Kapitola 3

Pozorování nepozorovatelného

> Od doby, co Američané vynalezli Richarda Nixona,
> neměl nikdy nikdo tolik trablů.
>
> Richard Condon, *Bandicoot*

Tohle bude krátká kapitola, ale obtíže, na něž tu narazíme, by stačily na celou knihu. V předchozí kapitole jsme hovořili o našem hlavním problému, o tom, že díky velikosti „přirozených systémů" v sociálních vědách není možné docílit úplného popisu a úplné analýzy reality. Redukce je naším nejdůležitějším, ne však jediným epistemologickým problémem. Tím druhým je problém **transformace informací**.

V tomhle typu nepříjemností nejsme sami. Naprostá většina věd nedělá závěry přímo z analýzy reality, ale na základě záznamu, popisu této reality. To platí i pro přírodní vědy. Představa vědce, který zatřepe zkumavkou, obsah zkumavky změní barvu a vydá dramatický kouř a vědec začne tančit radostí, že jeho teorie byla potvrzena, patří dnes spíše do světa televizní fikce. I zde jsou konečné závěry navrženy na základě popisu, na základě protokolů o mnoha experimentech či pozorováních. A popis nutně znamená transformaci, překlad reality do symbolů jazyka.

Cvičení 3.1

Některé vědy se obejdou bez tohoto kroku, bez transformace reality do jejího popisu. Zkuste jmenovat nějakou takovou šťastnou vědu a vysvětlete, proč nemusí analyzovat pouhý popis.

Problém je v tom, že transformace informací je v různých vědách a pro různé pozorované problémy různě obtížná a různě spolehlivá.

Určit pohlaví respondenta není obvykle obtížný problém. Ale mnohé důležité sociologické koncepty nejsou dostupné přímému pozorování, například koncept odcizení. Jiné jsou dostupné pozorování, a přece je sledujeme obvykle jenom nepřímo. Jistě by bylo možné zjistit

přímým pozorováním bydliště respondenta, počet dětí, povolání, vlastnictví auta, pohlaví atd., ale my se většinou na tyto věci respondenta jenom ptáme, prostě proto, že přímé sledování by bylo příliš časově náročné a nákladné.

3.1 A proč se neptat?

Většinou nám nezbude nic jiného než se zeptat. Jenom nesmíme zapomenout, že každé nepřímé měření zvyšuje riziko zkreslení, a to tím více, čím více je kroků mezi zkoumaným fenoménem a naším protokolem, záznamem o tomto fenoménu.

I v něčem zdánlivě tak jednoduchém, jako je proces odpovědi na otázku, je těch transformací požehnaně dost. Podívejme se silným zvětšovacím sklem na následující transformaci:

otázka → odpověď

Mnoho kroků mezi těmito dvěma body má pro nás dosti důležitý význam. Měli bychom je mít stále na mysli, když kupř. připravujeme otázky pro interview nebo (nedej bože) pro dotazník.

(1) Především, zkoumaná osoba musí naší otázce správně porozumět. To není vždycky tak jednoduché. „Správně porozumět" znamená, že respondent pochopí otázku přesně tak, jak to měl výzkumník na mysli, a dále to znamená, že všichni dotazovaní musí pochopit otázku naprosto stejným způsobem. Jde nám přece – alespoň v kvantitativním výzkumu – o srovnatelnost dat.

Zkreslení může vyvolat i představa pozitivní sankce. Respondent volí nepravdivou odpověď, aby se ukázal v nejlepším světle, nebo chce prostě tazatele potěšit. To dobře ilustruje naše pravdivá Pohádka 3.

Pohádka pro odrostlejší děti 3

O vzorném tazateli a vzorné zkoumané osobě

Byl jednou jeden rozsáhlý výzkum o využití volného času ve velkém průmyslovém městě. Bylo to velmi náročné pro respondenty. Žádali jsme je, aby po řadu dnů zaznamenávali v 15minutových intervalech, co dělali od probuzení do usnutí. Tazateli byli studenti naší fakulty. Mezi nimi Jirka. Na Jirku připadla také jedna velice stará paní. Když uslyšela, o co jde, posílala Jirku za sousedkou. „Ta je mladá, pořád je někde, chodí na schůze a tak…" Náš Jirka jí laskavě a moudře vysvětlil teorii konstrukce vzorku, význam sociologického výzkumu, a nakonec babička podlehla, a když Jirka po několika dnech přišel, aby odnesl záznam, byl moc překvapen. Babička byla v divadle, na výstavě, dokonce i na schůzích. Jirka ji moc pochválil a stará paní řekla: „Ale jsem moc ráda, že to mám za sebou. Já bych už víc toho chození po těch candrbálech nevydržela."

Tak to je vzácný případ toho, jak výzkum může zkreslit ne odpověď, ale chování zkoumané osoby. Všimněte si, že jsem známého sociologa průmyslu, dr. Jiřího Orta, vůbec nejmenoval.

(4) I verbalizace odpovědi, zejména v rozhovoru, může být spojena s některými problémy. Pro respondenta může být nepřijatelné **vyslovit** určitou nepříjemnou alternativu, použít neobvyklé slovo. Obvykle jsme však schopni těmto situacím předejít: i v řízeném rozhovoru můžeme u kritických otázek předložit dotazované osobě tištěný lístek se seznamem alternativních odpovědí a zeptat se, která alternativa nejlépe vystihuje její situaci. Pro respondenta může být příjemnější odpovědět třeba číslem 5 než vyslovit přímo, že jeho manželství je velice nešťastné a že již často uvažoval o rozvodu.

(5) Záznam odpovědi může také vyvolat zkreslení. U otevřených otázek a zejména v kvalitativním výzkumu proces záznamu odpovědí – a někdy potřebujeme velice přesný záznam – může narušit přirozený charakter interakce mezi tazatelem a zkoumanou osobou. V standardizovaném rozhovoru je hlavním zdrojem zkreslení seznam nabídnutých kategorií. Neúplný soubor alternativních kategorií a kategorie, které se překrývají, patří mezi naše

obvyklé hříchy. K tomu ještě přistupuje méně častý omyl, výběr chybné kategorie, ať již tazatelem, nebo respondentem.

Vidíme tedy, že i zdánlivě jednoduchý proces, jako je odpovídat na otázku, představuje celou řadu transformací, z nichž každá může být spojena s považlivým zkreslením. Co si pak počneme se zkoumáním konceptů, které nemají přímou empirickou reprezentaci? A v sociologii máme takových konceptů opravdu mnoho. Odcizení, prestiž povolání, třídní vědomí, autorita a stovky jiných konceptů, se kterými pracuje sociologická teorie, je téměř nemožné pozorovat přímo.

3.2 Jak nám respondent řekne to, co neví

Ve většině věd jsou pozorování a měření převážně nepřímá. Studovaný fenomén je reprezentován nějakým **indikátorem**. Kupř. tělesná teplota je indikována pozicí rtuti na stupnici teploměru, věk zkoumané osoby je reprezentován její odpovědí na příslušnou otázku. Čím větší je vzdálenost mezi tím, co chceme studovat, a indikátorem, který jsme schopni pozorovat, tím větší je riziko zkreslení. V sociologii je tato vzdálenost značná a většinou zahrnuje dlouhou řadu transformací.

Nalezení spolehlivého ukazatele je jednou z nejkritičtějších operací ve výzkumu. Dopustíme-li se omylu v této fázi, všechny následující operace budou neplatné. Měli bychom zde aplikovat rčení populární mezi lidmi kolem počítačů od dob, kdy počítače byly velké jako dům a byly jenom dva. Dnes se vlastně nepoužívá rčení, ale jen zkratka: GIGO. Znamená to „garbage in, garbage out". Volně přeloženo to znamená asi „svinstvo na vstupu dá svinstvo na výstupu". Výzkum můžeme pokazit kdykoliv, ale nezachycené omyly při konstrukci indikátoru jsou fatální.

Abychom se mohli produktivně bavit o problémech s tím spojených, musíme zavést dva nové pojmy: pojem validity a pojem reliability.

Cvičení 3.2

Je s následující otázkou všechno v pořádku?

– Jaký je nyní Váš příjem?

Zamyslete se nad tím na chvíli.

Častým hříchem při konstrukci otázek je, že mylně generalizujeme náš specifický okruh znalostí na celou populaci. Můžeme si být jisti, že opravdu každý ví, co je HZDS? Můžeme si být jisti, že opravdu každý používá určitý termín ve stejném smyslu jako my? Někdy můžeme být opravdu překvapeni:

Pohádka pro odrostlejší děti 2

Egri Bikaver aneb Jak to nevyšlo

Před mnoha a mnoha lety jsme dělali výzkum v jedné malé horské obci. Velice chudá obec to byla, s tak špatným JZD, že v něm nebylo ani co ukradnout. Byl to kvalitativní výzkum. Mimo jiné nás zajímal životní styl obyvatel. Ptali jsme se také: „Co nejčastěji pijete?"

Odpověď prvního družstevníka nás lehce překvapila. Odpověď druhého, třetího, desátého nás prostě šokovala. Všechny odpovědi byly stejné. Všichni dotazovaní odpověděli, že nejčastěji pijí Egri Bikaver. To byl pro nás šok: tohle víno, maďarská Býčí krev, bylo v té době pěkně drahé. Rozhodně nad poměry lidí v téhle obci.

Ovšem jsem si hned vykonstruoval teorii. Krásnou a chytrou, kompenzační. Velmi nízká životní úroveň deprimuje družstevníky. Pití je důležitou součástí místní kultury, a tak se lidé snaží udržet vysoký status alespoň v této oblasti, když všude jinde je jejich status zoufale nízký.

Tak jsem se šel hned pochlubit se svojí teorií osobě, která rozuměla obci nejlépe, místnímu panu faráři. Když se pan farář přestal smát – a smál se moc dlouho –, vysvětlil mi jednu zvláštnost místního slovníku: Egri Bikaver se tu říká hořčáku, v té době nejlacinějšímu, přišernému ovocnému vínu. Říká se tomu tak proto, že „když se toho napijete, bolí vás hlava tak, že řvete jako bejk". A to byl zase jednou konec mé geniální teorie.

(2) V dalším kroku se musí zkoumaná osoba rozhodnout, zda je vůbec ochotna odpovědět na naši otázku. Některé otázky mohou být pro respondenta nepříjemné, může se cítit jimi ohrožen, může chtít chránit své soukromí. Ne vždy jsme schopni takovou situaci předvídat.

(3) Pak se náš respondent musí rozhodnout, jak odpoví. Odpoví pravdivě? Zná vůbec pravdivou odpověď na naši otázku? Institut für Demoskopie zjistil, že hlavní příčinou nepravdivých odpovědí v Německé spolkové republice v 60. letech byl prostý fakt, že respondent **neznal** pravdivou odpověď (Noelle 1967). Neznalost respondenta je velice častým zdrojem zkreslení; zkoumaná osoba nerada přiznává, že nezná odpověď, že nerozumí některému výrazu v otázce, a tak stejně odpoví. Přirozeně taková odpověď je pro náš výzkum neplatná, ale my o tom nevíme.

Není příliš těžké empiricky dokázat, že riziko takového zkreslení je dost vysoké. Před několika lety moji studenti provedli následující test. Do krátkého dotazníku, zaměřeného na některé etické aspekty politického rozhodování, jsme zařadili vysloveně nesmyslnou otázku. Její český ekvivalent by zněl asi takto:

– Domníváte se, že koncept kufanditní pumprdentnosti je:

1. mravně akceptovatelný
2. akceptovatelný jen někdy
3. neakceptovatelný
4. nevím

Dotazník byl pak aplikován na malý a nereprezentativní vzorek posluchačů prvého ročníku a výsledky byly docela zajímavé. Jen menšina respondentů zvolila jedinou „pravdivou" alternativu čtyři. Nestálo by za to vyzkoušet, zda výsledky u nás by byly odlišné od výsledků z Kanady?

Ovšem někdy je nepravdivá odpověď záměrná. Obvykle proto, že určitá alternativa odpovědi je pro respondenta spojena s představou nějaké sankce. V Severní Americe může být obtížné získat pravdivé informace o příjmu a finančních záležitostech vůbec. „Co kdyby tazatel měl nějaké spojení s daňovými úřady nebo s konkurencí?" U nás, v ne tak dávné minulosti, hrála silnou zkreslující úlohu obava z politického postihu. Předpovědět možnost takového zkreslení je často nesnadné, v jednom z našich prvých výzkumů, asi v roce 1966, jsme dostali v malých vesnicích na moravskoslovenském pomezí řadu „radioaktivních" odpovědí. Ale v jedné z obcí naprostá většina respondentů odmítla říci, co měli v neděli k obědu. Proč, to nevím. Dobře to vystihuje výrok jednoho mého kanadského kolegy: „Typický respondent se spíše přizná k manželské nevěře, než by připustil, že si nečistí pravidelně zuby."

Validní měření
je takové měření, které měří skutečně to, co jsme zamýšleli měřit.

Reliabilní měření
je takové měření, které nám při opakované aplikaci dává shodné výsledky, pokud se ovšem stav pozorovaného objektu nezměnil.

Obě definice jsou jednoduché a nezní příliš vědecky. Ale uvidíme, že jsou velmi užitečným a pro výzkum nezbytným nástrojem.

Podívejme se však nejdříve na některé příklady narušení validity a reliability. Naše pohádka č. 4 ilustruje takový příklad.

Pohádka pro odrostlejší děti 4

O muzeu a validitě

Bylo jednou v jednom krásném městě krásné Muzeum objevů vědy a techniky. Jeho ředitelem byl podnikavý a sečtělý mladý muž, a ten se jednou dočetl o studii podniknuté v jednom muzeu v Chicagu, která použila opotřebení koberce před jednotlivými exponáty jako ukazatele jejich popularity. Dočetl se také, že tato technika byla použita i jinde, a tak se rozhodl, že ji použije i ve svém muzeu.

Výsledky výzkumu byly poněkud překvapivé. Jako nejúspěšnější byl vyhodnocen exponát věnovaný úspěchům čtyřletého plánu hospodářské výstavby, skříň s uprášenými diagramy a plakáty, věcmi, o které v té době předstírali zdvořilý zájem jen reprezentanti nejvyšších mocenských orgánů.

Výzkum byl publikován, ředitel dostal vyznamenání a muzeum podstatně zvýšený rozpočet. Jenom několik set zasvěcenců vědělo, že kolem nejúspěšnější expozice vede jediný přístup k jediným toaletám v muzeu.

Druhé velké nebezpečí je v narušení reliability. Je to podobná situace, jako bychom měřili délku měřítkem, které se tu zkracuje, tu prodlužuje, a my o tom nevíme. Uveďme si docela realistický příklad nereliabilní otázky, jakou čas od času potkáváme ve výzkumech:

Chodíte do kina
– často
– ne tak často
– občas
– zřídka
– vůbec ne

Odpověď „často" může znamenat pro respondenta A „jednou za měsíc", pro respondenta B „dvakrát týdně", pro mě „dvakrát za rok" a znám jednoho televizního dramaturga, u kterého si význam odpovědi „často" vůbec netroufám odhadnout. V každém případě je tato otázka velice nereliabilní.

