) от среднего значения. Стандартное отклонение позволяет, во-первых, судить, насколько далеки в среднем значения всех опрошенных от среднего арифметического; во-вторых, стандартное отклонение является величиной, которая лежит в основе многих видов многомерного анализа.

В настоящее время у исследователя нет необходимости вручную считать стандартное отклонение: практически все компьютерные программы обработки социологических данных позволяют автоматически получать значение этого показателя.

В задачу же исследователя входит, во-первых, самому обратить внимание на этот индикатор (в первую очередь на то, чтобы значение стандартного отклонения не превышало значение среднего арифметического); а во-вторых, обязательно приводить этот показатель в аналитических таблицах (там, где приводится значение средних арифметических, рядом необходимо привести значение стандартного отклонения).

Таким образом, мы видим, что, характеризуя данные, полученные по непрерывным шкалам (метрической и интервальной), исследователь, наряду со средним обобщающим показателем изучаемого признака, описывает и характер его разброса (изменчивость признака).

Форма распределения

Существует еще одна характеристика распределения данных, полученных по непрерывным шкалам, которую исследователь тоже должен обязательно учитывать. Это *форма распределения*.

Данные распределения старшеклассников по возрасту являются примером *нормального распределения*. Нормальным является такое распределение, при котором кривая построенного по его данным графика представляет собой колоколообразную симметричную кривую.

Например, если мы построим график по данным распределения старшеклассников по возрасту, то получим соответствующую колоколообразную кривую. Если же мы построим график по массиву третьеклассников и учителей, опрошенных в одной школе, мы получим две кривые. Нормальное распределение — это теоретическая кривая. Практически любые эмпирические данные в той или иной степени отклоняются от нормального - распределения вероятностей, закону которого подчиняются распределения случайных величин. Но поскольку все расчеты, включающие значение среднего арифметического и стандартного отклонения, основаны на теории вероятности, в аналитическую задачу исследователя входит оценка (по крайней мере, приблизительная) того, насколько правомерно использовать данный тип анализа к полученным результатам. Поэтому даже на уровне описания (не говоря уже о множественном анализе), прежде чем приводить данные по их средним значениям (среднее арифметическое и стандартное отклонение), необходимо оценить характер формы распределения — в какой степени она приближается к нормальному распределению.

Для этого используют показатели скоса (асимметрии, *skewness*) и эксцесса (*kurtosis*). В скобках указываются термины, которые обычно у разных авторов используются для обозначения одних и тех же понятий. В частности, здесь приведены англоязычные обозначения рассматриваемых характеристик, которые приводятся в компьютерной программе обработки и анализа социологических данных — SPSS.

Показатель скоса (skewness) позволяет оценить степень и направленность асимметрии кривой распределения. В случае идеального нормального распределения асимметрия равна нулю.

В эмпирической социологии идеальные нормальные распределения практически не встречаются. Но существуют методы оценки степени приближения полученного распределения к нормальному. Коэффициент скоса имеет числовое значение и знак, указывающий направленность скоса. Чем больше величина отличается от нуля, тем большая асимметрия у полученного распределения, и, соответственно, большая погрешность может проявиться при применении коэффициентов статистического анализа, формула которых включает значения стандартного отклонения.

Существуют специальные процедуры оценки степени допустимости такой погрешности, а также искусственной нормализации шкалы. Исследователь может, при необходимости, осуществлять преобразование данных. С различными способами преобразования данных можно ознакомиться в специальной справочной и учебной литературе; но исследователю необходимо обязательно оценить степень асимметрии. (Простейшим косвенным показателем, указывающим на асимметрию, является расхождение между значениями среднего арифметического, моды и медианы; при идеальном нормальном распределении все три показателя равны).

Показатель эксцесса (kurtosis) показывает, в какой степени «крутизна» полученной кривой приближается к нормальному распределению.

Показатели асимметрии и эксцесса необходимы исследователю в первую очередь для того, чтобы он мог установить — в какой степени в анализе может быть использовано стандартное отклонение.

Таким образом, на первом этапе анализа (описания данных), представляя данные, полученные с помощью непрерывных шкал (метрических и интервальных), исследователь должен руководствоваться следующими правилами.

1. Рассмотреть среднее арифметическое (mean), стандартное отклонение, коэффициент асимметрии и коэффициент эксцесса. Эти показатели позволяют решить, можно ли эти данные использовать в основном анализе или они требуют дополнительных преобразований.
2. Если исследователь не овладел техникой преобразования данных, ему следует отказаться от представления полученных по этому показателю значений среднего арифметического в итоговом научном документе (отчете, статье, таблицах и т.п.). В противном случае, информация будет носить неадекватный, искаженный характер.
3. В тех случаях, когда эти показатели удовлетворительны, данные по метрическим шкалам обычно представляют в итоговых таблицах по следующим параметрам: 1) N (численность группы, по которой считалось среднее значение); 2) величина среднего арифметического; 3) величина стандартного отклонения. Эти параметры позволяют в дальнейшем при необходимости посчитать статистическую значимость различий между средними значениями у различных групп населения. Допустим, если автор анализа приводит данные по средней заработной плате различных групп населения, то эти параметры позволят не только сравнить — насколько различается средняя заработная плата у каких-либо двух категорий работников, представленных в таблице, но и оценить — является ли это различие статистически значимым (с какой степенью

вероятности мы получим то же самое различие при повторных исследованиях на других аналогичных выборках).

Такое представление данных позволяет переходить на следующий уровень анализа — объяснение и интерпретация полученных данных.

Лекция 9.

Кросс-табуляция: анализ двумерных таблиц

1. Роль статистики в процессе социологического анализа

Возможности и ограниченность

Если на первом уровне анализа основной задачей исследователя являлось адекватное и корректное *описание* полученных фактов, то ко второму уровню автор переходит с целью поиска *объяснения* полученным фактам.

Следует сразу же подчеркнуть основную особенность статистического анализа данных, полученных в результате массового опроса. Сами по себе статистические процедуры не приводят к объяснению полученных фактов, они являются лишь инструментом проверки гипотез, выдвигаемых исследователем. Применяя те или иные статистические процедуры, автор получает ответ на свой собственный запрос — подтверждается или не подтверждается полученными в результате конкретного опроса данными проверяемая гипотеза.

Роль исследовательской гипотезы в процессе анализа данных

Если гипотеза подтверждается, то статистический инструментарий позволяет установить вероятность правильности вывода. Если гипотеза не подтверждается, то это еще не означает, что она была неверной, так как причина может заключаться в неполноте программы, неадекватности инструментария, погрешностях организации сбора данных и т.п.

Таким образом, основным инструментом содержательного анализа изучаемых социальных процессов являются аналитические возможности исследователя (его знание предмета исследования, логика, инструментальная компетентность), позволяющие ему в процессе теоретического анализа выдвинуть гипотезы, эмпирическая проверка которых открывает возможность объяснять причины исследуемых процессов или явлений. Статистический анализ позволяет лишь подтвердить (или не подтвердить) выдвинутые гипотезы эмпирическим материалом. Никакой статистический анализ не поможет объяснить факты причинами, которые не были заложены в инструментарий исследования. Исследователь всегда должен помнить, что реальные причины, влияющие на исследуемые факты, могли

остаться вне его поля зрения (или, по крайней мере, вне инструментария). Если автор не предусмотрел возможное влияние тех или иных причин, когда готовил инструментарий исследования, то эти причины и не будут обнаруживаться в процессе анализа полученных данных.

После этой преамбулы, суть которой не следует забывать в процессе статистического анализа полученных данных, рассмотрим основную технологическую схему двумерного анализа, позволяющую перейти от статистического описания результатов исследования к использованию статистических процедур в объяснении полученных фактов.

2. Кросс-табуляция

Зависимые и независимые переменные

Основой анализа связи между двумя переменными является подготовка двумерных таблиц (cross-tabulation), представляющая собой взаимное пересечение двух рядов распределений.

Переменные для рядов распределения определяются гипотезами исследователя. Одна переменная рассматривается как фактор (причина), влияющая на исследуемое явление; другая — как показатель этого явления. Один и тот же признак (представленный в инструментарии вопросом анкеты) в процессе анализа может выступать и в качестве показателя и в качестве фактора. Например, одна гипотеза может быть сформулирована так: «Уровень образования влияет на удовлетворенность работой: чем выше уровень образования, тем выше уровень удовлетворенности работой». Другая гипотеза, требующая проверки, может быть следующей: «Удовлетворенность работой влияет на общую удовлетворенность человека жизнью: чем выше удовлетворенность работой, тем выше удовлетворенность жизнью». В первом случае удовлетворенность работой выступает в качестве показателя, а во втором — в качестве фактора. Когда признак в процессе статистической проверки гипотезы рассматривается в качестве фактора, он называется *независимой переменной* (подлежащим); когда признак рассматривается в качестве показателя, в статистическом анализе он называется *зависимой переменной* (сказуемым).

Содержательный смысл таблицы, построенной на основе результатов кросс-табуляции, заключается в следующем. По признаку, рассматриваемому в качестве независимой переменной, исследователь разбивает всю совокупность опрошенных на отдельные группы (подгруппы). Наряду с общим результатом одномерного распределения по изучаемому показателю, исследователь параллельно получает результаты распределения мнений отдельных подгрупп населения, и эти результаты сравнивает между собой. Чем больше различие в характере ответов у этих групп, тем о большей связи между двумя переменными можно говорить. Обращаясь к приведенному выше примеру, можно сказать, что, проверяя первую гипотезу, исследователь всю совокупность разбивает на группы с различным уровнем образования (здесь, напомним, проверять гипотезу будет тем легче, чем более произведенная группировка соответствует выдвинутой гипотезе: если речь идет об **уровне** образования, то полученные группы должны различаться между собой именно по **уровню** образования); проверяя же вторую гипотезу, исследователь разбивает всю совокупность на группы людей, различающихся между собой по уровню удовлетворенности работой.

Построенная двумерная таблица позволяет «на глаз» увидеть, различаются ли ответы у разных групп населения. В таком виде обычно и рассматривает результаты исследования заинтересованный потребитель информации. Но результаты, оцениваемые «на глаз», часто могут вводить в заблуждение.

Поэтому основной задачей исследователя-аналитика, представляющего результаты опроса в виде двумерных распределений является *статистическая экспертиза* — оценка и отбор качественной информации. Просмотрев и статистически оценив сотни (а то и тысячи) двумерных таблиц, он должен для широкой аудитории (и для своего дальнейшего анализа) отобрать только те, которые соответствуют статистическим критериям качества информации. Другими словами, исследователь, прежде чем публиковать результаты двумерного анализа, должен осуществить статистическую экспертизу, позволяющую либо «пропустить» результаты исследования для широкого обнародования (в том случае, если полученные данные статистически обоснованы), либо «отбраковать» таблицы (в тех случаях, когда видимые «на глаз» различия статистически недостоверны в силу тех или иных причин, и могут привести к неверным выводам). Если исследователь не удовлетворен результатами статистической проверки (другими словами, интуиция подсказывает ему, что гипотеза была верной), он может в дальнейших исследованиях проверять ее, совершенствуя инструментарий и организацию сбора информации. Но он не должен представлять для широкой аудитории данные, не соответствующие статистическим критериям. Прежде чем рассмотреть особенности использования статистических показателей для последовательного осуществления шагов объяснительного анализа, рассмотрим двумерную таблицу, с точки зрения описания и интерпретации кросс-табуляции полученных данных.

Описание и статистическая экспертиза двумерных таблиц

Самая простая процедура — с помощью компьютера получить таблицу двумерного распределения. Допустим, исследователь хочет проверить гипотезу о влиянии материального уровня жизни на общую удовлетворенность жизнью. Более того, он считает, что на удовлетворенность жизнью влияет не столько реальный уровень материального благополучия, сколько самооценка человеком своего материального статуса. Допустим, что для измерения самооценки материального статуса в инструментарий был включен вопрос: «К какой группе людей по уровню материальной обеспеченности Вы бы себя отнесли?». Полученные данные показали, что из 1755 опрошенных 4 человека отнесли себя «к высокообеспеченным», 909 — к «людям среднего достатка», 842 — к «малоимущим». Объявляя этот признак независимой переменной, а ответы на вопрос об удовлетворенности жизнью (с веером ответов, включающим пять вариантов) — зависимой переменной, исследователь может получить два типа двумерных таблиц:

- 1) шкала ответов на вопрос об удовлетворенности рассматривается как дискретная (порядковая/ранговая);
- 2) шкала ответов на вопрос об удовлетворенности рассматривается как непрерывная (порядковая/интервальная).

Таблица 1 (первого типа)

**Влияние самооценки материального статуса
на удовлетворенность жизнью**

	Удовлетворены ли Вы своей жизнью в целом? (в % к каждой группе)					Всего
	1	2	3	4	5	
	Совершенно удовлетворен	Скорее не удовлетворен	Трудно сказать, удовлетворен или нет	Скорее удовлетворен	Полностью удовлетворен	
К какой группе людей Вы бы Себя отнесли?						
Высокообеспеченным	0,0	50,0	0,0	50,0	0,0	100,0
Среднего достатка	11,2	28,1	28,7	28,2	3,9	100,0
Малоимущим	23,0	37,1	26,5	12,2	1,2	100,0

В общем обзоре таблицы просматривается взаимосвязь между самооценкой материального статуса и удовлетворенностью жизнью. Однако количественное сравнение различных групп по их общей удовлетворенности жизнью при шкале с повышенной степенью точности (в данном случае — 5 градаций) довольно затруднительно.