Zkusme v následujícím cvičení, jak jsme porozuměli pojmům validity a reliability. Otázka v tomto cvičení je jednoduchá, ale není právě nejlehčí.

Cvičení 3.3

Může nereliabilní měření být validní?

A co opačná situace: Může měření, které je nevalidní, být přesto reliabilní? Asi ano. Je to takové měření, které nám dává konstantní, ale chybné výsledky. Je to podobné situaci ve fyzickém měření, kdy by naše měřítko bylo kalibrováno v jiných jednotkách, než se domníváme. Následující příklad je hodně přehnaný, ale vystihuje dobře náš problém: domníváme se, že měříme v centimetrech, a měřítko bylo kalibrováno v palcích. Při opakovaném měření stejného objektu dostaneme stejné výsledky, které však interpretujeme v centimetrech, a tak jsou naše závěry nevalidní.

Zajistit validitu pozorování je pro sociologický výzkum nesmírně důležité. My opravdu musíme často od zkoumané osoby získat informaci, kterou tato osoba vůbec nezná. Tedy alespoň ne v té formě, v jaké ji hodláme použít pro náš výzkum. Ptát se průměrného respondenta, zda se cítí alienován, bylo by zřejmě absurdní. Nejhorší na tom je, že mnoho respondentů by nám na tuto otázku stejně odpovědělo, třebas by nevěděli, co slovo alienace znamená; nechtěli by prostě přiznat svoji neznalost. I když mnozí z nás mají spíše skeptický názor na měření inteligence, IQ je stále užívaným konceptem. IQ měříme i u osob, které nikdy o tomto konceptu neslyšely. Na některé otázky dnes pravděpodobně nedostanete pravdivou odpověď. Jen málokterý respondent je ochoten přiznat, že je rasista, nebo že bije manželku. Proto se snažíme obdržet takovou informaci různými prostředky nepřímého měření. Povíme si o nich víc v příští kapitole. Ale už teď můžeme snadno odhadnout, že takové měření bude vystaveno značnému riziku zkreslení.

V takových případech je nezbytné kontrolovat, zda navrhované měření je skutečně validní. Většina z následujícího výčtu technik pro kontrolu validity není snadná. Zdaleka ne všechny techniky, zejména ne ty nejúčinnější, je možno aplikovat ve všech situacích. Testování validity přidává další krok do výzkumného procesu. Je to krok časově náročný, a tedy i nákladný. To by nás však nemělo odradit od aplikace těchto technik, kdykoliv to je jen trochu možné. Nezapomeňme na GIGO!

První a pravděpodobně nejmocnější skupina technik pro kontrolu validity je založena na **srovnání s nějakým vnějším kritériem**.

1. Validita založená na členství ve známé skupině

Logika této techniky je velice jednoduchá: zkoušíme náš nástroj na skupině, o níž víme, že má vlastnost, kterou má náš nástroj měřit. Tak např. otázky měřící rasismus by mohly být testovány na členech rasistických organizací. Jejich skóre by mělo být významně vyšší než skóre vzorku z obecné populace. Nepřímé otázky, které mají zjistit, zda respondent bije ženu, by mohly být vyzkoušeny na lidech odsouzených opakovaně pro tento delikt. Hlavní problém je zde existence a dostupnost takové skupiny.

2. Prediktivní validita

porovnává předpověď založenou na testovaném měření se skutečnými výsledky. Např. validita škály, která má předpovídat pravděpodobnost úspěchu středoškolských studentů na

vysoké škole, je testována porovnáním výsledků testu se známkami, které student na univerzitě skutečně docílil. Jen tehdy, je-li shoda obou souborů dat dostatečná, je test prohlášen za validní a bude i nadále používán. Je zřejmé, že tento postup může být smysluplně použit pro testování měření, která mají být používána opakovaně, a nelze jej použít pro jednorázový výzkum.

3. Souběžná validita

je testována tak, že měření téže vlastnosti je prováděno dvěma nebo více různými postupy. Můžeme třeba použít různé formulace nepřímých otázek, kombinovat je s projekčními otázkami (o těch hovoříme v příští kapitole) a můžeme v tomto souboru použít i přímou otázku. Čím větší je shoda mezi těmito různými přístupy k měření, tím je pravděpodobnější, že náš postup je validní. Velikou výhodou souběžné validity je, že tato technika je aplikovatelná takřka univerzálně. Určitou nevýhodou je však to, že plný test validity obdržíme teprve tehdy, když data byla sebrána a analyzována.

V další skupině máme jen jednu techniku:

4. Konstruovaná validita

V této technice je konstruován test hypotéz logicky spojujících zkoumanou vlastnost s technikou měření, kterou používáme. Ilustrujme si to na tomto příkladu:

Profesor P. I. Toma vyvinul škálu pro měření frustrace. Aby dokázal, že je tato škála validní, zkonstruoval soubor hypotéz předpovídajících, za jakých okolností můžeme očekávat vysokou úroveň frustrace:

Hypotéza 1: Očekávání těžkého úkolu v blízké budoucnosti zvyšuje hladinu frustrace.
Hypotéza 2: Čím negativněji je tento úkol vnímán, tím bude frustrace silnější.
Hypotéza 3: Čím méně zajímavý je očekávaný úkol, tím vyšší bude frustrace.

Studenti sociologie, očekávající zkoušku ze statistiky, byli testováni. Jejich frustrace byla významně vyšší než frustrace kontrolního vzorku studentů, kteří neočekávali žádnou zkoušku. Tomova škála byla přijata jako validní ukazatel frustrace.

(Adaptováno z Therese L. Baker, *Doing Social Research*, 1988)

Poslední skupinu testů validity představují testy založené na **obsahu** měřeného jevu:

5. Validita založená na mínění skupiny soudců

Skupina zkušených knihovníků, učitelů literatury, nám může pomoci posoudit, zda naše posouzení čtenářské vyspělosti zkoumaných osob založené na výčtu přečtených knih je přijatelné. Důležité je, aby každý z členů poroty pracoval nezávisle na druhých. Shoda mezi výroky různých soudců je pak dobrým ukazatelem validity.

6. Validita testovaná výčtem obsahu (Domain of Meaning)

To je v podstatě test, zda naše měření dostatečně kryje doménu jevu, který zkoumáme. Tento test může být sotva použit pro validizaci abstraktnějších konceptů. Výčet obsahu je užitečný zejména tehdy, kdy samotná definice zkoumaného konceptu není dostatečným nástrojem pro výzkumné operace. Kupř. srovnávat volný čas zkoumaných osob není snadný úkol. Prostý výčet aktivit, odpovídající definici tohoto testu validity, není dostatečným řešením: jedna a tatáž aktivita je pro jednoho respondenta milovanou rekreací a pro jiného velice nepříjemnou povinností. Proto jsme pro výzkum o využití volného času navrhli tento postup: Po určitou dobu budou zaznamenávány všechny aktivity zkoumaných osob a jejich trvání. Pak budou respondenti dotazováni na každou aktivitu, zda ji považují za volný čas. Takový postup může poskytnout validní, srovnatelnou definici rozsahu volného času.

7. Zjevná validita (Face Validity)

Většina učebnic výzkumných metod uvádí tuto techniku mezi jinými technikami testování validity. My se domníváme, že je to spíše popis situace, kdy kontrola validity není vyžadována. Kupř. jistě není třeba testovat validitu otázky v dotazníku zjišťující respondentovo pohlaví. Bohužel spoléhání na zjevnou validitu je aplikováno i na složitější situace. Babbie (1989, s. 124) uvádí tento příklad: Počet stížností podaných dělníky odborové organizaci by mohl být použit jako ukazatel pracovní morálky. Asi bychom se shodli, že to může mít něco společného s měřenou vlastností. Mohli bychom se shodnout také na tom, že je to lepší ukazatel než řekněme počet knih, které si dělníci vypůjčili z veřejné knihovny. Potud Babbie. Zjevná validita je založena na tom, že intuitivně spoléháme na to, že existuje dostatečná spojitost mezi zkoumaným jevem a naším měřením. Vyjádřeno velice prostě: „Face Validity" je eufemismus pro situaci, kdy kontrola validity nebyla provedena vůbec. Někdy, jako tomu bylo v prvém příkladu v tomto odstavci, není na tom nic špatného. Ve složitějších situacích to může být nebezpečné.

Kdyby se vám snad po téhle kapitole zdálo, že náš profesionální život je příliš jednoduchý, přidejme nakonec ještě jednu nepříjemnou připomínku: Neexistuje nic takového jako univerzálně validní měření určitého konceptu. **Validita prokázaná v určitém kontextu a pro určitý účel je automaticky platná jen pro tento kontext a shodný účel.** Indikátor určitého jevu, který je perfektně validní pro jeden účel, může být nevalidní pro měření téhož jevu v jiné souvislosti a naopak.

Cvičení 3.4

V této kapitole jsme diskutovali otázku měřící frekvenci chození do kina v kategoriích „často", „ne tak často", „občas" atd. Dospěli jsme k závěru, že tato otázka není reliabilní, a tedy není ani validní. Umíme si představit situaci, ve které by byla validním měřením?

To také znamená, že validní indikátory „necestují" dobře: exportujeme-li validní indikátor ze země do země, z jedné kultury do jiné, snadno může ztratit svoji validitu. Kupř. jedna škála měřící validně odcizení v USA obsahuje mimo jiné otázky: Čtete pravidelně Reader's Digest? Chodíte pravidelně do kostela? Negativní odpověď na tyto otázky signalizuje přítomnost odcizení. Když je tato škála exportována třeba jenom do Kanady, výsledky měření jsou poněkud zkreslené: Reader's Digest není v Kanadě příliš populární, chození do kostela se zdá mít v kanadském kulturním kontextu jiný význam než v USA. Importovat takovou škálu v její původní formě k nám by bylo prostě absurdní.

Uveďme si alespoň ještě jeden příklad: v době, kdy se odehrála naše pohádka č. 5, byla ve výzkumu vztahů na pracovišti v USA ještě hojně používána škála zvaná SRA Attitude Survey (Dabas 1958). Tato škála byla považována za spolehlivé měřítko pracovní morálky, za „morale audit".

Pohádka pro odrostlejší děti 5

Co by se mohlo stát...

Náš oblíbený profesor P. I. Toma se rozhodl studovat, jaký vliv na dělníky má participace v dělnické samosprávě. Jeho základní hypotéza byla, že čím vyšší je tato participace, tím vyšší bude pracovní morálka dělníků. Pro měření závisle proměnné použil škálu z SRA Attitude Survey.

Analýza dat potvrdila, že existuje skutečně souvislost mezi participací v samosprávě a morálkou dělníků. Vypočítané koeficienty byly překvapivě silné a velice významné. Problém byl v tom, že koeficienty byly negativní: Čím vyšší participace, tím nižší morálka.

Profesor P. I. Toma přijal hypotézu, že participace v dělnické samosprávě podvrací pracovní morálku, zvyšuje jejich alienaci, a je proto zavrženíhodná. Problém je v tom, že v SRA škále je považován za indikaci nízké pracovní morálky souhlas s tvrzeními jako:

29. Špatné pracovní podmínky mi znemožňují, abych podal nejlepší pracovní výkon.

42. Vedení nám nedává jasné příkazy a instrukce.

44. Vedení nás neinformuje o věcech, které bychom měli vědět.

62. Náš šéf by měl být přátelštější k zaměstnancům.

Náš P. I. Toma si neuvědomil, že tato tvrzení, stejně jako mnoho jiných ve škále, vyjadřují postoje, které by měl mít každý člen dělnické samosprávy. Zřejmě mu na jeho univerzitě utajili, že posláním samosprávy je zaujímat kritický postoj k formálnímu vedení závodu. Domníval se, že měří souvislost mezi členstvím v samosprávě a pracovní morálkou. Ve skutečnosti však měřil souvislost mezi členstvím a postoji, které jsou očekávány od členů samosprávy. Není tedy divu, že jeho koeficienty byly vysoké: koreloval jen dvě části jednoho jevu.

Výzkum popsaný v naší pohádce nám tedy neřekl nic o vlivu samosprávy na pracovní morálku. Ale není bezcenný. Ukázal nám zřetelně, že měření, jehož validita byla uspokojivě prokázána v jednom sociálním prostředí, může být tragicky (či komicky) nevalidní na jiné populaci.

Do stejné kategorie patří i problémy třídního, sexistického a rasového zkreslení testů inteligence. Jejich výsledky snad mohou být validní pro bílé studenty z rodin vyšší střední třídy,

tedy v prostředí, ve kterém byly vyvinuty a ve kterém byla jejich validita nejčastěji testována. Aplikujeme-li však tyto testy na černého kluka z ghetta, jehož negramotní rodiče živoří na veřejné podpoře, nebo na romskou ženu z východoslovenské cikánské vesnice, výsledky mohou být daleko méně validní.

Pokud si snad myslíte, že našich trampot s logikou empirického výzkumu v sociologii není ještě dost, máme tu ještě jednu obzvláště pěknou Jobovu zvěst. Měli bychom si ji velice dobře pamatovat:

> Všechny závěry z empirického výzkumu, statistické důkazy atd. se nevztahují na jevy, které chceme studovat, ale jen na jejich indikátory.

Řekněme, že sledujeme vliv vzdělání na politickou participaci. Všechny pozorované vztahy, souvislosti, statistické důkazy se skutečně nevztahují na tyto dva koncepty, ale jen na jejich ukazatele. Skutečně tedy měříme souvislost mezi léty formálního vzdělání a faktem, zda respondent hlasoval v posledních volbách, přispíval finančně na volební fond nějaké strany, chodil do průvodu a prapory vyvěšoval nebo přilepil volební heslo na nárazník svého auta. Někdy je rozdíl mezi indikátorem a studovaným konceptem zanedbatelný. Častěji je však zdravé pořádně se zamyslit nad vzdáleností mezi konceptem a jeho ukazateli. Ale k tomuto problému se v naší knize ještě několikrát vrátíme.

Řešení úkolů z kapitoly 3

Cvičení 3.1

Do této kategorie patří formální vědy jako matematika a formální logika. Jejich předmětem je čirá abstrakce, a tak kategorie překladu „reality" do jejího popisu tu není aplikovatelná.

Cvičení 3.2

To je opravdu špatná otázka. Nemůžeme vědět, zda částka, kterou respondent uvedl, představuje jeho čistý nebo hrubý příjem, zda zahrnuje také příjem z vedlejšího povolání. Nevíme, zda je to týdenní, dvoutýdenní nebo měsíční příjem. Někteří respondenti nám odpovědí hodinovým příjmem. Rozhodně nám tato otázka neposkytne výsledky srovnatelné pro všechny osoby v našem vzorku.

Cvičení 3.3

Měření, které není reliabilní, nám nedává stabilní výsledky. V každém pozorování měří něco jiného, nebo měří jinak. Taková operace tedy měří něco jiného, než má. Je tedy nevalidní.

Cvičení 3.4

Tato otázka je nevalidní jakožto ukazatel kvantity chození do kina. Mohla by však být validním ukazatelem toho, jak je nasycena respondentova potřeba chodit do kina. V tom případě by bylo užitečné měřit frekvenci jinou, objektivnější otázkou, třeba: „Kolikrát jste byl za poslední měsíc v kině?"

4

Jak si nakreslit plán aneb Na co jsou hypotézy

dedukce
indukce
výzkum jako testování hypotéz
pracovní hypotézy
operační hypotézy
funkce pracovních hypotéz
test zkoumatelnosti
optimalizace redukce
rozhodnutí o technice sběru dat
rozsah výzkumu

Kapitola 4

Jak si nakreslit plán aneb Na co jsou hypotézy

> I'm short enough and ugly enough to succeed on my own.
>
> Woody Allen, *Play it again Sam*

Až dosud jsme si jenom stěžovali. Stěžovali jsme si na nezvládnutelnou velikost přirozených systémů v sociálních vědách, na trampoty, které z toho vyplývají: na nutnost pracovat s neúplně popsaným a ještě neúplněji analyzovatelným systémem, na to, že jsme v důsledku této **redukce** vystaveni mnoha typům zkreslení. Stěžovali jsme si na to, že jsme schopni potvrdit existenci příčinné závislosti téměř výhradně jen experimentem.

A stěžovali jsme si také na náš druhý veliký problém: na nutnost nepřímého měření, na to, že **transformace** studovaného jevu do indikátorů, ve kterých tento jev pozorujeme, měříme a analyzujeme, je mnohem delší a mnohem riskantnější než ve většině jiných věd. Naříkali jsme nad nejistou validitou našeho měření.

V téhle kapitole se podíváme na to, co se s tím dá dělat. Neočekávejte tu žádné zázračné recepty. Udělat slušný empirický výzkum, to dá fušku. Mnoho a mnoho hodin přemýšlení a psaní ještě předtím, než máme vůbec možnost ušpinit si ruce fakty. Dlouhé dny v terénu, dlouhé dny strávené analýzou dat, a to se nakonec může stát, že naše hypotéza bude fakty vyvrácena. Alespoň jsme zjistili, že tudy cesta nevede. Horší je, když ji naše výsledky ani nepotvrdí, ani nevyvrátí. I to se někdy stane. Ale přesto přese všechno je to veliká krása, dělat výzkum. Vždyť studovat vztahy mezi lidmi je velké dobrodružství. Pevnina sociálních vztahů a struktur je ještě jen málo prozkoumaná a zaplňovat prázdná místa na mapě je veliké dobrodružství. Zkuste to, a uvidíte, že se vám to bude líbit. Ale pozor, je to návykové.

V téhle kapitole se budeme zabývat spíše logickými pravidly výzkumu než podávat recepty, jak technicky připravit výzkum. Tím se zabývá celá řada jiných knih. Věříme však, že tato

pravidla jsou důležitou pomůckou i pro optimalizaci rozhodnutí o mnohých technických krocích ve výzkumu.

Ještě jednu důležitou poznámku, než opravdu začneme: v této kapitole, stejně jako v předchozích, se budeme zabývat hlavně kvantitativním výzkumem, a to ještě většinou v jeho silně standardizované verzi. Není to proto, že bychom tento typ výzkumu považovali za nadřazený. Jediný důvod pro tento popis je fakt, že v kvantitativním výzkumu jsou všechna naše zranitelná místa jasně viditelná.

4.1 Výzkum jako idiot

Můžeme začít třeba pravdivou pohádkou z počítačového pravěku:

Pohádka pro odrostlejší děti 6

O tom, co je počítač

Před dávnými a dávnými časy byl představen tisku jeden z prvních počítačů. Byl obrovský, nesmírně drahý, pomalý a neuměl o moc víc než trochu lepší kalkulačka. Ale protože to bylo dávno a dávno, novináři byli okouzleni. Hovořili s jedním z tvůrců toho počítače:

Novinář: „Pane profesore, ten počítač je génius!"

Profesor: „Ale pánové, to se mýlíte. To je úplný idiot. Umí jen jednu věc: rozeznat nulu od jedničky. Ale umí to velmi rychle a nedělá chyby."

Empirický kvantitativní výzkum v sociologii je také takový idiot. Umí jenom rozeznat, zda existuje souvislost mezi dvěma nebo více proměnnými. Bohužel neumí to rychle a dělá mnoho chyb.

Kvantitativní výzkum není nic jiného než testování hypotéz. Podívejme se blíže, co to vlastně znamená v termínech epistemologie vědy. Při produkci poznání zde užíváme v podstatě dva odlišné postupy: indukci a dedukci.

Deduktivní metoda vychází z teorie nebo z obecně formulovaného problému. Teoretický nebo praktický problém je přeložen do jazyka hypotéz. Hypotézy navrhují, jaké spojení mezi proměnnými bychom měli najít, je-li naše hypotéza pravdivá. Pak následuje sběr dat. Odpovídají-li závislosti mezi sebranými daty vzorci předpovězenému v hypotézách, přijmeme hypotézy jako platné. Jinak musíme hypotézy odmítnout.

Induktivní metoda je značně odlišná. Začíná pozorováním, ve kterém pátráme po pravidelnostech, vzorcích, které snad existují v objektivní realitě. Objevené pravidelnosti popíšeme ve formě předběžných závěrů. Ty pak ověřujeme dalším pozorováním. Konečným produktem je nová teorie.

A teď si zkusme, zda jsme rozdílům mezi dedukcí a indukcí dobře porozuměli:

Cvičení 4.1

Kterou z těchto metod používal velmistr dedukce, slavný detektiv Sherlock Holmes?

Tabulka 4.1

DEDUKCE	INDUKCE
Teorie	Pozorování
▼	▼
Hypotézy	Nalezené pravidelnosti
▼	▼
Pozorování	Předběžné závěry
▼	▼
Přijaté/zamítnuté hypotézy	Teorie

Kvantitativní výzkum používá deduktivní metodu. To znamená, že je schopen řešit jen určitou kategorii problémů:

> Kvantitativní výzkum může nalézt řešení jen pro takové problémy, které je možno popsat v termínech vztahů mezi pozorovatelnými proměnnými.

Přeformulovat problém tak, aby vyhovoval těmto podmínkám, přeložit teoretický problém do jazyka výzkumu, nemusí být snadné. Naštěstí máme nástroj, který nám tuto operaci podstatně ulehčí. Jsou to **pracovní hypotézy**. Je to velice užitečný nástroj, který nám pomůže minimalizovat problémy vyplývající z redukce a transformace informací, nevyhnutelné v sociologickém výzkumu.

4.2 Aby hypotézy nebyly jen hypotetické

Cesta od teorie k výzkumu je dlouhá a musíme se snažit, aby nebyla příliš klikalá. První krok od teorie k výzkumu je obvykle formulace základních, ještě dost obecných hypotéz. Taková hypotéza může být definována třeba takto: *„Hypotéza je očekávání o charakteru věcí, vyvozené z teorie"* (Babbie 1979). Uveďme si alespoň jeden příklad takové hypotézy:

> Hypotéza 1
>
> Čím více podnětů týkajících se politiky osoba dostává, tím vyšší je pravděpodobnost, že bude participovat v politice.
>
> Lester W. Milbrath, *Political Participation*, Chicago 1965, s. 39

Tato hypotéza je mnohem konkrétnější než většina hypotéz na této úrovni. Přesto nemůže být ještě bezprostředním východiskem pro technickou konstrukci výzkumu a výzkumných nástrojů. Je třeba ji rozložit do souboru **pracovních hypotéz**.

My definujeme pracovní hypotézu takto:

Tabulka 4.2

1. Pracovní hypotéza je tvrzení předpovídající existenci souvislosti mezi dvěma nebo více proměnnými.
2. Všechny proměnné zmíněné v hypotéze musí mít validní operační definici.
3. Soubor pracovních hypotéz musí zahrnovat nejen proměnné reprezentující zkoumané koncepty, ale i ty proměnné, které mohou významně zkreslit interpretaci testovaných vztahů.

Především si musíme ujasnit, co znamená termín „operační definice". Začněme jednoduchým příkladem. Co je vlastně kyselina? Chemik by nám nabídl definici popisující „kyselinu jako sloučeninu s vodíkovým atomem, který je schopen odštěpit se ve vodním roztoku jako vodíkový kation H+ a udělit roztoku kyselou reakci..." (Příruční slovník naučný, 1963, II. díl, s. 741).

Operační definice kyseliny by mohla znít asi takto: *Kyselina je taková tekutina, která změní barvu lakmusového papírku z modré na červenou.* Vidíme, že operační definice je mnohem méně precizní, daleko volněji spojená s teoretickými koncepty. **Nevysvětluje nám vůbec, co kyselina je.** Má však jednu velice užitečnou vlastnost: dává nám bezprostřední **návod, jak poznat, co je kyselina**.

V operační definici je koncept vyjádřen popisem operací, kterými bude měřen.

Ovšem i operační definice musí být zakotvena v teorii. Problém je, že v sociálních vědách je vzdálenost mezi teoretickým konceptem a jeho operační definicí značná, a to může ohrozit

validitu našich operací. Ale to všechno už známe: diskuse o operačních definicích je v podstatě opakováním naší diskuse o indikátorech z předchozí kapitoly.

Podívejme se zde alespoň na jeden klasický příklad operacionalizace. Emile Durkheim uvedl ve své klasické knize *Sebevražda* (1897) koncept anomie. Anomie je obvykle definována jako nepřítomnost norem. V dobách dramatických sociálních změn staré sociální normy ztrácejí platnost a nové normy nejsou ještě obecně přijímány. Jednotlivec si již není jist, co bude spojeno s negativní sankcí a co s odměnou. Durkheim zvolil sebevraždu, „výsostně individualistický a osobní akt", za ukazatel anomie. Nesmírně komplexní teoretický koncept anomie byl silně **redukován**: anomie určité oblasti je operačně definována jako počet sebevražd za rok připadajících na každých sto tisíc obyvatel. I stručná úvaha nám ukáže, jak mnoho zajímavých a důležitých dimenzí původního konceptu je v této operacionalizaci ztraceno.

Redukce informací je nevyhnutelné břemeno, které nás bude provázet na každém kroku výzkumných operací. To však nemusí být fatální. Zhoubné by to bylo tehdy, kdybychom uvěřili, že operační definice je perfektní reprezentací teoretického konceptu, že můžeme skutečně položit rovnítko mezi anomii a frekvenci sebevražd.

Naštěstí praxe empirického sociologického výzkumu se obvykle zabývá koncepty značně jednoduššími, než je Durkheimova anomie. Uvidíme také brzy, že empirický výzkum není s to testovat platnost „grand theories" (a většinou ani teorií středního rozsahu) jako celku. Tyto operace jsou jednoznačně doménou teorie.

A teď se už můžeme vrátit k mnohem skromnějším konceptům a zamyslet se nad jejich překladem do jazyka pracovních hypotéz. Vraťme se k Milbrathově hypotéze o participaci v politice:

> Hypotéza 1
> Čím více podnětů týkajících se politiky osoba dostává, tím vyšší je pravděpodobnost, že bude participovat v politice.

Podívejme se, jak tato hypotéza odpovídá naší definici pracovní hypotézy uvedené v tabulce 3.2. Hypotéza předpovídá existenci spojení mezi dvěma proměnnými, ale pro žádnou z nich

nemáme operační definici a ještě nevíme, co znamená třetí podmínka zmíněná v tabulce. Je však zřejmé, že hypotéza 1 není pracovní hypotézou, že musí být do ní teprve přeložena.

Nejdříve můžeme začít s pátráním po měřitelných ukazatelích politické participace. Není těžké vyjmenovat jich celou řadu. Za krátkou dobu bychom mohli skončit se seznamem obsahujícím několik desítek přijatelných indikátorů. Ale které z nich vybrat? Takový seznam může obsahovat všechno, od faktu, že respondent je třeba předsedou vlády, kandiduje na politickou pozici, přes účast na demonstracích, až po prosté hlasování ve volbách a všichni víte z nedávné minulosti, že stávka, účast na mši, na určitém divadelním představení nebo koncertě může být navýsost politickým aktem. Stejně dobře si umíte představit, že záměrná neúčast v oficiálních kanálech politické participace může být velice silným a významným vyjádřením politického postoje.

Co tedy vybrat? Záleží především na cíli výzkumu, na teoretickém rámci, ve kterém chceme politickou participaci studovat. Milbrathova hypotéza by mohla být použita právě tak ve studii zabývající se psychologickými a filozofickými dimenzemi politické participace jako ve výzkumu snažícím se porovnat participaci ve standardních politických kanálech u různých skupin obyvatelstva. Operační definice politické participace by ovšem musela být v obou případech různá.

A ovšem definice závisí na konkrétním politickém a historickém kontextu. Operační definice politické participace obyvatel hlavního města Prahy 1991 by se podstatně lišila od definice participace na téže populaci o pět let dříve. Jiné indikátory musí být použity pro měření politické participace v různých politických systémech. Indikátory použitelné v demokratických zemích by sotva daly smysluplné výsledky, kdyby byly aplikovány v zemi, kde je demokracie silně omezena. I v zemích se shodným politickým systémem mohou být ukazatele platné v jedné zemi neplatné v jiné, stejně demokratické. Prostě se tyto země liší v politické kultuře, různé politické akty mohou mít různý význam anebo nemusí mít v určité zemi vůbec žádný význam.

Participace v běžných formách politiky je v Severní Americe měřena často ukazateli podobnými těm uvedeným v naší tabulce 4.3. Tyto ukazatele bývají často exportovány do jiných zemí, bohužel často jen mechanicky, a to může vyvolat povážlivé zkreslení, nebezpečné zejména ve srovnávacím výzkumu. K tomu se vrátíme v jedné z našich budoucích pohádek. Bude to smutná pohádka, protože je pravdivá.

Naše tabulka je adaptací ukazatelů hierarchie politické participace navržených Milbrathem (Milbrath 1982, s. 18). Indikátory jsou řazeny sestupně. První skupina zahrnuje aktivity, které jsou v angloamerické literatuře označovány jako „gladiátorské" aktivity, poslední skupina „divácké" aktivity. Prostřední reprezentuje přechod mezi oběma vyhraněnými typy.

Tabulka 4.3

Výkon veřejné funkce nebo funkce v politické straně Kandidování na takovou pozici Vybírání peněz pro politické účely Participace na politických schůzích plánujících strategii Aktivní členství v politické straně Aktivní účast v předvolební kampani
Účast na politických schůzích a shromážděních Peněžní dary politické straně nebo kandidátovi Psaní dopisů veřejným nebo politickým funkcionářům
Politické plakáty na nárazníku auta, na trávníku před domem Přemlouvání druhých, aby hlasovali určitým způsobem Iniciování politické diskuse Hlasování

Ponechme stranou otázku platnosti hierarchie jednotlivých ukazatelů a podívejme se na problém validity těchto ukazatelů v našich poměrech.

Cvičení 4.2

Zkuste odhadnout, které z těchto ukazatelů by mohly vyvolat zkreslení, kdyby byly použity pro výzkum politické participace v našich poměrech.