Обычно, представляя данные по дискретным шкалам, производят укрупнение групп по зависимой переменной (для удобства сравнения между группами по независимой переменной), преобразуя вид таблицы следующим образом.

	Удовлетворены ли Вы своей жизнью в целом? (в % к каждой группе)			Всего
	Не удовлетворен	Трудно сказать, удовлетворен или нет	Удовлетворен	
К какой группе людей Вы бы Себя отнесли?				
Высокообеспеченным	50,0	0,0	50,0	100,0
Среднего достатка	39,3	28,7	32,1	100,0
Малоимущим	60,1	26,5	13,3	100,0

Преобразованная таким образом таблица позволяет не только «увидеть», что среди лиц «среднего достатка» больше людей, удовлетворенных жизнью, и меньше неудовлетворенных по сравнению с «малоимущими», но и подсчитать, насколько (или во сколько раз) их больше. Если исследователь, рассматривая рабочую таблицу (в которой указывается не только процент, но и абсолютное значение численности полученных

групп), не обратит внимание на то, что к «высокообеспеченным» себя отнесли всего 4 человека (в двумерных распределениях, подготовленных к публикации, абсолютные значения часто не указываются), то он может пытаться найти объяснение тому, что среди «высокообеспеченных» «половина» людей не удовлетворена жизнью, и делать далекоидущие выводы о том, что «не в деньгах счастье», и что повышение уровня жизни выше среднего достатка приводит к повышению доли лиц, не удовлетворенных жизнью и т.д., и т.п.

На самом деле, данные этого опроса не позволяют делать никаких выводов в отношении лиц, относящих себя к высокообеспеченным, поскольку таковых в выборке оказалось всего 4 человека.

Важнейшим правилом при представлении данных многомерных распределений является необходимость учитывать численность полученных групп. Конечно внимательный исследователь даже на уровне здравого смысла догадается, что группа численностью в четыре человека не может репрезентировать разброс мнений данной категории населения. Возникает вопрос — какой же численности должна быть полученная группа, чтобы по данным ее ответов можно было делать выводы о соответствующей категории населения, сравнивая ее с другими категориями (20, 30, 50, 100 и т.д. человек?). На этот вопрос нет однозначного ответа, поскольку на репрезентативность полученных данных влияют разные факторы: это и количество характеристик, по которым выборочная группа репрезентирует данную категорию населения, и разброс мнений по исследуемому показателю, и форма распределения и т.д. Но в статистическом анализе существует такой критерий, как **значимость различий**. Не вдаваясь в математические тонкости формулы расчета значимости различий долей (процентов), отметим лишь **необходимость** расчета статистической значимости различий в любом случае, когда автор рассуждает о различии в мнениях сравниваемых групп населения; и поясним в самом общем виде, о чем говорит критерий значимости различий долей.

В выборочном социологическом опросе исследователь обычно опрашивает определенную часть населения (выборку), а выводы делает относительно всей исследуемой совокупности (генеральной совокупности); эти же допущения лежат и в основе групп, которые получает автор в процессе анализа (ведь когда он говорит о «высокообеспеченных» или о «малоимущих», он имеет в виду все население, ограниченное лишь рамками генеральной совокупности). Предполагается, что выборку автор составил достаточно обоснованно. Но по тем же критериям отбора он может подготовить еще одну выборку. Разумеется, маловероятно, что он получит абсолютно идентичные результаты опроса. Формула расчета значимости различий позволяет учесть доверительную вероятность повторения такого же результата на идентичных выборках: сколько раз будет получен тот же самый результат, если бы исследователь повторил опрос 100 раз на идентичных выборках? Обычно принято указывать достоверность различий путем расчета коэффициента, показывающего вероятность неправильного решения. Если мы говорим о том, что различия значимы на уровне 5% (0,05), то это означает, что в 95 случаях из 100 мы получим тот же результат.

В формулу расчета включаются такие показатели, как численность анализируемой группы, а также доля (процент) лиц давших такой вариант ответа, по которому фиксируется различие. В настоящее время нет необходимости вручную рассчитывать коэффициент значимости различий долей (обычно все программы статистической обработки и анализа позволяют рассчитывать его автоматически). Задача же исследователя заключается в необходимости установления коэффициента значимости различий, прежде разговора о самих различиях. В социологическом анализе принято в

качестве верхнего предела рассматривать 5%-ый уровень значимости. Другими словами, при анализе социологических данных, различия в результатах, соответствующий коэффициент для которых больше значения 5% (0.05), признаются статистически незначимыми. Следует, конечно, осознавать условность этой границы. Исследователь может представлять данные и с более низким порогом значимости, если считает их достаточно важными. Но в таких случаях он обязательно должен оценить значимость различий процентов, учитывать при общем анализе материала и обязательно указать ее в тексте итогового документа. Но это можно делать только в случаях, когда готовится научный документ для специалистов. В общем случае, различия, полученные в результате двухмерного анализа, признаются статистически недостоверными, что и требуется подчеркивать, приводя данные в виде двумерных таблиц.

Порядковая шкала, которую автор условно принимает за интервальную, позволяет представить данные двумерного анализа в несколько ином виде (см. таблицу 2). Такой вид данных предпочтительнее, когда шкала ответов имеет достаточно высокую точность: в таком виде легче сравнивать различные группы между собой, так как сравнение можно проводить по одному числу; в таком виде легче увидеть определенные тенденции. Кроме того, анализируемые группы не разбиваются на дополнительные подгруппы в соответствии с ответами по зависимой переменной.

Таблица 2 (второго типа)

Влияние самооценки материального статуса на удовлетворенность жизнью

К какой группе людей Вы бы себя отнесли?	Уровень удовлетворенности жизнью (в баллах; шкала: 1-5)		
	N	Среднее арифметическое	Стандартное отклонение
Высокообеспеченным	4	3,00	1,33
Среднего достатка	909	2,85	1,07
Малоимущим	842	2,31*	1,00

** Различие между группами «среднего достатка» и «малоимущими» значимо на уровне 0.01*

Какие основные требования следует учитывать, представляя данные в таком виде?

Поскольку исследователь данную шкалу определил как интервальную достаточно условно, он должен быть уверен, что полученное распределение по своей форме приближается к нормальному. Мы уже говорили о тех показателях, которые позволяют оценить форму распределения (асимметрия, эксцесс). Но если даже исследователь их не вычислил, то, по крайней мере, он может сделать самую грубую прикидку — обратить внимание на то, чтобы стандартное отклонение не превышало величину средней. Если стандартное отклонение превышает значение среднего арифметического, то результаты анализа в таком виде (значение средней) представлять нецелесообразно.

Так же, как и в случае сравнения данных в процентах, исследователь должен учитывать уровень значимости различия средних. В формулу расчета входят значения таких параметров, как численность группы, среднее арифметическое и стандартное отклонение.

Как следует оформлять таблицу, если она приводится в контексте анализа? Это зависит от целевого назначения текста.

Если таблицы приводятся в документах сугубо научного или отчетного характера, то все эти три параметра обязательно должны включаться в соответствующую таблицу, чтобы читатель мог сам пересчитать значимость различий, если его заинтересовали полученные данные.

В научных публикациях автор прежде всего несет ответственность за те основные положения, которые он обосновывает полученными данными. Поэтому в тех случаях, когда различия имеют принципиальный для анализа характер (автор отмечает их в тексте, или сами по себе они представляют социальный интерес, например, сравнивается доверие к различным политическим лидерам и т.п.), исследователь считает значимость различий и указывает результаты расчета в таблице (см. таблицу 2). Обычно принято одной звездочкой указывать различие, уровень значимости которого 0.05, двумя звездочками — различие, уровень значимости которого 0.01, и тремя звездочками обозначать различие, уровень значимости которого 0.001. В соответствующей части текста обычно принято употреблять слово «различие» («различаются»), если это различие значимо на уровне 0.05; и «существенное различие», если видимое в таблице различие значимо на уровне 0.01 (и тем более на уровне 0.001).

Если вернуться к данным таблицы 2 и посмотреть полученные результаты — в одном случае без учета коэффициента значимости различий, а в другом, учитывая этот статистический критерий, то выводы будут различаться. В первом случае, если автор не будет учитывать значимость различий, он может сделать следующий вывод: «Данные позволяют обнаружить тенденцию повышения уровня удовлетворенности жизнью с повышением оценки материального статуса, причем, с переходом на уровень «высокообеспеченных», уровень удовлетворенности жизнью возрастает незначительно». Но на самом деле, полученные данные не позволяют делать именно такой вывод; как и следовало ожидать, ввиду малочисленности группы «высокообеспеченных», ее отличие по показателю жизненной удовлетворенности статистически не значимо не только по отношению к группе «среднего достатка», но и по отношению к «малоимущим». Из этого не следует, что группа «высокообеспеченных» не отличается от других групп населения по уровню жизненной удовлетворенности; а следует только вывод, что полученные данные не позволяют судить об уровне удовлетворенности жизнью «высокообеспеченных», в том числе и сравнивать этот показатель с другими группами.

Сравнительный анализ малочисленных, с точки зрения возможностей статистических процедур, социальных групп, как уже отмечалось, очень распространенная ошибка в публикациях, посвященных изложению результатов социологических опросов.

Поэтому, когда исследователь в результате двумерного анализа обнаруживает, что ответы на тот или иной вопрос анкеты представителей малочисленных групп существенно отличаются от ответов других групп населения, он всегда должен посчитать значимость этих различий. И если различия статистически незначимы, он, на основании данных только своего исследования, не может сказать, отражают ли полученные данные реальную картину, или они вызваны нерепрезентативностью выборочной подгруппы по отношению к соответствующей категории населения. Короче, на основании данных опроса, исследователь не может делать **никаких** выводов в отношении групп, численность которых в выборочной совокупности не позволяет делать **статистически обоснованных** выводов.

Что же должен делать исследователь, если малочисленность группы не позволяет приходиться к статистически обоснованным выводам?

Первое, самое простое, решение — убрать из результатов двумерных таблиц, подготавливаемых к широкому обнародованию, данные по группам, численность которых не позволяет делать обоснованных выводов. В противном случае, читатель, видя лишь представленные процентные распределения, будет введен в заблуждение. Конечно, при таком решении, часть информации «к размышлению» теряется. Но это лучше, чем формирование у неискушенного читателя искаженного представления об особенностях общественного сознания.

Второй путь, который может избрать исследователь, это допустимое укрупнение анализируемых групп. Например, возрастную группу 18-20 лет можно укрупнить (18-25 лет, или 18-30 лет и т.д.), увеличивая тем самым ее численность. В приведенной выше таблице 1 группу «высокообеспеченных» можно объединить с группой «среднего достатка», и в сравнительном анализе сопоставлять ответы «малоимущих» с людьми «среднего достатка и выше». При этом в тексте анализа следует отметить такое укрупнение, уточнив, сколько именно процентов от общей численности выборки составляет присоединенная группа. Разумеется, далеко не все малочисленные группы правомерно объединять с другими категориями. В некоторых случаях это достаточно очевидно: наверняка, ни один исследователь при сравнительном анализе различных профессиональных категорий не будет объединять в одну группу «военнослужащих», «студентов» и «предпринимателей». Несколько сложнее решать проблему укрупнения групп, полученных по условно порядковым шкалам. Например, градация ответов на вопрос анкеты, касающийся уровня образования, является случаем условно порядковой шкалы; и исследователь часто затрудняется с решением проблемы, с какой группой можно объединить категорию лиц, имеющих неоконченное высшее образование (эта категория тоже, как правило, в репрезентативных региональных выборках слишком малочисленна для многомерного анализа). В таких случаях укрупнение групп — самостоятельная творческая проблема, требующая обоснованного решения, которое не следует отдавать на откуп респондентам: произвести укрупнение в анкете (дать не все градации), а там — «пусть респондент сам решает, к какой группе себя отнести».

Существует и третий путь решения проблемы возможностей анализа малочисленных групп — подготовка «выпуклой» выборки (см. лекцию 4, раздел «Объем выборочной совокупности»).

3. Статистический анализ связи между двумя переменными

Стадии анализа

Построение двумерной таблицы и аналитическая интерпретация полученных данных — это лишь первая стадия анализа. Если основная задача исследователя ограничивается представлением результатов опроса общественного мнения как населения в целом, так и основных социально-демографических групп, то, в принципе, анализ может этим и ограничиться. Но если основной целью исследования является поиск взаимосвязи между социальными явлениями, объяснение причин полученных данных, то внимание исследователя переключается на изучение взаимосвязей между признаками. Поэтому основной задачей двумерного (как и любого многомерного) анализа на уровне объяснения

является оценка величины и направленности связей между признаками, как статистической основы причинного объяснения полученных результатов.

Как следует проводить статистическую оценку результатов двумерного анализа?

Общая схема анализа связей между признаками позволяет сориентироваться в назначении наиболее распространенных статистических коэффициентов связи. Как уже отмечалось, для социолога-аналитика в настоящее время не так важно разбираться в математической основе формул расчета, как в понимании содержательного смысла и, следовательно, аналитического назначения и правил использования различных статистических коэффициентов.