I když bychom pravděpodobně pro výzkum u nás některé z těchto ukazatelů vypustili, je zřejmé, že naše závisle proměnná, politická participace, musela být měřena v celé řadě proměnných. Protože si nemůžeme být jisti, že každý z indikátorů odráží politickou participaci ve stejné míře, **každý z indikátorů musí být použit ve zvláštní pracovní hypotéze**.

Podívejme se teď na naši nezávisle proměnnou – expozici politickým stimulům. Ta by mohla být sledována kupř. následujícími ukazateli:

– frekvence sledování politických pořadů v TV
– frekvence sledování politických pořadů v rozhlase
– frekvence čtení politických úvodníků
– politické debaty v rodině
– politické debaty s přáteli
– odebírání nebo pravidelná četba politicky orientovaných periodik

Jistě by bylo možno v tomto výčtu pokračovat. Tradice politické participace v rodině může být velmi silným politickým stimulem. Tento jev by ovšem musel být operacionalizován do několika proměnných (kupř. kdo z rodiny participoval v politice; jak; jak intenzivně; kdy atd.). Jiným, pravděpodobně silným, politickým stimulem by mohla být četba knih s politickou tematikou, a tak bychom mohli ještě hodnou chvíli pokračovat v našem výčtu.

Pro jednoduchost předpokládejme, že pro popis politických stimulů použijeme jenom prvých šest proměnných. Každá z nich ovšem musí být použita v samostatné hypotéze; každý z těchto politických podnětů může mít různě silný vliv na různé typy politické participace. Řekněme, že bychom pro popis politické participace použili třeba jen šest proměnných. Některé z nich jsou u nás sotva validní, jak jste sami zjistili ve cvičení 4.2. U jiných je frekvence výskytu tak nízká, že je v řadě výzkumů můžeme ignorovat (kupř. v USA v prvých třech typech „gladiátorské" participace působila méně než 2 % populace).

Máme tedy 6 proměnných popisujících politické stimuly a 6 dalších, popisujících politickou participaci. Naše pracovní hypotézy by mohly vypadat třeba takto:

– Osoby, které čtou pravidelně politické úvodníky, hlasují ve volbách častěji než osoby, které je nečtou.

– Osoby, které čtou pravidelně politické úvodníky, jsou častěji aktivními členy politické strany než osoby ostatní.

- Čím častěji respondent sleduje televizní politické pořady, tím vyšší je pravděpodobnost, že se bude aktivně podílet na předvolební agitaci.

Abychom popsali všechny vlivy stimulů na všechny typy participace, museli bychom zformulovat 36 podobných pracovních hypotéz. Vidíme, že nám počet hypotéz požehnaně roste, a to ještě nejsme zdaleka u konce s naším úkolem. Zatím jsme se ještě ani nestačili vypořádat s třetí podmínkou z naší definice pracovní hypotézy: s nutností zahrnout do souboru hypotézy o proměnných, které by mohly podstatně zkreslit naši interpretaci vztahů mezi diskutovanými proměnnými.

A opět je tu naše stará bolest: komplexnost přirozených systémů v sociálních vědách. Podívejme se alespoň na nejmarkantnější příklady:

I kdyby čtení úvodníků nemělo vůbec žádný vliv na aktivní členství v politické straně, je vysoce pravděpodobné, že souvislost mezi čtením a členstvím stejně naměříme. Je to projev naší staré známé z prvé kapitoly – nepravé korelace. Ve většině demokratických společností vzdělanější osoby participují v politice více než lidé méně vzdělaní. Ale více vzdělaní respondenti budou také s daleko větší pravděpodobností vystaveni politickým stimulům. Je tedy možné, že souvislost mezi politickými stimuly a participací je zcela nebo částečně vyvolána vlivem vzdělání na obě diskutované proměnné. Podobnou roli může hrát kupř. i sociálně-ekonomický status. Osoby v určitém ekonomickém postavení mohou být angažovány v politice, protože je to dobré pro jejich obchod. Z téhož důvodu pátrají po informaci o politických problémech a záměrně se tedy vystavují politickým stimulům.

Takových situací, kde nějaká vnější proměnná ovlivňuje vztah mezi studovanými jevy, může být mnoho. Naštěstí statistická analýza je schopna odizolovat vliv těchto vnějších proměnných. Můžeme toho však docílit jen tehdy, víme-li o každém respondentovi, jaké je jeho vzdělání, socioekonomický status atd. I kdyby náš výzkumný cíl byl velice úzký, kdybychom se zajímali jenom o vztah mezi politickými stimuly a participací, musíme do našeho souboru pracovních hypotéz zahrnout třeba tyto:

- Čím vyšší je jednotlivcovo vzdělání, tím je vyšší pravděpodobnost, že bude navštěvovat politické schůze a shromáždění.
- Osoby s vyšším vzděláním čtou politické úvodníky častěji než osoby méně vzdělané.

- Osoby, které vlastní v obci obchod, budou participovat v předvolební kampani pro místní volby častěji než osoby ostatní.

Takových hypotéz budeme muset formulovat mnoho: víme z literatury i z denní zkušenosti, že účast v politických aktivitách je silně ovlivněna proměnnými jako pohlaví, věk, sociabilita jedince, jeho sociální sebedůvěra a mnoha dalšími a většina z těchto proměnných může být asociována s tím, jak silně je jedinec vystaven politickým stimulům.

> Soubor pracovních hypotéz musí zahrnout hypotézy o všech proměnných, které mohou mít významný vliv na závisle proměnnou a jsou přitom asociovány také s nezávisle proměnnou.

A ještě jedna důležitá terminologická poznámka: **termín „závislá" a „nezávislá" proměnná neznamená existenci kauzální závislosti** a často ani ne očekávaný směr vlivu. Většinou je to jen terminologická konvence, jejíž užitečnost vyplyne, až se budeme zabývat statistickou analýzou. Kupříkladu v našem případě kdyby naše obecná hypotéza byla potvrzena daty, tento výsledek může být interpretován v termínech expozice stimulům ovlivňujícím participaci stejně dobře jako vliv participace na expozici stimulům. Ale to všechno už dobře známe z diskuse kauzality a experimentu v prvé kapitole.

4.3 Je to k něčemu?

Teď už jasně vidíme, že konstrukce pracovních hypotéz je nesnadný úkol již jen vzhledem k pouhému počtu nezbytných hypotéz. Je to úkol náročný, většinou spíše nudný, avšak bohužel nezbytný. Pracovní hypotézy hrají ve výzkumu několik velice důležitých rolí:

1. Formulace pracovních hypotéz je testem, zda je výzkum vůbec možný

V kvantitativním výzkumu můžeme zkoumat jen takové problémy, které je možné přeložit do jazyka pracovních hypotéz, tj. jen takové problémy, jež je možno vyjádřit jako vztahy mezi proměnnými, pro které máme validní operační definici.

2. Pracovní hypotézy jsou důležitým nástrojem pro optimalizaci redukce informací

Lék na naši nejnebezpečnější chorobu, nutnost pracovat s neúplným popisem reality, nebude asi nikdy vynalezen. Musíme se tedy naučit, jak optimálně s touhle nemocí žít. Musíme si být rozumně jistí, že jsme vybrali alespoň ty nejdůležitější proměnné, které ovlivňují zkoumaný problém, a zároveň se vyhnout sbírání informací, které nutně nepotřebujeme. Víme už sice, že zavedení proměnných, které nepatří do zkoumaného systému, nevyvolá zkreslení, ale každá proměnná navíc znamená investici času, peněz a energie, a to je něco, co si prostě nemůžeme dovolit. Většinou jsou naše limity překročeny ještě dříve, než jsme zahrnuli do výzkumu proměnné, které považujeme za nezbytné. Omyly tohoto typu jsou bohužel dosti časté. Náš výzkum má totiž ještě jednu populární dětskou nemoc; začíná zpravidla formulací otázek pro interview nebo dotazník. To je zajímavá, vzrušující a tvořivá činnost; zdaleka ne tak nudná jako formulace pracovních hypotéz. Skončíme se seznamem chytrých a zajímavých otázek, které s naším problémem **nějak** souvisejí. Bohužel při analýze dat zjistíme, že pro řadu proměnných nemáme žádné použití. To není moc dobré. Ale horší je, když při analýze zjistíme, že kdybychom položili jednu jedinou hloupou otázku navíc, mohli jsme opravdu něco objevit. To je opravdu nepříjemné, já vím. Sám jsem si to taky vyzkoušel.

Ovšem existence pracovních hypotéz, jak svědčí naše pravdivá pohádka č. 7, nemusí být automatickou pojistkou proti katastrofám tohoto typu.

Pohádka pro odrostlejší děti 7

Pohádka v prvé osobě jednotného čísla

Bylo, nebylo, byl jsem kdysi zodpovědný za metodologii mezinárodního srovnávacího výzkumu. V průběhu se k nám připojili kolegové z jedné půvabné země. Převzali většinu našich pracovních hypotéz, přidali několik svých, změnili několik operačních definic tam, kde by naše původní indikátory byly v jejich zemi neplatné, a pečlivě testovali validitu nových indikátorů. Připravili takřka perfektní výzkumný nástroj.

Ale když přišla data na zpracování do Toronta, nevěřil jsem svým očím: Hledal jsem a hledal, ale nikde jsem nemohl najít údaj o pohlaví respondenta. Při přepisování dotazového archu vypadla rubrika pro záznam pohlaví. Nikdo si toho nevšiml, až bylo pozdě. U některých osob bylo možno pohlaví odhadnout. Uvedl-li respondent povolání „žena v domácnosti", pravděpodobně to nebyl muž. Ale ve většině případů museli tazatelé znovu do terénu. Nechtěl bych být v jejich kůži. Nemusí být právě příjemné zaklepat znovu na téměř tři sta dveří s pošetilou otázkou: Jste muž, nebo žena?

A tak, máme-li smůlu, ani pracovní hypotézy nepomohou.

3. Pracovní hypotézy obsahují základní informaci pro optimální rozhodnutí o technikách výzkumu

Tak jsme zase nazpátek u naší dětské nemoci „začněme s dotazníkem". Nejenže dotazník se v podmínkách postindustriální společnosti stává pomalu, ale jistě nepoužitelným monstrem, ale na samém začátku výzkumu nemáme prostě dostatek informací k rozhodnutí, kterou z dost bohaté pokladnice technik sběru informací vybrat. Teprve pracovní hypotézy nám umožní se optimálně rozhodnout. Tomuhle problému však věnujeme později celou kapitolu. Pracovní hypotézy, protože popisují strukturu spojení mezi proměnnými, ukáží nám, alespoň v hrubých obrysech, jaký typ statistických operací budeme potřebovat. I v tomto časovém stadiu výzkumu tak budeme schopni odhadnout, zda pro analýzu dat použijeme kalkulačku, osobní počítač, nebo zda budeme muset – je-li náš vzorek opravdu obrovský – bojovat o čas na střediskovém počítači.

4. Pracovní hypotézy jsou základem pro odhad rozsahu výzkumu

Většinou teprve potom, když jsme připravili soubor pracovních hypotéz, jsme schopni alespoň přibližně odhadnout, jak náročný bude náš výzkum. Obvykle budeme překvapeni; teprve proces formulace pracovních hypotéz nám ukáže, jak mnoho proměnných, a ještě více analyzovaných vztahů mezi nimi, potřebujeme pro řešení zdánlivě jednoduchého problému. Pracovní hypotézy jsou důležitým výchozím bodem pro časové, a tedy i finanční plánování výzkumu. A na závěr ještě jedno velice nevědecké, ale užitečné doporučení pro odhad času, který budeme potřebovat: udělejme nejdříve spíše pesimistický odhad, násobme odhadnutý čas třemi, a když všechno dobře půjde, snad tento nový termín splníme. Tenhle spíše prostoduchý recept se nám už často dobře osvědčil.

Vidíme tedy, že ty nepříjemné pracovní hypotézy mohou být velice užitečné. Neměli bychom se jim vyhýbat. Alespoň ne ze začátku, alespoň ne po prvých deset dvacet let výzkumné praxe. Ale pak už je z toho návyk; dost zdravý a hlavně bezpečný.

Řešení úkolů z kapitoly 4

Cvičení 4.1

Bohužel tento mistr dedukce používal induktivní metodu. Snažil se najít vzorec, vysvětlení, které existuje ve fyzických stopách chování. Podívejte se pro potvrzení na kteroukoliv z Doyleových povídek, ve které tento hrdina vystupuje, a pokud by vám tento důkaz nestačil, přečtěte si Skandál v Čechách, kde Holmes explicitně popisuje svoji gnozeologii. Ale s tím se v téhle knížce ještě setkáme.

Cvičení 4.2

Některé z ukazatelů prostě neodpovídají naší politické kultuře a tradicím, nicméně jejich použití by nevyvolalo vážné zkreslení. Jen frekvence pozitivních odpovědí by byla nízká. Psaní dopisů veřejným činitelům, finanční příspěvky na politické účely jsou u nás mnohem řidší než v USA, kde byla tato stupnice vytvořena. Jedna z nejběžnějších forem amerického politického výrazu, „bumperstickers", a volební tabule umístěné na trávníku před rodinnými domky u nás téměř neexistují. Když však jsou takové indikátory použity beze změn ve srovnávacím výzkumu, vyvolají nebezpečné etnocentrické zkreslení. S jedním takovým případem se později seznámíme.

5

Kolik vran musíme pozorovat?

populace
vzorek
hladina významnosti
směrodatná chyba
pravděpodobnostní výběr
náhodný vzorek
interval spolehlivosti
jednoduchý náhodný výběr
systematický výběr
náhodný stratifikovaný výběr
vícestupňový výběr
kvótní výběr
anketa
technika sněhové koule
teoretická nasycenost vzorku
zkreslení samovýběrem

Kapitola 5

Kolik vran musíme pozorovat?

Když nevíš, co děláš, zeptej se někoho, kdo to ví.

Jerry Pournelle, téměř v každém čísle magazínu *BYTE*

Tohle je opět kapitola o redukci informací. Je to kapitola přece jen radostnější než ty předchozí. Redukce populace na vzorek má dobře propracovanou teorii i dobře vypracované a spolehlivé recepty. Některé operace tu nejsou snadné, ale je mnoho lidí, kteří je znají a mohou nám poradit. Buďte tedy zadobře se statistiky.

Touto kapitolou vstupujeme do spíše technické oblasti výzkumu. K tomu se nám může hodit dobrý pomocník. Dovolte, abych vám představil Dr. Watsona.

Dr. Watson je svým způsobem chytrý muž na systemizovaném místě pitomce. Je to někdo, koho každý profesor touží mít ve třídě. Doktor Watson vždycky navrhne nějakou zdánlivě zřejmou, ale ve skutečnosti pitomoučkou odpověď, čímž umožní profesorovi nabídnout správnou odpověď, a tak se zaskvít svojí moudrostí a učeností. Budeme služeb Dr. Watsona hodně používat.

5.1 Vzorek z nouze

Začneme spíše stupidní otázkou: „Kolik vran musíme pozorovat, abychom mohli říci, že všechny vrány jsou černé?" Odpověď je tak jednoduchá, že po ní nemusíme pátrat na konci kapitoly, a přirozeně zní: „Všechny!" Na druhé straně asi nikdo nikdy nepozoroval všechny vrány. Nezbývá nám nic jiného než se spokojit s tvrzením, že „většina vran je černých". Opět je to něco, co už známe: redukovaná analýza reality vede k tvrzením pravděpodobnostního charakteru.

Skupiny, o které se v sociologickém výzkumu zajímáme, nejsou malé. V kvantitativní verzi výzkumu jsme schopni zkoumat celou skupinu jenom výjimečně. Pravidelně jedině sčítání lidu je studií celé populace. Většinou studujeme jen některé členy skupiny a doufáme, že naše závěry budou aplikovatelné i na ostatní, na ty nestudované. To nás přivádí k dvěma základním termínům, které potřebujeme pro tuhle kapitolu: populace a vzorek (výběrový soubor). Jejich definice je jednoduchá:

VZOREK:	skupina jednotek, které skutečně pozorujeme
POPULACE:	(neboli základní soubor) je soubor jednotek, o kterém předpokládáme, že jsou pro něj naše závěry platné

Náš stěžejní úkol je najít postup, aby výsledky, které získáme na vzorku, byly co nejvíce podobné těm, které bychom získali na celé populaci. První věc, která nám přijde na mysl, je snažit se mít vzorek co největší. Ale naše následující pravdivá pohádka nám ukáže, že to není jen tak:

Pohádka pro odrostlejší děti 8

O hodně velkém vzorku aneb Jak to nevyšlo

Byl jednou v Americe velice rozšířený týdeník, který se jmenoval Literary Digest. Byl u svých čtenářů hodně oblíben. Byl proslulý také tím, že spolehlivě předpovídal výsledky prezidentských voleb. Jeho předpovědi byly založeny na obrovském vzorku dvou milionů voličů. (Dnes jsou podobné předpovědi založeny na vzorku tisíckrát menším.) Vzorek byl zkonstruován z mnoha zdrojů. Literary Digest si opatřil adresy voličů z celých USA. Používal pro to zdroje jako telefonní seznamy, městské adresáře, adresy držitelů řidičských průkazů, členské seznamy organizací, seznamy předplatitelů novin a časopisů atd.

Předpovědi byly přesné a úspěšné ve volbách 1920, 1924, 1928, 1932. A pak přišly volby v roce 1936. Literary Digest předpověděl, že prezidentský kandidát Landon porazí Roosevelta rozdílem 14 %. Přišel volební den a s ním i konec slávy Literary Digestu: Franklin Delano Roosevelt zvítězil drtivou většinou.

Cvičení 5.1

Reprezentoval vzorek použitý Literary Digestem dobře celou populaci voličů v USA?

To nebylo tak těžké, že? Trochu složitější je otázka, jak je možné, že vzorek, který prakticky vyloučil z výzkumu voliče náležející k nižším sociálním třídám, fungoval dobře v předchozích volbách? Klíčem k řešení je rok: v roce 1935 vrcholila v USA hospodářská krize, a to vedlo k ostré polarizaci podle vertikální stratifikační osy. Předtím sociálně-ekonomický status nehrál příliš důležitou roli v otázce volebních preferencí. Daleko větší úlohu hrály takové faktory jako náboženství, zeměpisná poloha atd. Krize to všechno změnila: sociální status začal hrát důležitou roli. Pravděpodobně nejdůležitější bylo to, že krize přivedla k volebním urnám příslušníky nižších sociálně-ekonomických vrstev, kteří předtím příliš často nehlasovali. Můžeme tedy říci, že v letech 1920–32 předpovědi Literary Digestu vyšly jenom náhodou. Abychom byli schopni z chování vzorku předpovídat chování populace, **musí struktura vzorku imitovat složení populace tak přesně, jak je to jen možné**.

Dr. Watson:

Ale to je přece docela lehké! Když je v populaci řekněme 51 % žen, tak vyberu také 51 % žen do vzorku, a když je v populaci 12 % osob nad 65 let věku, vyberu také stejné procento starých osob do vzorku atd.

Tentokrát má Dr. Watson pravdu. Technika konstrukce vzorku, tak jak ji popsal, se opravdu používá. Říká se tomu **kvótní výběr**.

Kvótní výběr imituje ve struktuře vzorku známé vlastnosti populace.

Bohužel má tato technika některé nepříjemné vlastnosti. Jedna z nich souvisí se slovem „známé" v naší definici. Pro většinu populací není problém zjistit jejich skladbu podle

pohlaví, věku, vzdělání, povolání atd. Lze si snadno představit problém, pro který jsou důležitější jiné vlastnosti, takové, o kterých běžná statistická šetření údaje neshromažďují (kupř. věk, ve kterém se respondent poprvé zamiloval).

Na další problém snadno přijdete sami:

Cvičení 5.2

Navrhněte prosím kritéria pro konstrukci kvótního vzorku pro populaci veksláků.

Kvótní výběr může být použit jen na populaci, o které jsme dobře informováni, a to zdaleka není každá populace. Další obtíž je spojena s praktickou stránkou výběru přímo v terénu. Poslední krok obvykle závisí na tazateli, který vybírá jedince podle dané instrukce. Taková instrukce by mohla vypadat třeba takto:

Jméno tazatele: *Dr. Watson*

Respondent č. 1

muž, věk 30–40, dokončené středoškolské vzdělání, povoláním úředník, ženatý, ale bezdětný, bydlící v rodinném domku, žijící v našem městě alespoň 5 let, ale který se narodil v obci pod pět set obyvatel...

Respondent č. 2

žena, věk 60–65, alespoň s dokončeným základním vzděláním, důchodkyně, která, pokud byla ještě ekonomicky aktivní, měla dělnické povolání, která žije sama, v bytě alespoň o dvou místnostech a bydlí od narození v našem městě...

Tak to si od nás Dr. Watson opravdu nezaslouží. Umíte si představit, na kolik dveří by musel zaklepat, než by našel osoby odpovídající zmíněným charakteristikám? Třeba by je nenašel vůbec, možná že vůbec neexistují. Ve skutečnosti je instrukce v kvótním výběru mnohem skromnější. Navrhuje jen několik málo proměnných, takových jako pohlaví, věk a povolání. Lokalita a typ obce jsou obvykle dány působištěm tazatele. Jinak nejsou tyto proměnné vázány do určitých kombinací. Instrukce by mohla znít takto: „Hovořte s deseti

osobami, z toho se šesti ženami a čtyřmi muži. Vyberte tři osoby ve věku pod 20 let, pět ve věku 21–50." O ostatních, pro nás třeba daleko důležitějších proměnných můžeme jenom doufat, že budou ve vzorku dostatečně správně reprezentovány.

Dr. Watson:
Co si s tím ale počneme?

Odpověď nám nabízí titul následujícího paragrafu.

5.2 Hodíme si korunou aneb Pravděpodobnost pro Dr. Watsona

Představme si, že máme velikou krabici plnou kuliček a že všechny kuličky jsou zelené. Dobře krabicí zatřepeme a poslepu vybereme jednu kuličku. Jakou máme šanci, že vybraná kulička bude zelená? To byla ale pitomá otázka, že ano? Tak si teď zkusme něco trochu složitějšího: Máme teď jinou populaci kuliček, sestávající ze zelených a červených kuliček. Těch zelených je 80 % a těch červených je ovšem 20 %. Ale počkejte, já se vás nebudu ptát, jaká je pravděpodobnost, že si náhodně vyberete červenou kuličku. To by byla otázka jen o málo méně pitomá než ta první, a všichni víme, že ta pravděpodobnost je 20 %, a chceme-li to vyjádřit učeněji, můžeme říci, že p = 0,20.

My tu máme jiný úkol: zjistit, jaká je skladba populace, aniž bychom prohlíželi všechny kuličky. Jinými slovy, hledáme metodu, jak vytvořit vzorek, který by dobře reprezentoval celou populaci kuliček. Můžeme zkusit třeba toto: Opět začneme tím, že krabicí dobře zatřeseme. To není vtip, to je opravdu nutné: **každá kulička musí mít stejnou pravděpodobnost, že bude vybrána**. (Co kdyby všechny červené kuličky byly navrchu?) A teď vybereme poslepu 10 kuliček. Uvidíme třeba, že jsme vybrali 6 červených a 4 zelené. To je dost daleko od dobré reprezentativity. Perfektní vzorek by měl přece obsahovat 20 % červených a 80 % zelených. Tedy vybereme opět poslepu dalších deset kuliček. Třeba 6 z nich bude zelených a 4 červené. Přidáme je k našemu původnímu vzorku. Nový, větší vzorek sestává z 10 červených a 10 zelených kuliček. Teď bychom odhadli, že v populaci je stejné procento červených jako zelených kuliček. To ještě není vůbec dobré. Museli bychom tedy pokračovat, přidávat další a další kuličky. Brzy bychom zpozorovali zajímavou věc:

S rostoucí velikostí vzorku se rozdíl mezi strukturou populace a vzorku zmenšuje.

Nejdříve rychle, pak pomaleji a pomaleji. Úplné shody mezi strukturou populace dosáhneme teprve tehdy, když jsme zahrnuli všechny elementy populace do vzorku.

Dr. Watson:
„Ale to je všechno nesmysl! Když je to pravda, jak je potom možné, že obrovský vzorek použitý Literary Digestem vedl k tak nesprávným výsledkům?"

Asi už víte, co bychom mohli odpovědět na tuhle námitku: „Ale to je přece elementární, Watsone. Ti lidé z Literary Digestu zapomněli pořádně zatřást krabicí." Voliči z nižších socioekonomických vrstev měli mnohem menší šanci být vybráni do vzorku než voliči ze středních a vyšších vrstev, což dramaticky zkreslilo výsledky.

My jsme tu totiž, aniž bychom o tom věděli, vytvořili náhodný vzorek „populace" kuliček. A náhodný vzorek, to je aristokrat mezi vzorky; má mnoho jedinečných a pro nás důležitých vlastností. Všechno, co budeme v tomto odstavci probírat, se týká jenom vzorků, které byly vytvořeny opravdu náhodným výběrem. Termín „náhodný" neznamená výběr nazdařbůh. I když náhodný výběr může být, jak brzy uvidíme, technicky velmi obtížný a často i nemožný, jeho definice je jednoduchá:

Náhodný (pravděpodobnostní) výběr je takový výběr, ve kterém každý element populace má stejnou pravděpodobnost, že bude vybrán do vzorku.

To se lépe řekne, než se to udělá. Ale dovolte, abych vás ještě dříve, než budeme mluvit o řadě trampot, dobře naladil popisem pozoruhodných vlastností náhodného vzorku. Snad

nejdůležitější z nich, alespoň pro nás sociology – statistik by s námi možná nesouhlasil –, je tato vlastnost:

> Náhodný vzorek reprezentuje všechny známé i neznámé vlastnosti populace.

A ještě dříve, než Dr. Watson začne namítat, uveďme si jednoduchý příklad. Máme teď novou populaci kuliček. Jsou opět červené a zelené. Ale mají ještě jednu zajímavou vlastnost, o které my nevíme: Jsou duté a uvnitř každé je malý papírek a na každém tom lístku je něco napsáno. (Znáte „fortune cookies" z čínských restaurací?) Třeba nějaké neslušné slovo. Když jsme vybrali dobrý náhodný vzorek kuliček, budou reprezentovat celou populaci kuliček nejen vzhledem k distribuci barev, ale i vzhledem k distribuci neslušných slov, i když o tom nevíme a třeba nikdy nebudeme vědět. Uveďme si jiný, užitečnější příklad. V náhodném vzorku obyvatelstva hlavního města Prahy budeme mít slušnou reprezentaci populace vzhledem k věku, pohlaví, vzdělání, povolání, politické orientaci, vzhledem ke všem postojům, ale i reprezentaci třeba vzhledem k oblíbeným jídlům, počtu zubních kazů, věku, kdy se lidé poprvé zamilovali, množství vypitého piva, počtu milenek, počtu veksláků, peněžní hodnotě nakradeného zboží, číslům bot, prostě vzhledem ke všemu. To neznamená, že tohle všechno budeme schopni měřit, to je jiný problém. Ale znamená to, že ať už je naším cílem cokoli, víme, že **proměnné, které jsou pro nás relevantní, budou mít v našem vzorku podobnou distribuci, jaká existuje v celé populaci, a naše závěry jsou tedy na tuto populaci aplikovatelné**.

Náhodný výběr má ještě jednu pozoruhodnou vlastnost:

> U náhodného vzorku jsme schopni odhadnout, jak se vzorek liší od populace.

Jinými slovy, jsme schopni určit, jak dobrý je náš vzorek. Teď je načase naučit se několik slov z odborné hantýrky, jednak abychom mohli oslnit přátele, jednak abychom rozuměli správně významu publikovaných statistických dat. Podívejme se na následující tabulku:

Tabulka 5.1

Velikost vzorku a konfidenční interval
na 95% hladině významnosti pro alternativní znaky při distribuci 50:50

Velikost vzorku	Konfidenční interval
100	±10 %
400	±5 %
1600	±2,5 %

Adaptováno z Babbie, *Social Research for Consumer*, 1982

To vypadá dost učeně, že? Ale nebojte se. Pochopit princip a vědět, jak se taková věc aplikuje, není těžké. Trochu obtížnější je statistické zdůvodnění. Ale takové vysvětlení necháme pro někoho jiného, kdo vás uvede do zajímavého světa **skutečné** statistiky.

Řekněme, že jsme vybrali náhodně 400 kuliček a zjistili jsme, že ve vzorku (neboli ve výběrovém souboru) je 78 % zelených kuliček. Protože jsme nevybrali všechny kuličky, musíme předpokládat, že jsme se dopustili určité chyby, že pozorovaná relativní četnost zelených kuliček ve vzorku se liší od procenta, které skutečně existuje v celé populaci (základním souboru). My však potřebujeme vědět, jak moc se mýlíme. A v tom nám pomůže ta nepřátelsky vyhlížející tabulka. Pozor! **Tahle tabulka je jen ilustrací a platí jen tehdy, je-li v populaci právě tolik zelených jako červených kuliček. Platí jen pro alternativní (binomické) proměnné**, to je pro takové znaky, které mají jen dvě kategorie, jako ANO a NE. V našem případě zelená a „nezelená" kulička.

Velikost našeho vzorku je 400 a této velikosti vzorku odpovídá konfidenční interval (interval spolehlivosti) 5 %. Odečteme tedy tuto hodnotu od pozorovaných 78 % a dostaneme tedy 73 %. Pak ji opět přičteme k pozorované hodnotě a dostaneme horní mez. A teď víme, že skutečná proporce zelených kuliček v celé populaci je mezi 73 a 83 %. Jenomže to nevíme docela určitě, vždyť jsme nepozorovali všechny kuličky. Teď se dostáváme k tomu poněkud kryptickému výrazu v podtitulu naší tabulky: **hladina významnosti**.