Общая технологическая схема предполагает: 1) разграничение трех стадий анализа, 2) определение, какая из переменных будет выступать в анализе в качестве независимой, а какая — зависимой, 3) установление типа шкалы каждой из переменных.

Разграничивают три стадии анализа:

- 1) установление связи между признаками;
- 2) измерение связи между признаками;
- 3) объяснение связи между признаками.

Для решения каждой из поставленных задач используются свои статистические показатели, соответствующие типу шкалы.

Установление связи между признаками

Те исследователи, которым приходилось интерпретировать результаты, представленные в двумерных таблицах, знают: чем больше вариантов ответов на вопросы, выступающие в таблице в качестве независимых и зависимых переменных, тем труднее проследить тенденции влияния одного признака на другой — обнаружить закономерность в изменении частот зависимой переменной с изменением характеристики независимой переменной. В том случае, когда деление на группы по независимой переменной — дихотомическое (например, по полу или по самооценке материального статуса: «малообеспеченные» и «среднеобеспеченные»), и деление по зависимой переменной — также дихотомическое («доверяют президенту» — «не доверяют президенту», «удовлетворены жизнью» — «не удовлетворены жизнью»), то различия между группами сравнивать достаточно легко; можно даже дать количественную оценку — насколько больше доверяющих президенту в одной группе по сравнению с другой. Но чем в таблице больше сравниваемых между собой групп, и чем больше вариантов ответа, по которым сравниваются эти группы, тем сложнее визуально установить закономерность изменения в характере ответов.

Существует статистический критерий, позволяющий получить ответ на вопрос: есть ли связь между двумя признаками, представленными в таблице. Этот критерий называется «хи квадрат».

Установление величины этого коэффициента и его статистической значимости позволяет ответить на вопрос о наличии связи между исследуемыми признаками. Эти вычисления обычно входят в компьютерные программы обработки данных (SPSS, OCA). Если значение этого коэффициента статистически значимо, по крайней мере, на уровне 5%, то исследователь может сделать вывод о том, что между анализируемыми признаками есть определенная связь (чем выше значение данного коэффициента, тем более тесная связь между этими двумя признаками). Если значение хи-квадрата статистически не значимо на уровне 5%, то исследователь не может сделать определенных выводов о наличии связи между этими двумя параметрами. Расчетная формула данного коэффициента включает наполненность каждой из полученных в таблице подгрупп. Иногда исследователь «забывает», что, производя группировку всей совокупности, он фактически из одной достаточно большой совокупности получает много маленьких. Например, если опрошено 1000 человек, и исследователь хочет определить, как влияет на оценку того или иного социального явления род занятий, то он разбивает всю совокупность на группы по роду занятий. Допустим таких групп 10. Даже если они наполнены равномерно, то каждая подгруппа составляет 100 человек. Эти сто человек давая ответ на вопрос о степени доверия президенту по пятибалльной шкале, в свою очередь, разбиваются на подгруппы в зависимости от выбранного варианта ответа. Если они распределяются равномерно, то каждый вариант ответа отметят всего 20 человек (если какая-то социальная группа получается большей численности, или какой-либо вариант ответа отмечается большим числом респондентов этой группы, то, соответственно, численность других подгрупп уменьшается). Поэтому при расчете хи-квадрата исследователь должен осознавать два момента. Первый — чем меньше градации анализируемых переменных, и, соответственно, чем больше численность полученных в результате кросс-табуляции подгрупп (на языке статистики они называются «степенями свободы»), тем более вероятно, что данный коэффициент покажет величину и значимость связи между изучаемыми признаками, если такая связь есть в объективной реальности.

Второй момент, который следует отметить, может быть полезен для тех социологов, которые хотели бы разобраться в содержательном значении данного коэффициента. Хи-квадрат рассчитывается, исходя из следующего предположения: если изменение одного признака не приводит к изменению другого, то все ответы распределяются равномерно (так, как мы описали в вышеприведенном примере, — в каждой подгруппе будет по 20 человек). Чем больше влияние одного признака на другой, тем больше будет число и величина отклонений от такого линейного распределения, полученного в результате кросс-табуляции. Усредненное значение таких отклонений и лежит в основе расчета данного коэффициента.

Таким образом, хи-квадрат позволяет получить ответ на вопрос, есть ли связь между признаками.

Здесь следует подчеркнуть, что установленное наличие связи еще не говорит о ее направленности и, тем более, о причинности. Например, один исследователь выдвигает гипотезу о том, что удовлетворенность работой влияет на удовлетворенность жизнью; а другой исследователь считает, что общая удовлетворенность жизнью влияет на удовлетворенность человека своей работой. При проверке каждой из гипотез, коэффициент связи будет одинаков, так как он показывает **только взаимосвязанность признаков**.

Измерение связи между переменными

Если связь установлена, то исследователь может переходить к следующей стадии анализа — измерению связи.

Выбор статистических процедур измерения связи зависит от типа шкалы измерения признака. В зависимости от того, является шкала дискретной или непрерывной, исследователь должен выбирать те или иные статистические процедуры оценки связи. Принципы определения типа шкалы описаны в предыдущих разделах. Мы не будем детально останавливаться на математических и содержательных особенностях данных коэффициентов; читатель при желании всегда может подробно изучить соответствующие разделы в литературе по математической статистике. Задача данного учебного пособия — познакомить читателя с основными правилами использования наиболее распространенных коэффициентов связи при анализе первичного социологического материала.

Компьютерные программы статистической обработки (SPSS, OCA) позволяют автоматически рассчитать многие из коэффициентов, измеряющих связь между признаками. Однако, получая автоматически рассчитанные коэффициенты связи, исследователь не должен забывать об адекватности данного коэффициента типу шкалы (компьютер ведь, если его «попросить», может подсчитать и среднее арифметическое между мужчинами и женщинами, и на этой «основе» проводить дальнейший количественный анализ).

Поэтому прежде всего необходимо учитывать, что наиболее популярный среди начинающих социологов коэффициент корреляции Пирсона-Брава (r), который свидетельствует о наличии той или иной степени линейной зависимости между двумя переменными, может иметь смысл только в том случае, если обе переменные измеряются с помощью непрерывных шкал, а для шкал дискретных применяются другие коэффициенты: для ранговых порядковых — коэффициенты ранговой корреляции Спирмена и Кэндалла, а для номинальных шкал — коэффициенты ассоциации и контингенции (для дихотомических шкал) и коэффициенты сопряженности (Пирсона и др. для шкал большей размерности).

Все перечисленные выше коэффициенты позволяют определять только величину связи, но не ее направленность. Для измерения направленности (меры возможного влияния одной переменной на другую) при анализе дискретных шкал используют лямбда-коэффициент Гудмана, а при наличии двух непрерывных шкал о направленности влияния можно судить по уравнению линейной регрессии.

Но используя коэффициенты или уравнения, определяющие направленность связи, исследователь должен помнить, что соответствующие значения показывают только некоторую вероятность направленной связи, но не позволяют делать окончательные выводы о причинной зависимости переменных. Что является причиной, а что — следствием — решает исключительно сам исследователь, сначала формулируя гипотезу, а затем, интерпретируя результаты анализа.

Объяснение связи между переменными

Идеальной целью любого анализа является объяснение полученных в результате опроса фактов. Установление причинной зависимости между различными социальными и

социально-психологическими явлениями открывает возможности социального прогнозирования и управления изучаемыми процессами.

Статистика, как уже отмечалось, сама по себе объяснить ничего не может; она является лишь инструментом, позволяющим в определенной мере перевести в количественные параметры все многообразие социальных явлений и связей между ними; определить правила их сопоставления, чтобы обоснованно подтвердить (или опровергнуть) содержательные послышки и выводы аналитика.

Адекватность объяснения полученным результатам определяется в основном теоретическим уровнем подготовленности автора анализа (его знанием предмета и следованием законам логики). В какой мере статистические процедуры могут помочь в поиске каузальных (причинных) зависимостей? Мы не будем останавливаться на последних достижениях в области статистики, направленных на поиск процедур, позволяющих обеспечить каузальный анализ. Чтобы их освоить, необходимо в какой-то мере пройти весь тот путь, который проходила математическая статистика в своем развитии. В данном учебном пособии, предназначенном в основном для тех исследователей, которые только начали осознавать необходимость повышения своего уровня квалификации в области анализа, мы остановимся на тех возможностях, освоить которые позволяет элементарная логика и изложенные выше правила анализа двумерных распределений.

Трехмерный анализ: введение контрольной переменной

Приступая к логическому объяснению установленных взаимосвязей, следует учитывать, что обнаруженная связь может не только не являться причинно-следственной, но, более того, может ввести исследователя в заблуждение. Например, в одном из исследований в процессе двумерного анализа была установлена значимая связь между политическими ориентациями людей и уровнем их тревожности: сторонники социалистической и, особенно, коммунистической ориентации значительно отличались более высоким уровнем тревожности по сравнению с людьми, ориентирующимися на демократическое преобразование общества. Если полученные данные начинать интерпретировать, не учитывая того обстоятельства, что, возможно, существует третий фактор, который, одновременно влияет и на политические ориентации, и на уровень тревожности, то только из одной этой двумерной таблицы можно было бы сделать далеко идущие выводы. Однако, предположение о том, что такой фактор существует, позволило установить, что в данном случае основным фактором того и другого показателей (политических ориентаций и уровня тревожности) является возраст (со всеми социальными и психологическими особенностями, сопутствующими этой переменной): у молодых сторонников социализма уровень тревожности практически не отличался от уровня тревожности лиц, придерживающихся других политических взглядов. Подобная картина наблюдалась и у пожилых людей с разными политическими ориентациями. Основная же причина обнаружения более высокого уровня тревожности в группе сторонников «левых сил» состояла в том, что в этой группе доля пожилых людей была существенно выше, чем в группе приверженцев идеи демократического преобразования общества.

Введение в двумерный анализ **контрольной переменной** позволяет избежать ошибок объяснения установленных в результате двумерного анализа связей. В качестве контрольных переменных следует в первую очередь использовать те демографические

характеристики, которые были заложены в основу подготовки репрезентативной выборки. Это объясняется тем, что когда исследователь обосновывает принципы отбора, он прежде всего заботится о том, чтобы выборочная совокупность репрезентировала генеральную по тем параметрам, которые в наибольшей степени могут влиять на показатели изучаемого явления. Если изучается связь между исследуемым показателем и различными социально-демографическими факторами, и установлено, что в той или иной степени каждый из них влияет на исследуемый показатель, то иерархизируя их по величине связи, следует убедиться, что каждый из этих факторов оказывает влияние сам по себе, а не только опосредуется другими.

Введение контрольной переменной приводит к построению трехмерных таблиц.

Эти таблицы можно анализировать визуально, а можно и статистически. Визуальный просмотр таблицы позволяет «увидеть», действительно ли зависимая переменная изменяется с изменением независимой, если группы уравниваются по третьему параметру.

В качестве примера введения третьей переменной приведем фрагмент анализа факторов самочувствия (самооценки здоровья). Допустим исследователь построил две двумерные таблицы: Таблица А (в качестве независимой переменной — пол, зависимой — самочувствие); и Таблица Б (в качестве независимой переменной — возраст, зависимой — самочувствие).

Таблица А

Влияние пола на самооценку здоровья (самочувствие), %

Самочувствие	Пол	
	Мужчины	Женщины
Плохое	24	38
Среднее	48	47
Хорошее	28	15
Всего %	100	100
N	794	1007

Таблица Б

Влияние возраста на самооценку здоровья (самочувствие), %

Самочувствие	Возраст		
	18-30	31-55	56-80
Плохое	12	23	61
Среднее	46	57	35
Хорошее	42	20	4
Всего %	100	100	100
N	446	834	497

Визуальный обзор и сопоставление этих двух таблиц показывает, что на самочувствие влияют оба фактора: пол (среди женщин больше тех, кто чувствует себя плохо, и меньше тех, кто чувствует себя хорошо, по сравнению с мужчинами) и возраста (чем старше возрастная группа, тем больше доля «чувствующих себя плохо» и меньше — «чувствующих себя хорошо»). Данные таблицы также позволяют «увидеть», что возраст оказывает более сильное влияние на самочувствие, чем пол: величина зависимой переменной в большей степени изменяется с изменением независимой в таблице Б. Статистическая проверка по хи-квадрату подтверждает вывод, сделанный на основе визуального сравнения: в таблице А хи-квадрат равен 67,3, в таблице Б — 197,4 (оба коэффициента значимы на уровне 0.01).

Внимательный аналитик не может не поставить перед собой вопрос: «Действительно ли пол сам по себе влияет на самочувствие. Может быть, поскольку продолжительность жизни женщин выше, их средний возраст выше, чем у мужчин; и в таблице А влияние пола опосредуется возрастным фактором?». Поэтому, прежде чем делать окончательные выводы о влиянии пола на самочувствие, необходимо проверить, влияет ли пол на самочувствие независимо от возраста. Для проверки этой гипотезы строится трехмерная таблица, позволяющая нивелировать влияние возраста: самочувствие мужчин и женщин сравнивается в пределах каждой возрастной группы (см. таблицу С).