V našem případě to znamená, že skutečná proporce, která existuje v populaci, se nalézá s 95% pravděpodobností uvnitř vypočítaného intervalu spolehlivosti. Kdybychom vytvořili 100 vzorků obdobné velikosti, jen v 5 vzorcích by bylo možné, že skutečná proporce zelených kuliček leží pod nebo nad vypočítaným konfidenčním intervalem. O tom, jakou hladinu zvolit, rozhodne výzkumník, a podle tohoto rozhodnutí je interval vypočítáván. Toto rozhodnutí je svobodné ovšem jen z hlediska statistické teorie; ve skutečnosti je vázán míněním přijatým v příslušné vědecké komunitě. V sociologii je to obvykle 95 nebo 98 %. (Vidíte, i v sociologii máme malý kousek paradigmatu.)

A teď se podívejme, jak by se takový interval **mohl** vypočítat. Není to tak, jak se to opravdu dělá. Ve skutečnosti neznáme distribuci proměnné, která existuje v populaci. Ale náš popis výpočtu nám dá alespoň nějaký vhled do logiky, která je skryta za pozoruhodnými vlastnostmi náhodného výběru. Protože jsem vám slíbil, že v naší knize nebudou (skoro) žádné vzorečky, popíšeme si výpočet slovně. Nejdříve musíme vypočítat veličinu, která má opravdu zajímavé vlastnosti a které se říká **směrodatná chyba**. Uvidíte, že je to nejen snadné vypočítat, ale také že není těžké rozumět většině kroků v tomto výpočtu.

Výpočet směrodatné chyby:

CO UDĚLÁME	CO TO ZNAMENÁ
Nejdříve vynásobíme proporci zelených kuliček v populaci proporcí červených. Tato proporce musí být vyjádřena jako desetinný vzorek, ne v procentech. (Tedy kdyby v populaci bylo 50 % červených a 50 % zelených, budeme počítat 0.5krát 0.5.)	Homogenita vzorku má vliv na velikost chyby. Čím nerovnoměrnější je distribuce ve vzorku, tím menší bude chyba a tím užší bude interval spolehlivosti. Kdyby například v populaci bylo 90 % zelených kuliček a velikost vzorku by byla 100, vypočítaný konfidenční interval by byl ±6 %. Kdyby ve stejně velkém vzorku byl stejný počet zelených jako červených kuliček, konfidenční interval by byl mnohem širší: ±10 %.

Vypočítaný násobek vydělíme velikostí vzorku.	Čím větší vzorek, tím menší je směrodatná chyba a tím užší bude konfidenční interval. V případě, že by v populaci byla stejná proporce zelených a červených kuliček, ve vzorku 100 pozorování by interval byl ±10 %; ve vzorku 400 pozorování by byl mnohem užší: ±5 % a ve vzorku 1000: ±3 %.
Nakonec vypočítáme druhou odmocninu z výsledku dělení.	To je transformace do čísla zajímavých vlastností. Ti, kdo jsou trošku seznámeni se statistikou, vidí už teď souvislost s konceptem směrodatné odchylky. My ostatní to pochopíme trochu lépe, až budeme mluvit o směrodatné odchylce v naší statistické kapitole.

A teď nám už zbývá jen jedno. Rozhodnout se, jakou hladinu významnosti chceme přijmout, a pak vypočítat interval spolehlivosti.

Směrodatná chyba má jednu pozoruhodnou vlastnost: do intervalu vymezeného ±1 standardní chybou od hodnoty pozorované ve vzorku připadne správná hodnota existující v populaci přibližně v 68 případech ze sta. Tak bychom dostali interval spolehlivosti na 68% hladině významnosti. To ovšem není zdaleka dost vysoká pravděpodobnost. Abychom vypočítali interval spolehlivosti na úrovni, jaká je vyžadována v našem oboru, musíme přičíst a odečíst směrodatnou chybu dvakrát. Jinými slovy: **interval spolehlivosti na 95% hladině významnosti je dán rozmezím ±2 směrodatné chyby od hodnoty naměřené ve vzorku**. Rozmezí ±3 směrodatné chyby nám definuje ještě mnohem striktnější interval na hladině 99.9 %. Ten je užíván zejména v přírodních vědách.

A teď už víme dost, abychom mohli představit další, opravdu překvapivou vlastnost náhodného výběru:

> Velikost směrodatné chyby, a tedy i konfidenční interval (interval spolehlivosti) nezávisí vůbec na velikosti populace.

Jedině velikost vzorku a jeho homogenity ovlivňují velikost chyby.

Dr. Watson:

Počkejte, počkejte! Chcete mi namluvit, že řekněme vzorek 300 respondentů vykáže stejnou chybu, když reprezentuje populaci továrny s 800 dělníků, jako stejně velký vzorek, který reprezentuje město s 50 000 obyvatel, nebo dokonce zemi s 200 000 000 občanů? Já tomu prostě nevěřím!

Neuvěřitelné, a přece je to pravda, pokud ovšem distribuce zkoumané proměnné je ve všech těch populacích stejně homogenní. A pokud mi ještě nevěříte, podívejte se znovu na popis výpočtu směrodatné chyby. Najdete tam zmíněnou proporci zelených a červených kuliček, velikost vzorku a to je vše. Ani zmínka o populaci.

To, co víme, by nám mohlo dát dostatečnou informaci, abychom mohli **navrhnout velikost vzorku, jakou potřebujeme vzhledem k velikosti chyby, jakou jsme ochotni riskovat.** V praxi to však není snadné: pro výpočet směrodatné chyby potřebujeme znát homogenitu populace vzhledem k našim proměnným, rozptyl těchto proměnných. Většinou tuto znalost nemáme. Existují sice techniky, které nám umožní tuto informaci odhadnout, ale tyto techniky jsou buďto nákladné, nebo nepřesné.

A tak v tvrdé praxi denního života výzkumníka spoléháme na zkušenost a na zdravý rozum. Můžeme se třeba zamyslit nad tím, které kombinace proměnných jsou pro nás nejdůležitější. Představíme si, kolik polí bude mít tabulka (nebo tabulky), a navrhneme, kolik pozorování musí každé pole v těchto tabulkách obsahovat – prázdná pole nebo pole s málo pozorováními mohou podstatně zkreslit výsledky statistické analýzy. Zaměřme se raději na dost vysoké minimum; někdy navrhovaný průměr 10 pozorování na jedno pole tabulek může být nezdravě optimistický. Data ve skutečnosti nebudou do všech polí rozdělena rovnoměrně; některá pole budou přeplněna a jiná téměř prázdná. Nadto v každém výzkumu máme mnoho proměnných, s různým počtem kategorií, někdy nevíme předem, které kombinace proměnných přinesou nějaké zajímavé výsledky, a tak si zaslouží hlubší analýzy atd. Zkrátka teoretizování o velikosti vzorku patří spíše na stránky učebnic než do praxe sociologického výzkumu. Tam aplikujeme následující, velice nevědecké, ale velice praktické pravidlo: **Snažme se vytvořit**

co největší vzorek, jaký nám naše časové a finanční podmínky dovolují; ne však za cenu vážného narušení pravidel náhodného výběru. Doba pro aplikaci naší znalosti o intervalech spolehlivosti přichází v praxi teprve v etapě statistické analýzy sebraných dat. Pak je to ovšem velice důležité.

A teď ještě jedno důležité varování:

> Velikost směrodatné chyby se týká jen zkreslení vyvolaného rozdíly mezi vzorkem a populací. Nevztahuje se, bohužel, na zkreslení vyvolané jinými typy redukce a transformace informací. Tato zkreslení jsou pro nás většinou mnohem nebezpečnější a my nemáme žádný nástroj, jak měřit velikost těchto omylů.

5.3 Jak správně házet korunou

Dr. Watson:

Já už vidím, že náhodný výběr je výborný. Hned to začnu používat. Vždycky jsem chtěl vědět, co si lidé v Praze myslí o mojí politické straně. Hned začnu pracovat na hypotézách a otázkách pro rozhovor. Od pondělí budu každé dopoledne na Václaváku a budu se vyptávat náhodně vybrané osoby...

Pokud náš pošetilý přítel doufá, že jeho výsledky budou reprezentovat mínění pražské populace, je ještě mnohem pošetilejší, než jsme si mysleli. Víme přece, že při náhodném výběru každý člen populace musí mít stejnou pravděpodobnost, že bude vybrán. Watsonův vzorek by byl silně zkreslený.

Cvičení 5.3

Navrhněte prosím, jak by se Watsonův vzorek lišil od pražské populace.

Tedy jasně vidíme, že tento vzorek by snad mohl být reprezentativní pro populaci definovanou asi takto: osoby, které se nacházejí na Václaváku ve všední den dopoledne, v dané roční době. Pro nějaké speciální účely by mohla být taková populace zajímavá: kupř. pro

plánování obchodních strategií pro obchody na Václaváku, rozhodně však ne pro problémy spojené s politickou orientací obyvatel. Ale i tak by byla **náhodnost**, a tedy i reprezentativnost takového výběru problematická. Dr. Watson, protože je v podstatě konzervativní, by se mohl ostýchat oslovit méně konvenčně oblečené osoby. Kdyby takový výběr prováděl můj syn, půvabné mladé ženy by byly ve vzorku přereprezentovány. Kdybych prováděl výběr já, pak by byly podreprezentovány, protože jsem stydlivý. Ono se vůbec zdá, že lidská mysl není schopna pracovat opravdu náhodně.

Můžeme si to dost snadno vyzkoušet. Požádejte větší skupinu lidí – třeba třídu studentů –, aby každý napsal na kousek papíru jakékoliv číslo mezi 1 a 10. Bez dlouhého přemýšlení musí napsat to, co jim přijde na mysl. Je-li skupina dost velká, je vysoká pravděpodobnost, že číslo 7 bude mít daleko nejvyšší frekvenci. Proč, to nevím, a předem můžete zavrhnout teorii vlivu sedmy v naší mariášnické kultuře; v Kanadě to funguje také, a jak! Snad to má něco dělat s tradiční mystikou čísel, ale v každém případě to krásně dokumentuje, že náš mozek je velice špatným generátorem náhodnosti. Musíme jej nahradit něčím neosobním. Hodit si korunou? Zatřepat krabicí?

Pomůcky, které v praxi při výběru náhodného vzorku používáme, skutečně imitují takové mechanismy. Mohli bychom třeba napsat jména všech členů populace na papírky, dát do klobouku, kloboukem pořádně zatřepat a pak poslepu vytáhnout tolik papírků, kolik osob potřebujeme do vzorku. Ovšem většinou by to muselo být pěkně velký klobouk a v každém případě je to dost nepohodlný postup. Můžeme jej však dobře imitovat. Prostě jednotlivce v seznamu populace očíslujeme a pak použijeme „něco", co produkuje náhodná čísla, a vybíráme ty jedince, jejichž číslo se s těmi náhodnými shoduje. Říká se tomu

prostý náhodný výběr

Jednoduchá však v tom není generace těch náhodných čísel. Kdysi se k tomu užívala taková podivná „kostka", mnohohran s deseti stejnými plochami, na každé z nich byla jedna z číslic od 0 do 9. Prý bylo obtížné vyrobit takovou „kostku", aby byla „poctivá", to je, aby každá číslice měla stejnou pravděpodobnost, že „padne". Ještě donedávna jsme používali tabulky náhodných čísel, dost tlusté knihy číselných skupin, o nichž nám matematici řekli, že v nich za takových a takových okolností nebyli schopni objevit žádnou pravidelnost. Dodnes jsou výtahy z těchto tabulek přetiskovány téměř v každé učebnici výzkumných metod. Jejich správné

používání rozhodně není nejzábavnější kratochvíle, ale někdy nám prostě nezbude nic jiného. Naštěstí dnes každá lepší kalkulačka a ovšem každý, i nejmenší osobní počítač umí produkovat náhodná (matematik by řekl „kvazináhodná") čísla. Tenhle přístup má velikou výhodu: program produkuje náhodná čísla jenom v tom rozsahu, v jakém je potřebujeme. Řekneme počítači, jak je velká populace, třeba 300, a program pro nás vyprodukuje náhodná čísla jenom v rozsahu od 1 do 300. Tabulky náhodných čísel jsou nejméně pěticiferné. Pro naši velikost populace použijeme ovšem jen první nebo poslední tři sloupce číslic, ale i tak sedm z deseti nalezených nebudeme s to použít. Kalkulačka nebo počítač jsou mnohem efektivnější, a když si s tím nevíte rady, obraťte se na sousedova syna, a pokud by neměl takový program, většina těch chytrých holek a kluků, kteří vlastní třeba i ten nejmenší Sinclair, je schopna napsat takový program v Basicu za několik minut.

Dr. Watson:

Ale já nemám kalkulačku a všichni sousedi jsou bezdětní. Tak bych si to chtěl zjednodušit. Populace má 500 členů a já chci vzorek ve velikosti 100. Proč bych nemohl vzít jednoduše každou pátou osobu ze seznamu?

Tentokrát Dr. Watson promluvil pro změnu moudře. Technika, kterou navrhl, se opravdu používá. Říká se jí systematický výběr. Nenechte se však zmást tím názvem; je to opět technika náhodného výběru.

Systematický výběr

V systematickém výběru je do vzorku zahrnuta každá N-tá jednotka ze seznamu. Velikost kroku (N) dostaneme, když vydělíme velikost populace velikostí požadovaného vzorku. Důležité však je, aby první jedinec byl vybrán náhodně a teprve od tohoto výchozího bodu budeme vybírat každou N-tou jednotku.

Tento postup však nemůžeme použít, když jsou seznamy řazeny podle nějakého systematického schématu. Naše pohádka ilustruje něco, co se v praxi opravdu stává.

Pohádka pro odrostlejší děti 9

O výběru, který byl příliš systematický

Bylo, nebylo, kdesi existovalo malé království, které se jmenovalo Org. Bylo to království, kde všechno bylo velice dobře zorganizováno, a přesto byl každý šťastný a spokojený. Každý, až na vojáky základní služby. Ti si stěžovali na plat, na stravu, na zacházení od představených, na všechno, a protože vše bylo dobře zorganizováno, vláda pozvala zahraničního odborníka, profesora P. I. Tomu, aby provedl výzkum postojů v armádě.

P. I. Toma přijel, zkonstruoval výborný dotazník a vyzkoušel jeho validitu. Protože to království bylo tak malé, že se tam ani počítač nevešel a místní knihovny neměly tabulku náhodných čísel, rozhodl se použít pro konstrukci vzorku techniku systematického výběru. Armáda toho malého království byla taky malá, důstojníci, poddůstojníci i mužstvo dohromady jen 12 000 osob. Profesor P. I. Toma odhadl, že vzorek 200 osob mu poskytne přijatelný interval spolehlivosti a zvolil tedy krok 60. Náhodně vybral prvého jedince. Byla to osoba č. 31 a pak vybíral každého dalšího šedesátého vojáka. Výsledky výzkumu byly prostě náramné. Ještě nikdo nikde nezkoumal tak spokojenou armádu. Každý byl šťastný v tom malém šťastném království – až do příštího jara, kdy začalo krvavé povstání vojáků základní služby.

Ale vy už víte, co se stalo: Prostě v království Org bylo vše dobře organizováno. I seznamy členů armády byly uspořádány po četách, v každé četě nejdříve dva důstojníci, pak tři poddůstojníci, pak mužstvo základní služby a každá četa měla ne více, ne méně než 30 osob. I náš profesor měl smůlu, protože zvolený krok se shodoval přesně nejen s dvojnásobkem velikosti čety, ale také proto, že první náhodně vybraná osoba byl důstojník a tedy každá následující osoba musela být také důstojník. Poddůstojníci a vojáci základní služby nebyli zahrnuti do vzorku vůbec.

Nemysleme si, že takové zkreslení patří jen do absurdního světa pošetilých pohádek. Mnohé ze seznamů populací jsou systematicky uspořádány, kupř. žáci škol podle tříd, dělníci podle dílen atd. Někdy systém, podle kterého je seznam organizován, nemusí být na prvý pohled zřejmý. Kupř. byty na sídlištích ve velkých obytných budovách bývají identifikovány třícifernými čísly. Prvá číslice definuje podlaží, druhé dvě byt na podlaží. Protože půdorys se na každém podlaží opakuje, byty se stejnými posledními číslicemi budou mít obdobné vlastnosti, budou třeba větší či menší než byty ostatní, a to by opět při systematickém výběru mohlo produkovat zkreslení.

Podívejme se teď na jiný typ náhodného výběru, který by býval mohl zachránit profesora P. I. Tomu před zmíněnou blamáží:

Náhodný stratifikovaný výběr: Populace je rozdělena do skupin homogenních vzhledem k nějakému jasnému kritériu a jedinci jsou vybíráni do vzorku náhodně z těchto skupin.

Profesor Toma měl začít s třemi seznamy; se seznamem populace důstojníků, s jiným, zahrnujícím jen poddůstojníky, a konečně se seznamem vojáků základní služby. Z každé populace by pak byl vybrán náhodný vzorek, třeba technikou systematického výběru, a v našem malém království by k povstání třeba nedošlo. Ve skutečném světě, například při výzkumu studentů určité školy, bychom vybírali jedince zvlášť pro každý ročník. Při jiných výzkumech by populace mohla být stratifikována podle volebních obvodů, při výzkumu zaměstnanců továrny by mohl být výběr prováděn zvlášť mezi dělníky a zvlášť pro administrativu.

Stratifikovaný náhodný výběr má ještě jednu dodatečnou výhodu; snižuje velikost směrodatné chyby, a tedy i interval spolehlivosti. Třeba si ještě pamatujete, že chyba klesá s rostoucí velikostí vzorku a s přirůstající homogenností populace. Logika toho je zřejmá: když v populaci je pro kandidáta A 98 % voličů a pro kandidáta B jen 2 %, předpovědět, kdo vyhraje volby, je mnohem snadnější, než kdyby preference bylo třeba 55 % pro A a 45 % pro B. Ve stratifikovaném výběru jsou vzorky podskupin zcela homogenní vzhledem k proměnné, podle které byly stratifikovány: ve skupině jsou jenom vojáci základní služby nebo jenom posluchači druhého ročníku atd. Pro stratifikační proměnnou je tedy směrodatná chyba nulová a pro všechny jiné proměnné, které jsou s touto proměnnou asociovány, bude tato chyba podstatně menší.

A teď se podíváme na velmi zvláštní typ výběru, na vícestupňový náhodný výběr. Je to technika velice pracná, náročná a drahá, ale, jak hned uvidíme, velice důležitá a nenahraditelná.

Vícestupňový náhodný výběr

se provádí ve dvou nebo více krocích. Nejdříve jsou náhodně vybrána určitá přirozená seskupení, a pak teprve jsou náhodně vybíráni jedinci z oněch vybraných seskupení.

K čemu je to dobré? Pro ilustraci jednoho aspektu vás pozvu na výlet na jiný kontinent. Představte si, že bychom měli dělat výzkum na náhodném vzorku reprezentujícím dospělé obyvatelstvo Kanady. Kanada má něco přes dvacet milionů obyvatel, ale její plocha je téměř

10 000 000 čtverečných kilometrů. Řekněme že velikost vzorku by byla 1000 jedinců, a tak bychom teoreticky měli jednoho respondenta na deset tisíc čtverečných kilometrů. Ve skutečnosti by to bylo mnohem méně, obrovské rozlohy země jsou prázdné. Ale i tak jsou rozměry země obrovské a takové by byly i náklady. Při dané velikosti vzorku bychom měli nejmenší potíže s nejlidnatějšími provinciemi. V Quebeku bychom měli asi 290 respondentů, v Ontariu přibližně 350. Ale v Northwest Territories jednoho nebo dva a ti by nás přišli pěkně draho. Pokud bychom neměli velké štěstí, museli bychom, abychom je zastihli, najmout hydroplán, helikoptéru nebo psí spřežení. Ale i v nejlidnatějších provinciích anebo i v prostorově malé zemi s tak vysokou hustotou obyvatelstva, jako mělo Československo, rozptyl populace v prostoru podstatně zvyšuje náklady a nesmírně ztěžuje organizaci výzkumů. (Kupř. tazatelské týmy jsou organizovány a školeny lokálně; to snižuje cestovní náklady. Ale je jen omezený počet terénních center, která jsme schopni organizovat a financovat.) Tady je právě oblast uplatnění vícestupňového náhodného výběru. Můžeme postupovat třeba takto:

1. Nejdříve vybereme náhodně reprezentativní soubor okresů.
2. Pak v každém z vybraných okresů provedeme náhodný výběr obcí.
3. Ve velkých vybraných obcích zařadíme ještě další mezistupeň výběru: vybereme náhodně menší prostorové jednotky, třeba volební obvody.
4. Teprve pak vybíráme jedince.

Tímto způsobem obdržíme mnohem kompaktnější vzorek. Respondenti nejsou rozptýleni po celém teritoriu, ale jsou koncentrováni do zvládnutelného počtu regionů. Je-li takový výběr proveden správně, žádné závažné zkreslení reprezentativnosti nehrozí.

Nicméně existuje ještě jedna, dokonce důležitější doména použití tohoto výběru. Největším problémem pro použití pravděpodobnostního výběru v sociologii je fakt, že pro mnoho zajímavých populací žádný seznam neexistuje. Pro mnoho těchto situací je vícestupňový náhodný výběr jediným řešením. Řekněme, že bychom chtěli vytvořit pravděpodobnostní vzorek celé země, a žádné spolehlivé seznamy obyvatelstva buď neexistují, nebo nejsou výzkumníkovi dostupné. To je mimochodem situace ve většině zemí světa.

Postup by byl shodný v prvých třech krocích s předchozí tabulkou, ale pak by následovaly dva další, logicky jednoduché, ale pracovně náročné kroky:

4. Ve vybraných malých obcích nebo městských obvodech je proveden soupis všech sídelních jednotek (bytů, rodinných domků).
5. Pak je vytvořen náhodný vzorek těchto jednotek.
6. Je vytvořen seznam osob žijících ve vybraných jednotkách a pak jsou opět náhodně vybráni jedinci (nebo obvykle jedinec) do vzorku.

Nejnáročnější je ovšem krok č. 4. Představuje obsáhlou práci jak v přípravě, tak i v terénu; záznamy se obvykle opožďují za skutečností, nemusí rozlišovat mezi jednotkami, které jsou obydleny, a těmi, které jsou používány pro jiné účely atd. Poslední krok je obvykle prováděn tazatelem přímo v terénu. Náhodnost musí být zaručena i při tomto kroku. Záznamový arch pro interview obsahuje instrukci, v jakém pořadí mají být členové domácnosti zaznamenáváni, a náhodně generované pořadové číslo osoby, která má být interviewována. Bez takové instrukce by tazatel vybral osobu, která je právě dosažitelná, aby se tak vyhnul nutnosti další návštěvy, nebo osobu, která je mu sympatická. Tak by byly kupř. podreprezentovány osoby, které během dne pracují mimo dům.

Někdy aplikace vícestupňového výběru nemusí být obtížná a je přitom velice užitečná. Chtěli bychom třeba studovat na celostátním vzorku mínění studentů dvou nejvyšších ročníků střední školy. Ústřední seznam středoškolských studentů asi neexistuje, ale existuje seznam všech středních škol a každá škola má seznam žáků, sestavený pravděpodobně podle ročníků. Výběr by mohl být prováděn třeba takto: Náhodně by byly vybrány okresy, pak vzorek škol v těchto okresech a jedinci do vzorku by byli náhodně vybráni ze seznamu žáků posledních dvou ročníků.

Před časem jsme zkoumali postoje starších osob k možnosti vstoupit do institucí pro staré občany (Disman & Disman 1989). Naším cílem bylo sledovat vliv etnické kultury na tyto postoje; porovnávali jsme postoje Portugalců a Italů žijících v Torontu, ve věku 65 nebo starších, s postoji stejně starých Kanaďanů, jejichž mateřským jazykem je angličtina.

Vytvoření vzorku nebylo snadné. Osoby starší než 65 let představují 11 % torontské populace, z těchto starších osob je jen 5 % Italů a 1 % Portugalců. (To znamená, že Portugalci ve věku 65 a více představují asi 0.11 % z torontské populace.) Kdybychom tedy chtěli interviewovat

100 Italů a 100 Portugalců, museli bychom kontaktovat asi 100 000 domácností, a to je ovšem nemožné přinejmenším z finančních důvodů. Naštěstí jsme měli k dispozici seznamy osob pro daňové účely a tyto seznamy zahrnují prakticky všechny dospělé občany. Nadto tyto seznamy zahrnovaly také informaci o věku. Tato informace podstatně zúžila velikost vzorku pro vyhledávací fázi výzkumu. Ale i tak, abychom vyhledali vzorek 100 portugalských respondentů, museli bychom kontaktovat asi 10 000 domácností a i to by bylo nemožné.

Zůstala pro nás tedy otevřena jediná možnost: kontaktovat osoby ze seznamu, jejichž jména znějí italsky nebo portugalsky. Jistě, tato metoda má některé nevýhody. Kupř. portugalské jméno může mít britská manželka portugalského manžela, ale tyto případy byly vyloučeny v předběžném rozhovoru. Do vzorku nebyly zahrnuty osoby s etnicky netypickými jmény, italské nebo portugalské manželky mužů jiného etnického původu atd. Nicméně toto zkreslení – zejména vzhledem k silné tendenci obou národnostních skupin uzavírat sňatek uvnitř etnické skupiny (endogamy) – nebylo příliš vážné. Ale i tak – zejména vzhledem k úmrtnosti mezi staršími osobami, vinou nepřesností záznamů a vzhledem ke značné horizontální mobilitě – bylo nutno kontaktovat 652 portugalských adres s výtěžkem 161 jmen respondentů, odpovídajících naší definici populace.

V tomto případě jméno jako kritérium pro výběr – doufejme – nezpůsobilo vážné zkreslení. Ale nemusí tomu tak být vždycky. Mezi americkými sociology koluje hezká historka, kterou uvedeme v naší pohádce č. 10.

Pohádka pro odrostlejší děti 10

O zrádném písmenu

Bylo před místními volbami v jednom velkém místě na východním pobřeží USA a skupina politiků si objednala výzkum, předpověď výsledků voleb. V té době mělo město dobrý seznam voličů řazený abecedně. Kartotéky zaplňovaly několik místností. Pro konstrukci vzorku byla použita technika vícestupňového náhodného výběru. Nejdříve byla vybrána náhodně místnost, pak kartotéční skříň a ze zásuvek této skříně byli vybráni technikou systematického výběru jedinci do vzorku.

Výzkum skončil neslavně: jako vítěze vyhlásil kandidáta, který skončil daleko vzadu v poli poražených. Prostě výzkumník se dopustil omylu, ale zejména měl smůlu. Náhodně vybral začátek písmena M, a tak se stalo, že voliči irského a skotského původu, jejichž jména velice často začínají na Mac a Mc, byli silně přereprezentováni. Hlasování ve volbách v USA a Kanadě velmi často sleduje etnickou linii. Není proto divu, že výzkum mylně předpověděl vítězství irského kandidáta. Tomuto zkreslení bylo snadné zabránit, kdyby byl stejný počet voličů vybrán z více kartotéčních skříní. Je ale také pravda, že kdyby bylo vybráno jiné písmeno, ke zkreslení by asi nedošlo.

5.4 Když koruna nepracuje

Zatím jsme viděli členy dvou rodin výběrových technik. Nejdůležitější jsou pravděpodobnostní techniky založené na náhodném výběru. Jsou velice mocné, zajišťuji, že budou dobře reprezentovány všechny známé i neznámé vlastnosti populace. Nadto jen u nich jsme schopni prostředky statistiky odhadnout, nakolik se vzorek liší od populace. Bohužel zdaleka ne vždy jsme schopni tyto techniky použít. Někdy třeba proto, že pracnost a nákladnost těchto technik přesahuje rámec našich možností. Jindy proto, že neexistuje žádný seznam cílové populace. Nejčastější překážkou je však kombinace obou těchto důvodů. Speciální populace, o kterou se zajímáme, může být rozptýlena mezi celou populací a mít velice nízkou frekvenci. Teoreticky by bylo jistě možné vytvořit veliký vzorek celé populace a pak, po předběžných rozhovorech, vybrat jen ty jedince, kteří odpovídají definici naší cílové populace. Jak jsme si ilustrovali na příkladu výběru starých Portugalců, z hlediska nákladů by to bylo prostě nemožné. My jsme měli štěstí, byli jsme schopni zimprovizovat seznam populace, ale to se stává spíše výjimečně.

Jako prvou techniku tvorby vzorku jsme v této kapitole diskutovali kvótní výběr. Reprezentuje druhou skupinu výběrových technik, které nejsou založeny na teorii pravděpodobnosti, ale na logickém úsudku. Kvótní výběr je pravděpodobně nejspolehlivější mezi těmito technikami,

ale opět ne vždy je možno jej použít. Může být aplikován jen tehdy, když máme dostatečnou znalost o populaci, abychom její strukturu mohli imitovat ve struktuře vzorku. Do téže skupiny patří účelový výběr, někdy i anketa, a bývá sem zařazována i technika sněhové koule (snowball sampling).

Účelový výběr
je založen pouze na úsudku výzkumníka o tom, co by mělo být pozorováno, a o tom, co je možné pozorovat.

Jak vidíte, není to příliš vědecký přístup, ale velice často jediný, který nám zbývá. Je používán i profesionálními agenturami, které provádějí za úplatu výzkum trhu. Řekněme že byste při sobotním nákupu v Bílé labuti byli osloveni mladým mužem a dotazováni na to, co si myslíte o určité skupině výrobků. Na jakou populaci se výsledky takového výzkumu vztahují?

Dr. Watson:
Ale to je přece jednoduché! Na lidi, kteří nakupují v obchodních domech!

Jako obvykle, Dr. Watson je příliš optimistický. Takto konstruovaný vzorek by reprezentoval přinejlepším populaci osob, které nakupují v Bílé labuti v sobotu dopoledne a právě v této roční době. A kdybychom měli být opravdu přesní, museli bychom ještě dodat, že se závěry vztahují jen na ty jedince z takto definované populace, kteří jsou ochotni odpovídat na otázky daného typu. Není to příliš široká a dobře definovaná populace, ale pro účely výzkumu trhu by mohly být takto získané informace určitě užitečné.