Таблица С

**Влияние пола на самочувствие
у различных возрастных контингентов, %**

Возраст	18-30		31-55		56-80	
Пол	Мужчины	Женщины	Мужчины	Женщины	Мужчины	Женщины
Самочувствие						
Плохое	9	15	17	29	53	65
Среднее	38	54	58	56	42	30
Хорошее	53	31	25	15	5	5
Всего %	100	100	100	100	100	100
N	212	234	394	439	175	320

Визуальный обзор полученной трехмерной таблицы позволяет сделать вывод о том, что фактор пола оказывает влияние на самочувствие, так как женщины ниже оценивают состояние своего здоровья в каждой из возрастных групп.

В тех случаях, когда анализ факторов самочувствия не является основной и единственной целью эмпирического исследования, социолог может ограничиться констатацией этого факта и ранжировать факторы. Если же исследование факторов и особенностей самооценки здоровья является основной целью исследования, то специальные процедуры позволяют количественно проанализировать особенности различий самочувствия у мужчин и женщин на различных возрастных этапах жизненного пути.

Одна из таких процедур называется «анализ фрагментов таблиц сопряженности». Эта процедура состоит в выделении отдельных фрагментов таблицы по принципу (2x2) и сопоставлении их друг с другом по значению хи-квадрат.

Существует две основные модификации этой процедуры: 1) можно сопоставлять значение каждого фрагмента (например, различия в самооценке здоровья между градациями

«хорошее» и «среднее» в группах «женщин до 30 лет» и «мужчин до 30 лет») с общим распределением. Другими словами, каждое частное распределение сопоставляется с общим. Это позволяет измерить вклад каждой из значимых статистических связей в общий результат, и в конечном итоге, проранжировав каждую группировку по значению «хи-квадрат», определить значение из каждого факторов на разных возрастных этапах.

Другая модификация сводится к взаимному сопоставлению различных фрагментов общей трехмерной таблицы сопряженности между собой. Например, «на глаз» видно, что в любой возрастной категории женщины ниже оценивают состояние своего здоровья, чем мужчины. Видно также, что хуже всего себя чувствуют женщины старше 55 лет; но без дополнительного статистического анализа нельзя сделать вывод о том, в какой из возрастных групп эти различия наиболее выражены: молодые ли женщины чувствуют себя хуже, чем мужчины, или женщины среднего возраста, или старшего. Разность процентов не всегда статистически достоверна, да и трудно интерпретировать эту разность сразу по трем градациям. Хи-квадрат позволяет оценить связи в целом, определить их статистическую значимость и сопоставить по количественному значению.

Что должен учитывать социолог, строящий трехмерные таблицы с включением контрольной переменной, позволяющие проверить правильность или ошибочность связи, установленной при двумерном анализе?

1. Для трехмерного анализа требуется выборка достаточно большого объема, поскольку, разбивая двумерное распределение на подгруппы по контрольной переменной, исследователь в процессе анализа уменьшает численность каждой из вновь полученных подгрупп в несколько раз (точнее, во столько раз, сколько градаций в контрольной переменной). Это обстоятельство существенно снижает возможности применения трехмерного анализа. Однако без применения этой процедуры — введения контрольной переменной — исследователь не может быть полностью уверен, что полученные в результате двумерного анализа связи являются причинными. Но увеличение объема выборки связано со значительным повышением трудовых и финансовых затрат. Если организационные и финансовые возможности не позволяют увеличить объем выборки (который, напомним, определяется в значительной степени планом анализа), необходимо исходить, в первую очередь, из основной цели исследования.

В тех случаях, когда основной целью исследования является опрос общественного мнения, выборка строится на основе принципа репрезентативности выборочной совокупности по отношению к генеральной. Результаты двумерного анализа, прошедшие статистическую экспертизу, могут быть обнародованы, но автор должен избегать категорических выводов о причинно-следственном характере обнаруженных взаимосвязей. В тексте анализа он должен чаще пользоваться словами и выражениями, подчеркивающими гипотетический характер суждений о направленности связей: «можно предположить», «возможно», «скорее всего» и т.п.

Если основной целью исследования является анализ взаимосвязей, причин, закономерностей (т.е. проводится именно социологическое исследование), то в данном случае следует отказаться от репрезентативной выборки, а использовать стратификационный подход, набирать равночисленные группы респондентов по основным параметрам (гипотетическим факторам), которые необходимо учитывать при анализе.

2. При переходе от двумерного к трехмерному анализу возникает проблема — по какому принципу отбирать признаки в качестве контрольной переменной. Совершенно очевидно,

что механически перебирать все имеющиеся в инструментарии показатели (или даже часть из них) — работа не только невероятно трудоемкая, но и малоэффективная. Главную роль в отборе признаков, вводимых в анализ в качестве контрольной переменной, играет общая компетентность исследователя в изучаемой проблеме (объем его знаний и общая способность к логическому анализу). Например, включая возраст в качестве контрольной переменной, исследователь мог исходить из знания о том, что продолжительность жизни женщин выше, чем у мужчин, а следовательно, в выборочной совокупности женщин удельный вес лиц старшего возраста выше, что и приводит к снижению самооценки здоровья в группе женщин в целом. Трехмерный анализ позволил установить, что пол влияет на самочувствие независимо от возраста. А при анализе уровня тревожности у лиц с разной политической ориентацией, наоборот, введение возраста в качестве контрольной переменной позволило опровергнуть гипотезу о связи между этими переменными.

Как уже отмечалось, исследователь сам решает, какие признаки следует ввести в качестве контрольной переменной в процессе причинного анализа. Но если обратиться к общим рекомендациям, то следует обратить внимание начинающих социологов на такие признаки, как пол, возраст и уровень образования. Нередко введение в анализ именно этих параметров в качестве контрольных позволяет уточнить взаимосвязи между полученными данными. Следует заметить, что достаточно распространенной ошибкой являются выводы, касающиеся влияния таких параметров, как тип поселения и национальность. Например, нередко в двумерных распределениях мнения (оценки, отношения) у украинцев и русских различаются, то же относится к различиям между городскими и сельскими жителями.

Трехмерный анализ в некоторых случаях подтверждает, что в основе различий лежит этнический или социо-культурный фактор; однако нередко обнаруживается, что причина различия в мнениях заключается не в национальности, а в уровне образования (в репрезентативной выборке уровень образования украинцев ниже, поскольку среди них больше жителей сел). Выравненные группы по уровню образования нивелируют различие в ответах у разных этнических групп. Поэтому в случаях анализа таких факторов, как национальность и тип поселения, прежде чем делать категорические выводы об этнических и социокультурных различиях, следует обратить внимание на контроль по основным независимым переменным.

Если при введении контрольного признака связь между двумя переменными нивелируется, то такой признак называется *объясняющей переменной*, поскольку установленная между двумя показателями связь объясняется тем, что оба эти признака являлись следствием воздействия другого фактора (в нашем примере — возраста). Этот фактор остается **латентным** (скрытым) от глаз исследователя, когда он рассматривает связи только между парами признаков, т.е. осуществляет двумерный анализ.

Поиск **латентных** переменных (основных факторов влияния) лежит в основе многих статистических процедур многомерного анализа, в частности, **латентно-структурного анализа**, направленного на построение объясняющих переменных.

Если целью социологического исследования является поиск причин и закономерностей изучаемых явлений, то возможность осуществления многомерного анализа планируется на этапе подготовки проекта и программы исследования: выдвигаются гипотезы относительно латентных факторов и строятся шкалы, которые позволят применить соответствующие статистические процедуры. Исследователь должен учитывать, что если латентно-структурный анализ он может проводить на основе дискретных шкал, то,

например, **факторный** анализ предполагает, в основном, использование непрерывных шкал.

Подводя итоги рассмотрения проблемы анализа результатов эмпирического социологического исследования, подчеркнем основные моменты, которые должен учитывать социолог, приступая к этому этапу работы.

4. Заключение: технологическая схема анализа социологических данных

1. При освещении результатов эмпирического социологического исследования анализ можно осуществлять на двух уровнях: описание установленных фактов и объяснение (поиск причин и закономерностей).
2. Выбор уровня должен определяться статистической компетентностью исследователя и типом исследования (опрос общественного мнения/социологическое исследование).
3. Анализ результатов опроса общественного мнения в основном ограничивается описанием данных с учетом результатов статистической экспертизы.
4. Социологическое исследование предполагает причинный анализ взаимосвязей изучаемого явления, приступая к которому необходимо учитывать следующие обстоятельства:
 - а) следует различать общий социологический анализ проблемы и статистический анализ данных, полученных в результате опроса;
 - б) общий социологический анализ включает построение теоретического конструкта: логическое обоснование взаимосвязей изучаемого явления;
 - в) каждая декларируемая в теоретической модели взаимосвязь должна быть подтверждена фактами; эти факты исследователь может привлекать из различных источников; если он не находит факты в доступных ему источниках информации, он должен их собрать сам — этой основной цели и служит эмпирическое исследование;
 - г) программа социологического исследования (построение инструментария и плана анализа) подчинено гипотезам, выдвинутым в результате теоретического осмысления предмета исследования;
 - д) статистический анализ является статистической проверкой выдвинутых гипотез (при отсутствии гипотез статистические манипуляции с первичными данными приводят к псевдоанализу проблемы);
 - е) на этапе планирования анализа исследователь в первую очередь обращает внимание на шкалы — тип шкалы должен соответствовать статистическим процедурам проверки гипотез.
5. Планируя уровень анализа, необходимо также учитывать особенности итогового документа, в котором будут представлены результаты исследования.

Лекция 10.

Итоговые документы

1. Виды итоговых документов

После завершения анализа данных полученные результаты оформляются в итоговые документы исследования. По форме и назначению различают три основных вида итоговых документов: 1) отчет, 2) научные публикации, 3) публикации в средствах массовой информации. Особенности каждого из этих документов обусловлены спецификой информационных запросов различных потребителей.

Отчет в первую очередь предназначен для заказчика эмпирического исследования, который заинтересован в получении детальной информации по изучаемому предмету и в практических рекомендациях, позволяющих корректировать свои действия в соответствующей социальной сфере. Научные публикации адресованы в основном специалистам, занимающимся изучением данной проблемы. Публикации в средствах массовой информации предназначены для широких слоев населения, которых волнуют наиболее актуальные проблемы общественной жизни.

Подготавливая итоговый документ, исследователь должен учитывать специфику потребителя и основные требования, предъявляемые к каждому из этих видов оформления результатов.

2. Отчет

Общие требования к отчету

Отчет является основным итоговым документом, включающим всю содержательную информацию, полученную в результате исследования.

Первой особенностью отчета является **полнота информации**. Если социолог проводил опрос по анкете, включающей несколько сотен содержательных вопросов, а в отчет включает информацию по нескольким характеристикам изучаемого явления, то, следовательно, значительную часть средств он израсходовал непродуктивно. Не следует забывать, что с каждым пунктом анкеты на плечи респондентов, интервьюеров, операторов и других «соучастников» сбора информации ложилась определенная дополнительная нагрузка, следовательно, она должна окупиться, по крайней мере, систематизацией соответствующей информации.

Второй особенностью отчета является **доступность**. Под доступностью подразумевается, что основные результаты исследования будут изложены таким образом, чтобы их достаточно легко мог понять человек, заинтересованный в соответствующей информации, но не являющийся специалистом в области эмпирической социологии. Учитывая, что полнота и доступность являются в значительной степени взаимопротиводействующими качествами, существуют определенные технологические особенности, позволяющие повысить доступность и полноту отчета.

Структура отчета

В первую очередь основное внимание следует сосредоточить на структуре (композиции) отчета. При подготовке анкеты социолог во многом руководствуется особенностями восприятия респондентов: определяя последовательность вопросов исследователь как бы втягивает в диалог человека, который будет отвечать на вопросы анкеты.

Последовательность вопросов анкеты определяется особенностями психологии рядового человека («человека с улицы»), которому в общем-то нет дела до конкретного социологического исследования, но ему не безразличны как основные социальные проблемы, так и личностные, касающиеся всех людей. Особенности психологии заказчика при чтении отчета определяются прежде всего тем, что у него сформирован информационный запрос к конкретному эмпирическому исследованию — он хочет получить ответы на интересующие его вопросы. Поэтому, работая над структурой отчета, исследователь должен в первую очередь предъявить основные (главные) результаты исследования, а затем подробнее остановиться на деталях, позволяющих глубже понять особенности изучаемого явления.

Оптимальной является следующая структура отчета:

- 1) «резюмирующая» часть;
- 2) данные одномерных распределений;
- 3) основная часть.

Резюмирующая часть представляет собой две-три страницы, на которых изложены наиболее интересные и ценные результаты проведенного исследования, выводы и рекомендации.

Текст этой части отчета выполняется в так называемом «газетном формате»: короткие параграфы (не более трех-четырех предложений в каждом) и достаточно лаконичные фразы. Каждый параграф желательно снабдить подзаголовком (ключевой фразой), отражающим суть информации.

После резюмирующей части читателю целесообразно представить таблицы одномерных распределений по всем вопросам анкеты (в том порядке и с теми полными формулировками вопросов и ответов, в которых они предъявлялись респондентам). Данные, представленные в таком виде, называются одномерками.

Эти две части («резюмирующую» и «одномерки») можно поставить в начале отчета; но опыт показывает, что целесообразно объединить их в отдельный «том», титульный лист которого содержит название исследования и ремарку — «Основные результаты».

Перед одномерками необходимо привести «паспорт» исследования. «Паспорт» включает следующие характеристики исследования:

- 1) наименование организации, которая проводила исследование;
- 2) наименование организации, которая заказывала опрос;
- 3) время (сроки) проведения опроса;

- 4) определение генеральной совокупности (указание, на какую именно часть населения можно распространять результаты исследования);
- 5) указание объема выборочной совокупности (сколько человек было опрошено);
- 6) допустимую ошибку выборки (при случайном отборе) или перечисление характеристик, по которым отбирались квоты (при направленном отборе).

Резюмирующая часть отчета позволяет заказчику легко найти, как ответы на свои основные вопросы («информационный запрос»), ради которых он заказывал эмпирическое исследование, так и познакомиться с наиболее важными результатами, которые были установлены в результате исследования, даже если они не предполагались информационным запросом.

«Одномерки», во-первых, позволяют читателю ознакомиться с анкетой, на вопросы которой отвечали респонденты (ее общей композицией, порядком следования вопросов, полными формулировками вопросов и ответов); во-вторых, получить представление обо всех основных данных. Когда заказчик будет читать основную аналитическую часть отчета, он всегда, в случае необходимости, может обратиться к «базовым» данным, представленным в «одномерках».

Композиция основной (содержательной) части отчета

Основная часть отчета включает две главы: методическую и содержательную. В методической главе описываются: 1) принципы и процедура подготовки выборки; 2) принципы подготовки анкеты: связь между исследуемыми показателями и вопросами анкеты, функциональное назначение вопросов, возможности интерпретации ответов.

Содержательная глава посвящена подробному изложению результатов опроса и анализа полученных данных. Содержательная глава включает многочисленные фрагменты, каждый из которых представляет собой освещение определенного результата.

Композицию каждого фрагмента лучше всего строить следующим образом: 1) основной тезис — вывод, полученный на основе результата проверки гипотезы, 2) таблица, иллюстрирующая связь между показателями, 3) дополнительные комментарии к таблице, позволяющие читателю обратить внимание на различные особенности, подтверждающие и дополняющие основной тезис-вывод. Лучше, если исследователи каждый из фрагментов снабжают соответствующим подзаголовком. Это помогает как самому автору отчета не отвлекаться от основной сути изложения, последовательно и целенаправленно освещая полученные результаты, так и читателю ориентироваться в структуре потока информации.

Текстовая часть фрагментов не должна быть большой по объему. Исследователь должен стараться писать простым лаконичным языком. Главным элементом содержательной части отчета являются аналитические таблицы. Фактически именно они содержат всю полноту информации (текст лишь позволяет обратить внимание на основной вывод, а также помогает заказчику убедиться, что он «правильно прочитал» и понял таблицу. Поэтому к оформлению таблиц социолог должен относиться особенно внимательно.

Таблицы, в которых приводятся средние значения признака, измеряемого по определенной шкале, будут иметь аналитический смысл лишь в том случае, когда наряду со средними приводятся и показатели вариации (стандартное отклонение). Если же приводится таблица с коэффициентами корреляции, то необходимо указывать их статистическую значимость, по крайней мере, на трех уровнях, традиционно используемых в анализе данных социологических исследований — 5%, 1% и 0.1%.

Анализ информации, выводы и рекомендации

Основной целью «заказного» социологического исследования являются выводы и рекомендации для заказчика на основе анализа полученной в исследовании социальной информации. Это не означает, что при подготовке отчета нельзя использовать обнаруженные результаты других исследований по данной проблеме или привлекать для иллюстраций информацию о предмете исследования, извлеченную из различного рода социальных наблюдений. Однако общим правилом разработки рекомендаций при подготовке отчета является их обоснование прежде всего на основе обобщений и выводов проведенного опроса. Если же ту или иную рекомендацию можно обосновать на основе другой социальной информации, то затраты на исследование, проведенное автором отчета, оказываются излишними.

Очень часто в отчетах содержатся рекомендации, основанные исключительно на здравом смысле или житейском опыте автора. В этом случае речь идет уже не только о напрасно потраченных средствах на исследование, но о серьезной угрозе для заказчика получить заведомо ошибочные рекомендации, поскольку и здравый смысл и индивидуальный опыт могут не соответствовать истинному положению дел в той или иной сфере социальных отношений.

К адекватным предмету исследования и запросам заказчика рекомендациям приводит следующая структура анализа: систематизированное представление и описание первичной информации — обнаружение социальных фактов и закономерностей — обобщение полученных результатов в общих выводах исследования — разработка практических рекомендаций на основе выводов и с учетом потребностей заказчика. В этой логической цепочке не должно быть места эмоциям, житейскому опыту и «партийной позиции» (политическим убеждениям) исследователя.

Эмоции автора могут существенно повлиять на восприятие аналитического текста даже в том случае, когда анализ ограничивается только описанием числовых данных. Сплошь и рядом приходится читать тексты отчетов, в которых автор восклицает: «всего лишь столько-то процентов опрошенных поддерживают такую позицию!». Тем самым автор навязывает читателю мнение о незначительности поддержки населением определенной позиции. При этом автор на свой собственный вкус определяет, какой процент избравших тот или иной вариант ответа является несущественным для внимания читателя. Чтобы эмоциональные восклицания типа «всего лишь» или «очень большая часть» имели право на существование в тексте отчета, необходим весьма кропотливый анализ с привлечением сопоставимых данных других исследований, свидетельствующих о том, что результат последнего опроса — действительно заслуживает эмоционального выделения. В противном случае автор может написать, к примеру, что «всего лишь 20% избирателей поддерживают оппозицию». Однако если учесть при этом, что в более ранних

исследованиях оппозицию поддерживали 10%, а партию власти в настоящее время — 8% (и к тому же, у оппозиции общий индекс доверия выше, чем у правящей партии), то «всего лишь» применительно к полученному результату является абсолютно неадекватной характеристикой.

Опора на индивидуальный опыт столкновения с социальной действительностью также является весьма популярным методом выработки практических рекомендаций. К примеру, автор, страдающий от нарушения графика работы городского транспорта, настоятельно рекомендует городским властям прежде всего улучшить работу транспорта, хотя запрос заказчика был направлен на выявление отношения населения города к работе правоохранительных органов. И хотя от такой рекомендации, казалось бы, всем только польза, в действительности, она наносит ущерб, подрывая доверие заказчика к социологической информации и результатам опросов общественного мнения.

Недопустимой является корректировка социологической информации в соответствии с политическими убеждениями автора или заказчика. Нередко в текстах отчетов приходится сталкиваться со стремлением автора (иногда даже не осознаваемым им самим) так подать данные исследования, чтобы независимо от истинного положения вещей высокопоставленный или «денежный» заказчик остался доволен результатами. Такие факты наблюдаются при сравнительной оценке рейтингов политических деятелей и предсказании результатов выборов. Причем желание угодить заказчику или лицу, находящемуся у власти, объективно им же и наносит ущерб, дезориентируя в оценке реального положения дел. Разумеется, чаще всего ошибки в предсказаниях результатов выборов являются следствием добросовестного заблуждения социологов, не сумевших правильно организовать исследование или проинтерпретировать его результаты.

Искажению выводов и рекомендаций исследования способствует и стремление автора к интерпретации данных в соответствии с собственными социально-политическими ориентациями. Так, например, сторонник радикальных экономических реформ «не замечает» данных, свидетельствующих о преобладании консервативных установок у тех или иных групп населения, а поборник социального равенства акцентирует внимание только на потребности большинства людей в государственном патернализме, игнорируя параллельно высказываемое большинством желание развивать экономику, основанную на частной предпринимательской инициативе.

Существенно снижает качество выводов и рекомендаций, завершающих отчет по заказным исследованиям, неумелая фантазия, присущая авторам, открывающим для себя в процессе овладения ремеслом социолога безграничные эвристические возможности анализа социальных явлений. Обнаружив, например, что женщины в большей мере, чем мужчины, доверяют определенному политическому деятелю, «фантазер» сразу же находит многочисленные объяснения, связанные с особенностями женской психологии или с незаурядными внешними данными политика и т.п., хотя в инструментарии исследования не было вопросов, способных подтвердить или опровергнуть такого рода выводы.

Автор отчета, формулирующий выводы и рекомендации на основе анализа результатов опроса общественного мнения должен избегать императивных суждений, имеющих безусловную директивную направленность. Речь идет о таких рекомендациях, в которых содержатся обязательные требования изменить те или иные социальные условия, принять конкретные политические, экономические или правовые решения. И хотя нередко именно этого от него ждет заказчик, социолог должен отдавать себе отчет в том, что такого рода рекомендации возможны только на основе тщательно проработанного социального эксперимента, который способен обнаружить возможные последствия тех или иных

социальных решений и действий. Все прочие методы социологического исследования, в том числе и опросы, позволяют лишь адекватно описать и объяснить социальную ситуацию, с тем чтобы на основе правильного понимания этой ситуации практические работники (политики, менеджеры, бизнесмены и т.п.) могли самостоятельно принимать решения. К примеру, обнаружив факт неудовлетворенности большинства работников размером заработной платы и даже показав связь уровня неудовлетворенности с отношением к работе, социолог не может, исходя из этого, рекомендовать администрации повысить заработную плату до такого размера, который удовлетворял бы большинство работников. Его задача заключается в том, чтобы в выводах привлечь внимание администрации к одному из возможных источников недостаточно эффективной работы персонала, по возможности определить величину ущерба (по тесноте связи неудовлетворенности с трудовой отдачей) и рекомендовать заказчику сопоставить величину данного ущерба с теми затратами, которые будут неизбежны при повышении заработной платы. Разумеется, социолог в выводах может представить и свой (гипотетический) прогноз развития ситуации в случае принятия тех или иных решений, но при этом он не должен принимать решения за заказчика, формулируя рекомендации в директивной форме.

Обобщая вышесказанное, основные требования к анализу информации, выводам и рекомендациям, представленным в отчете, можно сформулировать следующим образом.

- 1) Выводы и рекомендации должны быть основаны прежде всего на анализе информации, полученной в исследовании; при этом должны быть исключены суждения, опирающиеся исключительно на здравый смысл, житейский опыт, эмоциональные оценки и идеологические предпочтения.
- 2) При подготовке резюмирующей части отчета ни в коем случае нельзя исходить из того, будут ли выводы и рекомендации приятны или неприятны заказчику, поскольку его объективный интерес связан исключительно с соответствием социологической информации реальному положению вещей.
- 3) Рекомендации не должны быть представлены в директивной форме как однозначные решения, требующие от заказчика обязательных действий. Рекомендации социолога, если в основу их не положены результаты тщательно проработанного социального эксперимента, могут выполнять объяснительную и (отчасти) прогностическую функции, оставляя сферу практических социальных решений практикам (политикам, администраторам и т.п.), что, собственно, и является их профессиональной деятельностью.

3. Научные публикации

Большинство содержательных требований, предъявляемых к работе над социологическим отчетом, распространяются и на подготовку научных публикаций: объективность анализа, соответствие выводов полученным в исследовании результатам, адекватное описание объекта и предмета исследования, грамотное оформление основного текста и иллюстративных материалов. Вместе с тем, научные публикации, в отличие от отчетов, с одной стороны, допускают больше возможностей для теоретических изысканий, для постановки вопросов, не имеющих строгой верификации в эмпирических данных, но открывающих новые перспективы исследований, но с другой стороны, к анализу

эмпирического материала в научных публикациях предъявляются более строгие требования, связанные со статистическим обоснованием полученных результатов.

В научных публикациях вполне допустимы теоретические рассуждения, не имеющие непосредственного выхода в общественную практику, в сферу социально-политических решений, однако в них нет места «механическому перечислению» эмпирических данных одномерных и двумерных распределений частот и процентов, полученных в результате обработки первичной социологической информации. К сожалению, именно такого рода описательный подход довольно часто встречается в публикациях весьма солидных отечественных журналов, в монографиях и сборниках социологических трудов. Во многом это связано и с тем, что научные публикации готовятся на основе текстов отчетов, одной из составляющих которых является представление больших объемов информации без углубленного анализа значительной части полученных в эмпирических исследованиях данных.

Чтобы «эффект описательности» не воспроизводился в научных публикациях, в мировой практике приняты четкие требования к представлению основного вида научного текста — статьи в научном журнале или сборнике научных трудов. Что касается монографий, брошюр и других видов объемных научных текстов, то требования к их оформлению во многом являются производными от требований к статьям, которые в большинстве случаев являются первоначальной основой подготовки отдельных разделов монографических изданий. Разумеется, существуют оригинальные и даже экстравагантные по форме подготовки социологические книги и статьи, но они, как правило, являются теоретическими работами, выполненными на стыке научного и публицистического жанров, тогда как основная часть научных публикаций оформляется по достаточно строгим требованиям. Так, научная статья должна содержать следующие обязательные подразделы:

- 1) краткое описание содержания (abstract);
- 2) введение, содержащее постановку проблемы и основные результаты ее изучения другими исследователями (introduction);
- 3) описание метода исследования (method) — выборки (sample, subjects)
 - процедуры (prosedure)
 - методики исследования и обработки данных (instruments);
- 4) изложение результатов (results, findings);
- 5) выводы и их обсуждение (conclusions and discussion);
- 6) примечания (notes) и ссылки на источники (references).

4. Публикации в средствах массовой информации

Социолог и СМИ

Теме взаимоотношений социологов и средств массовой информации (СМИ) в отечественной методической литературе уделяется недостаточное внимание, хотя именно

эта часть работы социологов имеет особое общественное значение, так как именно телевидение, радио и пресса превращают социологическую информацию в объект пристального внимания многомиллионной аудитории. Это накладывает особую ответственность на социологов, обнародующих данные исследований, затрагивающих актуальные общественные проблемы. Что касается роли социологов в тиражировании социологической информации посредством СМИ, то основной их заботой должно быть не только грамотное и доступное для непрофессиональной аудитории изложение материалов исследований, но и контроль за тем, чтобы в процессе подготовки материалов журналистами, в социологические данные не вносились искажения, а их интерпретация соответствовала реальному смыслу и общественному назначению.

Как показывает наш опыт общения с журналистами, наилучшей формой подачи материала является подготовка специальных пресс-релизов, содержащих информацию по одному из вопросов, представляющих интерес для массовых аудиторий. Требования к пресс-релизу: 1) краткость (объем обычно не должен превышать двух страниц, включающих описание исследования, представление полученных данных, выводы и комментарии), 2) популярность изложения (отсутствие узкоспециальной терминологии и статистических выкладок), 3) актуальность (связь результатов исследования с интересами массовой аудитории). Распространение пресс-релизов может осуществляться как в процессе непосредственной доставки их в редакции, так и посредством специально организованных для представителей СМИ пресс-конференций, где социологи могут ответить на вопросы, возникающие у журналистов в связи с полученной ими информацией. Но последний путь требует существенных организационных затрат, и не всегда удается добиться широкой посещаемости такого рода пресс-конференций.

Прекрасной трибуной для социологов являются тематические телепередачи, куда приглашаются эксперты по тем или иным социальным проблемам. При таком общении социолога и СМИ риск искажения социологической информации минимален со стороны работников средств массовой информации. Но далеко не всегда информация не соответствует действительности по вине журналистов и редакций, вольно обращающихся с цифрами и фактами, полученными от социологов. Немало проблем с качеством социологической информации в СМИ возникают по вине тех, кто эту информацию производит и отдает в прессу, на радио и телевидение. Это обстоятельство вынуждает ставить вопрос о специальной подготовке журналистов, использующих социологические данные, для контактов с различными фирмами и группами, производящими социологическую информацию.

СМИ и социологическая информация

В периоды, когда за своевременную социологическую информацию платят деньги (иногда немалые), как грибы после дождя растут «фирмы», «службы», «центры» по изучению общественного мнения. В этих социологических службах нередко работают люди, которые о социологии еще вчера знали только понаслышке, а сегодняшняя конъюнктура побудила их активно осваивать новый «промысел».

Средства массовой информации в последние годы пестрят данными опросов населения, ошибки и погрешности в проведении которых видны специалисту невооруженным глазом. «Святая простота» новоявленных «социологов» приводит к тому, что они предоставляют журналистам для дальнейшего широкого обнародования данные, не только позволяющие обнаружить погрешности и ошибки своей работы, но и рассчитать их величину.

Профессиональный долг и профессиональный интерес толкают журналистов на поиски информации, связанной с результатами опросов общественного мнения. Найдя ту или иную организацию (или представителя той или иной организации), располагающую результатами опроса населения, журналист получает желаемую информацию. Конечно, его иногда по-человечески волнует — насколько можно доверять полученным данным? Каким образом можно убедиться, что опрос, результаты которого он предполагает осветить, был проведен достаточно компетентно? Возможно ли вообще проконтролировать качество подобной информации?

К сожалению, при подготовке журналистов этой проблеме — как контролировать социологическую информацию, получаемую в результате опросов общественного мнения — до последнего времени не уделялось внимания. Единственное, чем мог быть вооружен журналист в подобных ситуациях, это личная добросовестность и здоровый скептицизм. Думаю, что многие журналисты уже имели возможность убедиться в том, что это очень слабое профессиональное оснащение в работе с результатами социологических опросов.

В таком же положении находятся и политики. С одной стороны, они нередко оказываются в ситуациях, когда остро испытывают потребность в информации о состоянии общественного мнения (чувствах, мыслях, намерениях различных слоев населения), готовы нести существенные расходы за получение подобной информации. С другой стороны, заказывая исследование, не уверены, что полученная информация будет достаточно качественной.

В странах Запада, где опыт освещения и использования результатов социологических опросов населения значительно богаче, издаются специальные брошюры и методические пособия, предназначенные для оказания помощи журналистам и политикам, обращающимся к результатам массовых опросов.

Анализ этой литературы и собственный опыт контактов с журналистами позволяет сформулировать ряд вопросов, получив квалифицированный ответ на которые журналист может быть уверен в том, что полученная им информация соответствует элементарным профессиональным требованиям.

Вопросы журналиста к социологу

К **обязательным вопросам**, которые журналист должен задать социологу, предоставляющему ему результаты опроса, относятся следующие.

1. Кто (какая организация) проводил опрос?
2. Когда проводился опрос? (Сроки сбора информации)
3. Кого репрезентирует (представляет) опрошенная группа населения? (Население Украины в целом? Отдельные области? Отдельные города или другие населенные пункты? Определенные слои или социальные группы населения?)
4. Сколько человек было опрошено? (Объем выборочной совокупности.)

5. Какова допустимая (граничная) ошибка случайной выборки? (Вопрос об ошибке репрезентативности).

Ответы на эти основные вопросы позволяют правильно интерпретировать данные опроса общественного мнения. Поэтому журналист должен не только сам получить на них ответы, но и предоставить их читателю (слушателю), сопроводив этой информацией публикуемые цифры — содержательные результаты опроса.

Из этих вопросов главным, пожалуй, является вопрос о том, кого репрезентирует опрошенная группа населения. Странно читать на страницах газет или слышать по радио или телевидению, например, следующее: «Проведен опрос населения; 70% опрошенных поддерживают курс на усиление экономических связей с Россией». Сразу же не только у политиков, но и у рядового внимательного слушателя возникает вопрос: «О ком идет речь? О населении Украины в целом? Или о каких-то регионах? Или опрос проводился только в Киеве?» и т.д.

Принципиальным является вопрос о допустимой (граничной) ошибке выборки. Очень важно, чтобы журналисты, как и другие пользователи социологической информацией понимали содержание и значение этой ошибки.

Например, если в каком-либо выборочном опросе 65% опрошенных ответили, что они поддерживают политику, направленную на приватизацию земли, то эта цифра (65%) при переносе ее на всю обследуемую группу не должна прочитываться буквально. Учитывая ошибку репрезентативности (допустим, 5%), читатель знает, что политику приватизации земли поддерживают 65% плюс-минус 5%, т.е. от 60% до 70% исследуемой группы населения. Эта ошибка вызвана не погрешностями, допущенными организаторами проведения опроса, а спецификой **выборочного** подхода когда опрашивается часть группы, а выводы делаются в отношении группы в целом. Это означает, что, если бы исследователи повторили свой опрос в то же самое время, в тех же условиях, по тем же правилам, но в выборку попали бы другие люди, то с определенной вероятностью можно говорить, что полученный ответ колебался бы в указанных пределах. Зафиксированная цифра без указания допустимой ошибки может характеризовать мнение только тех лиц, которые непосредственно были опрошены. Просто смешно читать, что «такую-то партию (или такого-то политического деятеля) поддерживает большее число населения — 13%, по сравнению с другой партией (другим политическим деятелем), которую поддерживают 11%. О различии в степени поддержки при подобных данных можно говорить лишь в том случае, если ошибка выборки не превышает 2%. Тем не менее такого типа информация довольно часто появляется в средствах массовой информации. Это свидетельствует о том, что не только журналисты, но и многие социологи, предоставляющие им информацию, естественные колебания, порождаемые принципом случайной выборки, интерпретируют как установленное различие в отношении населения. Хотя даже здравый смысл должен подсказывать, что если бы опрос был повторен еще раз (в то же время и по тем же правилам), то вряд ли можно было бы получить совершенно идентичные результаты.

Вопрос журналиста организаторам опроса об ошибке выборки, кроме получения необходимой информации, может иметь еще одно назначение. Он является своего рода тестом на квалификацию исследователей. Социологические организации, работники которых не рассчитывают ошибку репрезентативности, являются непрофессиональными, доверять результатам опросов которых не следует. В стандартных социологических опросах в качестве максимальной допустимой ошибки принята величина в 5%. Для того чтобы допустимая ошибка не превышала этой цифры, необходимо опросить около 400 представителей обследуемой группы населения.

Это относится к мнению группы в целом. Когда данные опросов дают по различным группам населения (например, мужчины и женщины, представители различных возрастных категорий и т.п.), то ошибка увеличивается. Чтобы ошибка не превышала 5%, необходимо опросить около 400 представителей каждой из анализируемых подгрупп.

Даже если журналист не располагает данными об ошибке выборки, в любом случае свою информацию он должен сопровождать сведениями о том, сколько человек было опрошено. Если он публикует предоставленную ему таблицу, в которой приводятся данные в процентах, он обязательно должен в таблицу включить цифру, характеризующую численность людей, по отношению к которым этот процент был вычислен.

Вопрос о том, сколько человек было опрошено (об объеме выборки) позволяет примерно предвидеть ошибку репрезентативности выборки: чем больше объем, тем меньше ошибка. Естественно, что это справедливо только в тех случаях, когда организаторы опроса правильно осуществляли отбор респондентов. Но журналист, к сожалению, никогда не сможет это оценить. Для соответствующей экспертизы нужны не только специальные знания, но и рабочие документы исследования (инструкции, отчеты интервьюеров и супервайзеров и т.п.). Такие документы могут быть предоставлены заказчику, если у него появляются сомнения в качестве выполненной работы, но журналист не вправе требовать от социолога рабочую документацию. Здесь представителям средств массовой информации остается полагаться на квалификацию исполнителей.

Поэтому очень важно каждый раз, приводя данные опроса, сопровождать освещаемые цифры названием фирмы, которая проводила опрос. Со временем у широкой общественности сформируется представление о качестве работы различных социологических служб, — так как обнародованные результаты опросов будут либо подтверждаться, либо опровергаться последующими событиями. На наш взгляд, журналист должен не только ссылаться на название организации, которая проводила опрос, но и указывать фамилию сотрудника фирмы, от которого он непосредственно получил информацию. Это корректно, с этической точки зрения, в том случае, когда информация качественная, адекватно проинтерпретирована. Это важно еще и потому, что в последние годы журналистам информацию нередко поставляют лица, которые, хотя и являются сотрудниками той или иной социологической организации, но в силу своей квалификации и положения не имеют права доступа к имеющейся информации. Такие сотрудники, не владея техникой обработки и анализа первичного социологического материала, получают из компьютера искаженную информацию и без ведома авторов исследования и руководителей соответствующей организации в таком виде предоставляют ее журналистам. Чтобы избежать подобных казусов, журналисту лучше ссылаться на имя сотрудника организации, от которого он получил результаты исследования.

В заключение данного раздела необходимо подчеркнуть, что овладение необходимым минимумом знаний в области проведения опросов, содержащимся в этом курсе лекций, позволит представителям средств массовой информации и другим возможным заказчикам социологической информации достаточно квалифицировано заказывать проведение опросов общественного мнения и отбор необходимой социологической информации и при этом компетентно контролировать качество исполнения заказанной социологам работы.

Заключение

Ознакомившись с представленным курсом лекций, читатель получил сведения не столько об академических требованиях к социологическому исследованию, представленных в традиционных учебниках и пособиях, но прежде всего о тех знаниях и практических рекомендациях, которые определяются современной общественной ситуацией и современными особенностями теоретико-методического обеспечения и организации инициативных и заказных социологических исследований. При этом в лекциях читатель встречал и те канонические требования, соблюдение которых является необходимым условием получения качественной социологической информации. Но этот материал был представлен в контексте реального опыта исследований в современных условиях, свидетельствующего, что даже профессиональные социологи, не говоря уже о многочисленных любителях, нередко игнорируют многие устоявшиеся методические требования, что приводит к искажению социальной информации и образа социальной действительности, создаваемого на основе данных социологических исследований.

Учитывая, что многие из тех советов и рекомендаций, которые появились на страницах этой книги, еще не приобрели статус тривиальных и общепринятых, я была бы признательна читателям, которые откликнутся на опубликованное пособие по проведению эмпирического социологического исследования. Для меня одинаковую ценность будут представлять как критические замечания опытных квалифицированных социологов, так и вопросы и замечания тех, кто только осваивает профессию социолога и будут использовать данное пособие при подготовке и проведении собственных исследований.

Литература

І. Справочная литература и учебные пособия

Батыгин Г.С. Библиография социологической литературы // Социологический журнал. — 1994. — N 4. — С. 153-174.

Воронов Ю.П. Методы сбора информации в социологическом исследовании. — М.: Статистика, 1974. — 158 с.

Гречихин В.Г. Лекции по методике и технике социологических исследований. Учебное пособие. — М.: Изд-во Моск. ун-та, 1988.- 232 с.

Давидюк Г.П. Прикладная социология. — Минск: Вышэйшая школа, 1979. — 224 с.

Здравомыслов А.Г. Методология и процедура социологических исследований. — М.: Мысль, 1969. — 205 с.

Капитонов Э.А. Социология XX века. История и технология. — Ростов-на-Дону: Феникс, 1996. — 509 с.

Краткий словарь по социологии. — М.: Политиздат, 1989.- 479 с.

Лекции по методике конкретных социальных исследований. — М.: Изд-во Моск. ун-та, 1972. — 202 с.

Маккинни Д. Методология, процедуры и техника социологии //Беккер Г., Босков А. Современная социологическая теория в ее преемственности и изменении. — М : Изд-во иностр. лит., 1961. — С. 218-271.

- Методы социальной психологии.* — Л.: Из-во Ленингр. ун-та, 1977. — 175 с.
- Михайлов С.* Эмпирическое социологическое исследование. — М.: Прогресс, 1975. — 383 с.
- Ноэль Э.* Массовые опросы. Введение в методику демоскопии. — М.: Прогресс, 1978. — 380 с.
- Основы прикладной социологии. Учебник для вузов.* — М.: Центр социального прогнозирования и маркетинга, 1995. — Т.1 — 200 с.
- Процесс социального исследования.* — М.: Прогресс, 1975.- 576 с.
- Пэнто Р., Гравитц М.* Методы социальных наук. — М.: Прогресс, 1972. — 607 с.
- Попова І.М.* Соціологія. Пропедевтичний курс. — К.: Тандем, 1996. — 271 с.
- Рабочая книга социолога.* — М.: Наука, 1983. — 477 с.
- Рукавишников В.О., Паниотто В.И., Чурилов Н.Н.* Опросы населения. — М.: Финансы и статистика, 1984. — 208 с.
- Словарь прикладной социологии.* — Минск: Изд-во Минского ун-та, 1984.- 317 с.
- Социологический справочник.* — Киев: Политиздат Украины, 1990. 382 с.
- Социология.* — М.: Мысль, 1990. — 447 с.
- Ядов В.А.* Социологическое исследование: методология, программы, методы. — М.: Наука, 1987. — 248 с.
- Якуба О.О.* Соціологія. Навчальний посібник для студентів. Харків: Константа, 1996. — 191 с.
- Alreck P.L., Settle R.B.* The Survey Research Handbook. Homewood: Irwin, 1985. — 429 p.
- Jary D., Jary J.* The Harper Collins Dictionary of Sociology. — New York: Harper Perennial, 1991. — 602 p.
- Lazarsfeld P. et al. (eds).* The Uses of Sociology. — New York: Basic Books, 1967. — 913 p.
- Miller D.C.* Handbook of Research Design and Social Measurement. — Newbury Park, London, New Delhi: Sage Publ., 1991. — 704 p.

**II. Литература по отдельным аспектам
технологии социологического
исследования**

Лекция 1.

- Артемов В.А.* Структура исследования. Программа (проект) исследования //Бюджет времени (вопросы изучения и использования). — Новосибирск : Наука, 1977. — С.102-113.
- Бабин Б.А., Бестужев-Лада И.В.* Предпрогнозная ориентация. Программа исследования //Прогнозирование социальных потребностей молодежи. — М.: Наука, 1978. — С.7-44.
- Батыгин Г.С., Малинкин А.Н., Цыденкова А.Л.* Преобразование концептуальной модели в систему показателей //Социальные показатели образа жизни советского общества. — М.: Наука, 1980. — С. 65-101.
- Буданцев Ю.П.* Программа и этапы эмпирического социологического исследования //Социология. Учебное пособие. — М.: Знание, 1995. — С. 218-286.
- Головаха Е.И.* Обоснование гипотез и показателей //Он же. Жизненная перспектива и профессиональное самоопределение молодежи. К.: Наукова думка, 1988. — С. 43-60.
- Горшков М.К., Шереги Ф.Э.* Примерная программа социологического исследования производственной дисциплины на предприятии //Как провести социологическое исследование. — М.: Политиздат, 1985. С.142-150.
- Горшков М.К.* Общественное мнение, история и современность. -М.: Политиздат, 1988. — 383 с.
- Грушин Б.А.* Мнения о мире и мир мнений. — М.: Политиздат, 1967. — 400 с.
- Докторов Б.З.* Метрологическая карта исследования общественного мнения //Социолог. исслед. — 1984. — N 1. — С.115-123.
- Заславская Т.И.* Социологический мониторинг экономических и социальных перемен в России //Экономические и социальные перемены: мониторинг общественного мнения. — 1993. — N 1. — С. 3-8.
- Здравомыслов А.Г., Ядов В.А.* Методологические предпосылки, методы и техника исследования (программа исследования) // Человек и его работа. Социологическое исследование. — М.: Мысль, 1967. — С. 12-97.
- Козлов В.А., Суслов Ю.А.* Программа конкретно-социологического исследования //Конкретно-социологические исследования в области права. Учебное пособие — Л.: Изд-во Ленингр. ун-та, 1981. — С.45-68.
- Массовая информация в советском промышленном городе. Опыт комплексного социологического исследования.* — М.: Политиздат, 1980. 446 с.
- Рейзема Я.В.* Информационный анализ социальных процессов. — М.: Наука, 1982. — 198 с.
- Социальные исследования. Теория и методы.* Вып. 5. — М.: Наука, 1970. — 308 с.

Adorno T., Frenkel-Brunswik E., Levinson D., Sandford R. The Authoritarian Personality. — New-York: W.W.Norton&Company, 1993. — 987 p.

Maas H.S., Kuypers J.A. From Thirty to Seventy: A Forty -Year Longitudinal Study of Adult Life Styles and Personality. Washington: Jossey — Baas Publ., 1977. — 328 p.

Лекция 2

Аванесов В.С. Тесты в социологическом исследовании.- М.: Наука, 1982.- 200 с.

Анастаси А. Психологическое тестирование. — В 2-х кн. — М.: Педагогика, 1982.- Кн. 1. — 318 с.

Бадалян Л.Г., Криворотов В.Ф. Объективность и достоверность данных социологического исследования //Методы социологического исследования. 3-я Всесоюзная конференция. Выпуск IV. — М.: Ин-т социологии АН СССР, 1989. — С. 53-56.

Батыгин Г.С. Опыт построения типологии социальных показателей образа жизни // Социолог. исслед. — 1977. — N2. — С.131-139.

Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психологической диагностике. — К.: Наук. думка, 1989. — 199 с.

Бутенко И.А. Анкетный опрос как общение социолога с респондентами. — М.: Высшая школа, 1989. — 176 с.

Бутенко И.А. Трудный вопрос //Социолог. исслед. — 1989. — N 4. — С.112-117.

Волович В.И. Надежность информации в социологическом исследовании. Проблемы методологии и методики. — К.: Наук. думка, 1974. 135 с.

Гильяшева И.Н. Вопросы как метод исследования личности // Методы психологической диагностики и коррекции в клинике. — Л.: Медицина, 1983. — С. 62-81.

Голофаст В.Б. Континуум опросных методов //Человек социалистического общества и процессы массовой коммуникации. — Л.: Наука, 1981. — С.105-111.

Димитров Д. Проблемы достоверности социологической информации //Социолог. исслед. — 1979. — N 1. — С. 186-194.

Дубин Б.В. Затрудняющиеся с ответом и социокультурная стратификация ответов //Экономические и социальные перемены: мониторинг общественного мнения. — 1995. — N 3. — С. 30-34.

Жабский М.И. Возможности, границы и техника опроса // Социолог. исслед. — 1984. — N 3. — С. 164-172.

Жабский М.И. Принципы стандартизированного интервью //Социолог. исслед. — 1985. — N 3. — С. 131-136.

Клайн П. Справочное руководство по конструированию тестов. К.: ПАН Лтд, 1994. — 283 с.

Косолапов М.С. Методический эксперимент как метод изучения познавательных возможностей различных эмпирических индикаторов //Методы социологического исследования. 3-я Всесоюзная конференция. Выпуск У. — М.: Ин-т социологии АН СССР, 1989. — С. 26-29.

Маслова О.М. Роль сравнительного анализа при обосновании методики социологического опроса //Методологические и методические проблемы сравнительного анализа в социологических исследованиях. — М.: ИСИ АН СССР, 1982. — Кн. 1. — С.21-39.

Маслова О.М. Познавательные возможности открытых и закрытых вопросов //Социолог. исслед. — 1984. — N 2. — С.129-138.

Методы сбора информации в социологических исследованиях. — Книга 1. Социологический опрос. — М.: Наука, 1990. — С. 64-124.

Паниотто В.И. Качество социологической информации.- К.: Наук. думка, 1986. — 207 с.

Погосян Г.А. Метод интервью и достоверность социологической информации. — Ереван: Изд-во АН Арм. ССР, 1985. — 142 с.

Погосян Г.А. Форма вопроса и целевая установка исследователя //Социолог. исслед. — 1983. — N 3. — С. 163-167.

Саганенко Г.И. Надежность результатов социологического исследования. — Л.: Наука, 1983. — 189 с.

Саганенко Г.И. Структура эмпирического результата в социологии и проблема его надежности // Социология. 4М. — 1994. — NN 3-4. С.5-22.

Социальные исследования: построение и сравнение показателей. М.: Наука, 1978. — 319 с.

Федоров И.В. Причины пропуска ответов при анкетном опросе // Социолог. исслед. — 1982. — N 2. — С.185-189.

Шляпентох В.Э. Проблемы достоверности статистической информации в социологических исследованиях. — М.: Статистика, 1973. — 144 с.

Шуман Г., Прессер С. Открытый и закрытый вопрос //Социолог. исслед. — 1982. — N 3. — С.145-156.

Ядов В.А. Базовое интервью //Саморегуляция и прогнозирование социального поведения личности. — Л.: Наука, 1979. — С. 209-216.

Dillman Don A. Mail and Telephone Survey:The Total Design Method. — New york: Willey-Inter science, 1978. — 480 p.

Hodge R.W., Siegel P.M., Rossi P.H. Validity, invalidity and reliability //Borgatta E.F., Bohrnstedt M., (eds.) Sociological Metodology. — San Francisco: Jossey-Bass, 1970. — P. 104-129.

Лекции 3-4.

Гарипов Я.З., Рысков И.Н. Построение типологии национально-административных территорий с помощью машинной классификации // Социолог. исслед. — 1979. — №3. — С. 146-153.

Давыдов А.А., Крыштановский А.О. Ремонт выборки // Социолог. исслед. — 1989. — №5. — С.100-105.

Жабский М.И. Обоснование репрезентативности социологического исследования //Социолог. исслед. — 1983. — №2. — С. 148-158.

Жабський М.І. Репрезентативність у соціологічному дослідженні //Філософ. думка. — 1974. — №5. — С. 81-91.

Козеренко Е.В., Новиков С.Г., Шнейдер М.Г. Методы построения и реализации выборки //Экономические и социальные перемены: мониторинг общественного мнения. — 1993. — №1. — С. 9-10.

Козеренко Е.В., Новиков С.Г. Выборка мониторинга: апостериорный контроль //Экономические и социальные перемены: мониторинг общественного мнения. — 1993. — №5. — С. 8-10.

Кокрен У. Методы выборочного исследования. — М.: Статистика, 1976. — 440 с.

Корнев Ю.Ф., Паниотто В.И., Яковенко Ю.И. Система территориальных опросов населения страны на базе почтовой службы страны //Социолог. исслед. — 1986. — №4. — С.115-118.

Петренко Е.С., Ярошенко Т.М. Социально-демографические показатели в социологических исследованиях. — М.: Статистика, 1979. — 166 с.

Проектирование и организация выборочного социологического исследования. — М.: ИСИ АН СССР, 1977. — 167 с.

Территориальная выборка в социологических исследованиях. — М.: Наука, 1980. — 218 с.

Филиппова О.В., Шкаратан О.И. Приложение 1. Методика выборки //Рабочий и инженер. — М.: Мысль, 1985. — С. 213-227.

Чубарев В.Л., Хотеев В.О. Вибірка за методом квот //Філософ. думка. — 1985. — №4. — С. 76-84.

Чурилов Н.Н. Проектирование выборочного социологического исследования (некоторые методологические и методические проблемы). — К.: Наук. думка, 1986. — 183 с.

Чурилов Н.Н. Труднодоступные единицы наблюдения — источник систематических ошибок //Социолог. исслед. — 1986. — №1. — С. 116-125.

Шереги Ф.Э. Методические проблемы выборки и репрезентативности в социологической практике //Социолог. исслед. — 1977. N1. — С.112-122.

Лекция 5

Алексеев Б.К., Докторов Б.З., Фирсов Б.М. Изучение общественного мнения: вопросы организации исследований //Социолог. исслед. 1981. — N 1. — С. 78-85.

Андреев В.Г. Сравнительный анализ организационных принципов развития эмпирических социологических исследований в США //Методологические и методические проблемы сравнительного анализа социологических исследований. Кн. 2. — М.: ИСИ АН СССР, 1982. — С.187-216.

Белановский С.А. Свободное интервью как метод социологического исследования // Социология. 4М. — 1991. — N 2. — С.5-19.

Войнова В.Д., Чернакова Н.Е. Организационно-методические особенности группового опроса //Организационно-методические проблемы социологического опроса. — М.: ИСИ АН СССР, 1986. — С. 50-69.

Гудков Л.Д. Характеристики респондентов, отказывающихся от контактов с интервьюерами //Экономические и социальные перемены: мониторинг общественного мнения. — 1996. — N 4. — С. 45-50.

Докторов Б.З., Паніотто В.І. Поштове опитування у соціологічному дослідженні //Філософська думка. — 1982. — N 4. — С. 42-51.

Ермолаева Е.М. Сравнительный анализ методических и организационных особенностей телефонного интервью и других методов массового опроса (на материалах исследований в США) //Методологические и методические проблемы сравнительного анализа в социологических исследованиях. Кн. 2. — М.: ИСИ АН СССР, 1982. — С.145-161.

Куракина В. Фокусированные групповые интервью как метод сбора социологической информации //Философская и социологическая мысль. — 1993. — N 2. — С.151-156.

Лесохина Л.Н. Введение в профессию социолога-интервьюера. Методические рекомендации . — Л.: НИИ общ. образования взрослых, 1990. 32 с.

Нестеров Л.И. Опросы общественного мнения, проводимые Институтом Гэллапа //Экспресс-информация. Серия: Методология изучения общественного мнения. — Выпуск 6. — М.: НИИ статистики ГК СССР по статистике, 1990. — С.104-107.

Панина Н.В. Методический эксперимент в изучении общественного мнения // Проблемы изучения и использования общественного мнения в условиях демократизации советского общества. — Горький: Горьк. ун-т, 1990. — С.24-27.

Потехин А.В., Бетяр И.А. Общественное мнение США и создание ООН. — Киев: Наукова думка, 1993. — 127 с.

Романович Н.А., Доморацкая А.Л. Почтовый контроль как метод повышения качества информации в социологии // Социс. — 1994. — N 2. С.96-100.

Романович Н.А. Отказы респондентов в зависимости от восприятия ими опроса и его методов // Экономические и социальные перемены: мониторинг общественного мнения. — 1996. — N 4. — С. 42-45.

Стацевич Т.Л. Особенности работы интервьюеров при телефонных опросах // Социс. — 1992. — N 8. — С.72-77.

Тодчук Ю.Н., Турчанинова И.Б. Социологическая служба в регионе // Социс. — 1995. — N 1. — С.84-91.

Топюк Л.М., Полторак В.А. Стационарна система опитувань громадської думки в місті (З досвіду створення СОГД в Дніпропетровську) // Філософська думка. — 1988. — N 4. — С. 47-57.

Человек и его работа. Социологическое исследование. — М.: Мысль, 1967. — Приложение А. Документы, относящиеся к сбору первичной информации. — С.323-335.

Чурилов Н.Н. Формирование постоянно действующих опросных сетей (методологические и методические проблемы) // Методы социологических исследований. Тезисы докладов 3-ей Всесоюзной конференции. — М.: ИСИ АН СССР, 1989. — С. 20-25.

Яковенко Ю.И., Паниотто В.И. Почтовый опрос в социологическом исследовании. — К.: Наукова думка, 1988. — 139 с.

Ядов В.А., Саганенко Г.И. Методологические проблемы оценки надежности сдвигов при повторных исследованиях (по материалам сравнительного исследования отношения рабочих к труду 1962-1976 гг.) // Советская социология. Том II. — М.: Наука, 1982. — С.230-249.

Cantrell P.D. Opinion Polling and American Democratic Culture. // International J. of Politics, Culture, and Society. 1992. — Vol.5. — P. 405-438.

Greer B. First Days in the Field // Hammond P.E., ed. Sociologists at Work: The Craft of Social Research. — Garden City: Doubleday, 1967. — P.372-398.

Sudman S. Reducing the cost of surveys. — Chicago, 1967. 194 p.

Seidler J. On Using Informants: A Technique for Collecting Quantitative Data and Controlling Measurement Error in Organizational Analysis // American Sociological Review. — 1974. Vol. 39. — P. 816-831.

Лекция 6

Андреев В.Г., Чердиченко В.А. К вопросу о создании банка социологической информации // Социолог. исслед. — 1982. — N 1. — С.140-146.

Горшков М.К., Шереги Ф.Э. В содружестве с ЭВМ (обработка собранной информации) // Как провести социологическое исследование. — М.: Политиздат, 1985. — С.142-150.

Докторова Л.Д. Система основных процедур подготовки и анализа данных в комплексном многоступенчатом исследовании //Саморегуляция и прогнозирование социального поведения личности. — Л.: Наука, 1979. 237-247.

Информационное обеспечение в социологии: теоретические и практические аспекты. — М.: ИСИ АН СССР, 1987. — 178 с.

Крыштановский А.О. Банк данных мониторинга //Экономические и социальные перемены: мониторинг общественного мнения. — 1995. — N 8. — С. 27-28.

Малахов В.А. Группировка социологической информации //Социолог. исслед. — 1989. — N 2. — С. 98-102.

Маслова Л.Н., Суховский М.Л. Подготовка анкет к обработке //Бюджет времени (вопросы изучения и использования). — Новосибирск : Наука, 1977. — С. 266-272.

Методика и техника статистической обработки первичной социологической информации. — М.: Наука, 1968. — 327 с.

Проблемы переработки и автоматизированного поиска социологической информации. — М.: ИСИ АН СССР, 1977. — 180 с.

Саганенко Г.И. Социологическая информация. Статистическая оценка исходных данных социологического исследования. — Л.: Наука, 1979. — 142 с.

Терещенко О.В., Орехов Н.И. Контроль и корректировка больших массивов социологической информации //Экспресс-информация. Серия: Методология изучения общественного мнения. Вып. 6. — М.: НИИ Статистики Госкомстата СССР, 1990. — С.81-83.

Толстова Ю.Н. Обеспечение однородности исходных данных в процессе применения математических методов //Социолог. исслед. — 1986. — N 3. — С.149-154.

Шереги Ф.Э. Подготовка первичной социологической информации к обработке на ЭВМ //Основы прикладной социологии. Учебник для вузов. — М.: Центр социального прогнозирования и маркетинга, 1995. — Т.2 С. 6-13.

Лекции 7-9

Адамов С.Ю. Система анализа нечисловой информации «САНИ» // Социология. 4М. — 1991. — N 2. — С.86-104

Аптон Г. Анализ таблиц сопряженности. — М.: Финансы и статистика, 1982. — 140 с.

Бешелев С.Д., Гурвич Ф.Г. Экспертные оценки. — М.: Наука, 1973. — 159 с.

Біловіт К.Д., Паніотто В.І. Вимірювання у конкретно-соціологічному дослідженні //Філософ. думка. — 1976. — N 5. С.83-93.

- Благуш П.* Факторный анализ с обобщениями. — М.: Финансы и статистика, 1989. — 248 с.
- Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. М.: Наука, 1983. — 416 с.
- Гласс Дж., Стэнли Дж.* Статистические методы в педагогике и психологии. — М.: Прогресс, 1976. — 494 с.
- Гуревич К.М., Акимова М.К., Козлова В.Т.* Статистическая норма или психологический норматив? //Психологический журнал. — 1986. N3. — С.136-142.
- Гурьев В.И.* Основы социальной статистики: методы, система показателей, анализ. — М.: Финансы и статистика, 1991. — 177 с.
- Гуттман Л.* Основные компоненты шкального анализа //Математические методы в современной буржуазной социологии — М.: Прогресс, 1966. С.288-343.
- Давыдов А.А.* Анализ одномерных частотных распределений в социологии: эволюция подходов //Социс. — 1995. — N 5. — С.113-116.
- Дегтярев Г.П.* Факторный анализ в социологическом исследовании: вопросы интерпретации //Комплексный подход к анализу данных в социологии. — М.: ИС АН СССР, 1989. — С. 168-184.
- Докторов Б.З.* О надежности измерения в социологическом исследовании. — Л.: Наука, 1979. — 127 с.
- Дрейпер Н., Смит Г.* Прикладной регрессионный анализ. — М.: Статистика, 1973. — 392 с.
- Дюран Б., Оддел П.* Кластерный анализ. — М.: Статистика, 1977. 128 с.
- Дэвид Г.* Метод парных сравнений. — М.: Статистика, 1978. — 142 с.
- Елисеева И.И.* Статистические методы измерения связей. — Л.: Изд-во Ленингр. ун-та, 1982. — 134 с.
- Иберла К.* Факторный анализ. — М.: Статистика, 1980. — 398 с.
- Ильясов Ф.Н.* Экспериментальное обоснование количества делений шкалы //Социолог. исслед. — 1984. — N 4. — С.113-117.
- Кендэлл М.* Ранговые корреляции. — М.: Статистика, 1975. — 212 с.
- Кендэлл М., Стюарт А.* Статистические выводы и связи. — М.: Наука, 1973. — 899 с.
- Киселева И.П.* Информативно-целевой анализ текста свободного интервью //Социологический журнал. — 1994. — N 3. — С. 110-116.
- Клигер С.А., Косолапов М.С., Толстова Ю.Н.* Шкалирование при сборе и анализе социологической информации. — М.: Наука, 1978. — 112 с.

Лазарсфельд П. Логические и математические основания латентно-структурного анализа //Математические методы в современной буржуазной социологии. — М.: Прогресс, 1966. — С.344-401.

Лазарсфельд П. Измерение в социологии //Американская социология: перспективы, проблемы, методы. — М.: Прогресс, 1972. С.134-149.

Лакутин О.В., Толстова Ю.Н. Качественная и количественная информация в социологии //Социс. — 1992. — N 8. — С.72-77.

Малахов В.А. Специальные методы группировки //Социолог. исслед. — 1980. — N 3. — С.190-200.

Математические методы анализа и интерпретации социологических данных. — М.: Наука, 1989. — 171 с.

Математические методы в социальных науках. — М.: Прогресс, 1973. — 351 с.

Математические методы в социологическом исследовании. — М.: Наука, 1981. — 334 с.

Моин В.Б. Две стратегии измерения //Социолог. исслед. — 1989. N 6. — С. 112-119.

Мюллер П., Нойман П., Шторм Р. Таблицы по математической статистике. — М.: Финансы и статистика, 1982. — 278 с.

Осинов Г.В., Андреев Э.П. Методы измерения в социологии. — М.: Наука, 1977. — 182 с.

Панина Н.В. Методика измерения жизненной удовлетворенности //«Lifeline» и другие новые методы психолого-биографического исследования. — М.: Прогресс, 1993. — С.97-105.

Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. — К.: Наук. думка, 1982. — 272 с.

Рассел Д. Измерение одиночества // Лабиринты одиночества. — М.: Прогресс, 1989. — С. 192-226.

Ростегаева Н.И. О некоторых методических аспектах построения социологических индексов //Комплексный подход к анализу данных в социологии. — М.: ИС АН СССР, 1989. — С. 43-52.

Сиповская И.В. Выборочное статистическое наблюдение и некоторые вопросы его применения в социологических исследованиях. — Тбилиси: Изд-во Тбилис. ун-та, 1972. — 133 с.

Статистические методы анализа информации в социологических исследованиях. — М.: Наука, 1979.- 319 с.

Терещенко О.В. К вопросу об оптимальных размерах оценочных шкал //Методы социологического исследования . 3-я Всесоюзная конференция. Выпуск У. — М.: Ин-т социологии АН СССР, 1989. — С. 48-50.

Толстова Ю.Н. Логика математического анализа социологических данных. — М.: Наука, 1991. — 111 с.

Толстова Ю.Н. Существует ли проблема социологического измерения? // Социология. 4М. — 1995. — NN 4-5. — С.103-117.

Фелингер А.Ф. Статистические алгоритмы в социологических исследованиях. — Новосибирск: Наука, 1985. — 208 с.

Харман Г. Современный факторный анализ: Пер. с англ. — М.: Статистика, 1972. — 486 с.

Cronbach L. Essential of Psychological Testing.- New York: Harper, 1984.- 630 p.

Лекция 10

Анализ нечисловой информации в социологических исследованиях. М.: Наука, 1985. — 221 с.

Андреевков В.Г. Проблемы сравнительного анализа в современных социологических исследованиях //Методологические и методические проблемы сравнительного анализа в социологических исследованиях. Книга 1. — М.: ИСИ АН СССР, 1982. — С.5-20.

Батыгин Г.С. Обоснование научного вывода в прикладной социологии. — М.: Наука, 1986. — 271 с.

Интерпретация и анализ данных в социологических исследованиях. — М.: Наука, 1987. — 255 с.

Королько В. Основи паблік рілейшенз. Посібник. К.: Ін-т соціології НАНУ, 1997. — 334 с.

Паніна Н.В. (відп.ред.) Опитування громадської думки. Для журналістів і політиків. — К.: Демократичні ініціативи, 1995. — 48 с.

Практикум по прикладной социологии. — М.: Изд-во Моск. ун-та., 1987. — 261 с.

Типология и классификация в социологических исследованиях. -М.: Наука, 1982. — 295 с.

Тумаков С.В., Гаспарашвили А.Т. На опросных фронтах //Социолог. исслед. — 1993. — N 6. — С. 36-41.

Шереги Ф.Э. Интерпретация полученных данных //Основы прикладной социологии: Учебник для вузов. — М.: Центр социального прогнозирования и маркетинга, 1995. — Т.2 — С. 54-66.

Ядов В.А. Стратегия и методы качественного анализа данных // Социология. 4М. — 1991. — N1. — С.14-31.