Účelový výběr nám téměř nikdy neumožní nějakou opravdu širokou generalizaci našich závěrů, ale to neznamená, že tyto závěry nejsou užitečné. Jen nesmíme předstírat jiným, a především ne sobě, že tyto závěry platí pro každého jedince ve vesmíru.

Při použití účelového výběru musí výzkumník jasně, přesně a otevřeně definovat populaci, kterou jeho vzorek **opravdu** reprezentuje.

Užití účelového výběru je pro některé populace jediným řešením. To platí kupř. pro etnické minority; snad v žádné zemi neexistují spolehlivé a vyčerpávající seznamy takových skupin. Pak nezbývá nic jiného než použít jako výchozí bod seznamy členů etnických organizací. Takový vzorek jistě nebude reprezentovat ty příslušníky etnické skupiny, kteří nejsou v žádné z takových skupin organizováni. Je pak na výzkumníkovi, aby posoudil, s použitím znalosti skupiny, jak dalece jsou jeho závěry zobecnitelné. Kupř. Pejovič (1990) zkoumal vzdělávací aspirace středoškolských studentů chorvatského původu, žijících v Torontu. Jeho závěry jsou velice zajímavé a závažné. Zdá se, že pro tuto skupinu neplatí obvyklé socioekonomické determinanty aspirací, které americká sociologie má tendenci považovat za univerzální. Pejovič užil techniku účelového výběru. Východiskem bylo členstvo různých chorvatských kulturních a sociálních organizací, účastníci různých chorvatských společenských akcí atd. Pejovič nikde nepředstírá, že jeho závěry platí mimo jeho vzorek. Nicméně síla zjištěných souvislostí a známá fakta o kulturní a socioekonomické homogenitě této etnické skupiny naznačují, že je silně pravděpodobné, že podobné výsledky bychom mohli dostat i pro většinu jiných mladých Chorvatů žijících ve velkých kanadských městech. **Důkaz** pro to by ovšem mohl být získán jen opakováním výzkumu na vzorcích vytvořených z dalších populací. Tedy i technicky vzato nereprezentativní vzorek může někdy poskytnout hodnotné výsledky. Ne však vždycky a ne automaticky a musíme si být vědomi toho, že je jen náhražkou za pravděpodobnostní výběr.

Některé techniky vytváření účelového vzorku jsou velice problematické. Bohužel stále je hojně používána anketa.

V anketě je výběr jedinců založen na rozhodnutí respondenta zodpovědět otázky uveřejněné v masových sdělovacích prostředcích.

Definovat populaci, ke které se nálezy ankety vztahují, je skutečně nemožné. Nejsou to čtenáři určitých novin nebo časopisů. To by bylo ještě dobré. Vzorek se však liší od celé populace právě tím, že to jsou ti, kteří zodpověděli anketu. Maximálně můžeme říci, že lidé ve vzorku jsou více motivováni než ostatní čtenáři, a to je velice slabá definice. Ani velikost vzorku

nepomůže. Správně konstatuje Zich (1976, s. 207), že anketa Rudého práva, která získala vzorek větší než 110 tisíc, není reprezentativní, i když v základních demografických ukazatelích se dosti shodovala se strukturou celé dospělé populace. Problém je v samovýběru respondentů. Ale to už známe z našeho příkladu vojáků, filmu a postojů k USA. Poznávací hodnota ankety je podle mého názoru pod hodnotou dobře a zodpovědně napsaného fejetonu.

A konečně tu máme techniku „snowball sampling", techniku sněhové koule. Podle mého názoru tato technika vůbec do této kapitoly nepatří. Je to technika identifikace populace, a ne vytvoření reprezentativního vzorku. Ale všechny učebnice, které znám, ji zařazují mezi výběrové techniky, a tedy i my sledujeme toto schéma. Ale posuďte to sami.

„Snowball technique" spočívá na výběru jedinců, při kterém nás nějaký původní informátor vede k jiným členům naší cílové skupiny.

Nejlépe si to ukážeme na jednoduchém příkladu. Třeba bychom chtěli studovat mocenskou strukturu v malé obci. Identifikovat oficiální vlivné osoby, jejichž pozice je formálně definována, by nebylo těžké. U nás by to byl třeba starosta, v nedávné minulosti straničtí funkcionáři, SNB atd. Ale vliv v obci mohou mít osoby, jejichž vliv není definován funkcí, a tato část souboru vlivných osob se liší podle místních okolností. V některé obci může být vlivnou osobou ředitel školy či továrny, v jiné obci mohou být osoby v takových funkcích bez vlivu a významný vliv na rozhodování může mít kněz nebo vlivný a mocný rodák, který v obci již dlouhá léta nežije. To vše je pro výzkumníka, který přichází z venku, neviditelné. Tady je na místě uplatnit výběr technikou sněhové koule.

Výzkumník začne rozhovorem s jasně definovanou osobou, třeba starostou. V tomto rozhovoru požádá respondenta, aby jmenoval další vlivné osoby. Ty jsou pak interviewovány a každá z nich dostane i stejnou otázku o vlivných lidech. Po určitém počtu rozhovorů se již jména nových vlivných osob neobjevují. Výzkumník může prohlásit, že vzorek je „teoreticky nasycen". Populace vlivných osob v obci byla jasně identifikována a náš vzorek je totožný s touto populací.

Technika sněhové koule, kde jména dalších osob se v řetězci rozhovorů „nabalují" jako sněhová koule (taková, kterou je znázorňována lavina v kreslených vtipech), je nenahradi-

telným nástrojem pro zkoumání populací, které existovaly jen dočasně: účastníci určitých demonstrací, svědkové katastrofy nebo jiné řídké události atd. Zde většinou teoretické nasycenosti vzorku nedosáhneme a aplikace této techniky má opravdu charakter konstrukce účelového vzorku.

Termín „teoretická nasycenost" byl uveden Glaserem a Strausem (1967) v souvislosti s jejich konceptem „grounded theory", snad nejdůležitějším epistemologickým nástrojem pro kvalitativní výzkum. Technika sněhové koule hraje pod jménem „teoretický výběr" velice důležitou úlohu. Má zde do jisté míry funkci ověřování validity. Ale k tomu se ještě vrátíme s celou řadou podrobností. Doufám, že to bude docela zajímavé.

5.5 Koruna přece jen není všechno

Techniky náhodného výběru opravdu produkují nejlepší možnou reprezentaci populace. Jenomže je reprezentativní pouze za předpokladu, že všichni vybraní jedinci se opravdu na výzkumu zúčastnili, to jest, že například zodpověděli naše otázky. Výpočet směrodatné chyby je plně založen na tomto očekávání. V příští kapitole uvidíme, že je to příliš optimistický předpoklad. V současné době procento osob, které odmítly tazatele nebo nevrátily dotazník, téměř všude roste. U dotazníku návratnost často nedosáhne ani padesáti procent.

Dr. Watson:

Ale to přece vůbec není problém. Já vím, že nám v jedné z dalších kapitol řeknete, že dotazník je jednou z nejlevnějších technik sběru dat. Tak když potřebuji ve vzorku 300 jedinců, prostě rozešlu 900 dotazníků, a tak dostanu vzorek i větší, než skutečně potřebuji.

Jistě už víte, proč by tento recept nefungoval: populace, která odpověděla, není totožná s tou, která odmítla odpovědět. Liší se v něčem, co bylo důvodem pro toto rozhodnutí, a pravděpodobně ono „něco" je silně spojeno s problémy, na které je výzkum zaměřen. Obvykle jsme o těchto důvodech schopni jenom spekulovat. Obávám se, že tu musím uvést nový typ nepříjemné redukce informací:

Redukce negativním samovýběrem vzniká tehdy, když část jedinců, vybraných do vzorku, odmítla na výzkumu participovat. Tento typ redukce může vážně ohrozit reprezentativnost vzorku.

Toto je vážný problém. Tak vážný, že před několika lety byl ústředním tématem výročního zasedání Americké statistické společnosti. Vidíte, na začátku této kapitoly jsme si pochvalovali, že redukce populace na vzorek je logicky, technicky a metodologicky dobře propracovanou operací, kde riziko zkreslení je menší než v jiných výzkumných operacích. Je to stále pravda, ale přece i zde máme zranitelné místo. Neznáme žádný univerzální lék na tento neduh. Jediné řešení je usilovat o co nejvyšší návratnost. U některých technik sběru informací je to snadnější, u některých je to téměř nemožné. Ale tohle už patří do příští kapitoly.

Řešení úkolů z kapitoly 5

Cvičení 5.1

Jistěže ne. Lidé, kteří neměli telefon (a těch bylo v roce 1936 velice mnoho), ti, kteří nevlastnili auto, nebyli členy organizací, tedy lidé náležející do nižších socioekonomických vrstev byli ze vzorku opravdu vyloučeni.

Cvičení 5.2

Tohle nebyla poctivá otázka. Kvótní výběr může být aplikován jen na populaci, jejíž vlastnosti relativně dobře známe. V našem případě bychom mohli nanejvýše navrhnout něco o taxikářích, vrátných v hotelech a prý i příslušnících bezpečnostních orgánů, ale rozhodně by to nebylo dost pro konstrukci kvótního vzorku.

6

Jak se to vlastně dělá

technické kroky ve výzkumu
pilotní studie
předvýzkum
základní techniky sběru informací
rozhovor
dotazník
pozorování
analýza dokumentu
jev
indikátor

interference se zkoumaným objektem
použitelnost technik v experimentu

standardizované techniky
otevřené a uzavřené otázky
kategorizace
standardizované pozorování
zkreslení pozorovatelem
měření jako zdroj změny
stereotypní odpovědi

Kapitola 6

Jak se to vlastně dělá

6.1 Co dřív?

A teď už je načase se podívat na typický jízdní řád průměrného výzkumu. V kvantitativním výzkumu se obvykle setkáme s těmito kroky:

1. Formulace teoretického nebo praktického sociálního problému.
2. Formulace teoretické hypotézy.
3. Formulace souboru pracovních hypotéz.
4. Rozhodnutí o populaci a vzorku.
5. Pilotní studie.
6. Rozhodnutí o technice sběru informací.
7. Konstrukce nástrojů pro tento sběr.
8. Předvýzkum.
9. Sběr dat.
10. Analýza dat.
11. Interpretace, závěry, teoretické zobecnění.

První čtyři kroky už známe a z prvých tří kapitol víme už dost pro to, abychom učinili poslední krok – interpretaci a zobecňování dat. Podívejme se teď na pátý krok. Definice pilotní studie a předvýzkumu se u různých autorů liší. My zde nabízíme definice používané ve většině americké literatury, i když jsou odchylné od definic užívaných u nás. Účelem pilotní studie je zjistit, zda je náš výzkum v dané populaci vůbec možný. Naproti tomu účelem předvýzkumu (bod 8 v našem schématu) je odzkoušení nástrojů (kupř. dotazníku), které jsme pro náš výzkum zkonstruovali.

Podívejme se teď velice krátce na pilotní studii:

Pilotní studie
je prováděna na malé skupině vybrané z populace, kterou hodláme studovat. Technika tohoto kroku se podstatně liší od techniky, kterou hodláme použít ve vlastním výzkumu; nejčastěji zde používáme kvalitativní postupy (kupř. nestandardizovaný rozhovor). Cílem pilotní studie je zjistit, zda informace, kterou požadujeme, v naší populaci vůbec existuje a zda je dosažitelná.

Pokud nemáme opravdu hlubokou znalost o cílové populaci, zejména vzhledem ke studované problematice, pilotní studie je velice důležitá.

Dr. Watson:
Ale to my přece u nás vůbec nepotřebujeme! U nás přece neexistují žádné neznámé populace!

Bohužel mnoho lidí sdílí tenhle názor s Dr. Watsonem. Někdy to může být nebezpečné. Žádná společnost není kulturně homogenní a každý z nás má tendenci soudit jiné kultury či subkultury z pozic naší vlastní kultury. Říkáme tomu etnocentrismus a budeme o tom hovořit později. Když dva říkají totéž, nemusí to nutně totéž znamenat. Vzpomeňme si jen na náš příklad s Egri Bikaverem v naší pohádce č. 2. Dále to, co je v jedné skupině takovou informací, která může být snadno předána neznámému tazateli, v jiné subkultuře může být považováno za informaci, která „má zůstat v rodině". Různé profesionální subkultury mohou mít velice specifický systém hodnot, který je odlišný od hodnot existujících ve společnosti obecně. Tyto skupinové neformální normy často zabraňují tomu, aby se podávaly jakékoliv informace o nich všem nečlenům. Nemusí tu jít vůbec o nějaké ilegální skupiny: kupř. výzkum na populaci policistů je velmi obtížný všude na světě. Podobné problémy mohou být spojeny se zkoumáním etnických kultur. Kvantitativní akce kupř. na populaci Romů by byla bez seriózní pilotní studie nezodpovědná.

Pilotní studie je velmi často opomíjena, a to i v případech, kdy je to riskantní. Jako vždy je to problém času a peněz. Ale úspory v této etapě mohou být zaplaceny pochybnou validitou našich výsledků a mohou nakonec vyústit ve výzkumnou akci, která je zcela nevalidní.

Stejně nebezpečné je opominutí předvýzkumu:

> Předvýzkum
> je opět prováděn na malém vzorku naší cílové populace; tento vzorek je však obvykle větší než vzorek pro pilotní studii. Předvýzkum je testem nástrojů, které ve výzkumu hodláme použít.

Cílem tu je obvykle testovat srozumitelnost a jednoznačnost otázek. To je jedna z těch podivných věcí, které potkáváme ve výzkumné praxi: Když čteme záznamový arch, který konstruoval někdo jiný, je poměrně snadné poznat, že určité otázky nejsou zcela srozumitelné, že jiným otázkám je možno porozumět více než jedním způsobem. Na druhé straně je téměř nemožné objevit stejné neduhy v otázkách, které jsme formulovali sami. Já to vím, sám jsem si s tím nejednou udělal ostudu. Ale nespoléhejme s revizí otázek na vlastní kolegy. Oni totiž většinou patří ke stejné kulturní skupině jako my, mají přibližně stejný okruh znalostí, podobný systém hodnot. Značná část naší cílové populace může být velice odlišná. Zatímco určitá otázka je pro nás zcela přijatelná, pro jiné skupiny v populaci může být velice iritující. Něco, co považujeme za běžnou znalost, nemusí být tak docela běžné. Jen předvýzkum na vzorku cílové populace může minimalizovat, i když určitě ne vyloučit, omyly tohoto typu. Zatímco opominutí pilotní studie nás v některých případech nemusí vytrestat vážným zkreslením, opomenutí předvýzkumu je nesmírně riskantní: **předvýzkum by měl být nezbytnou součástí každé výzkumné akce.**

Když už pilotní studie byla skončena (a přirozeně ještě předtím, než můžeme začít s předvýzkumem), vstupujeme do velice zajímavé, ale velice rozsáhlé a komplexní oblasti konstrukce výzkumných nástrojů a do oblasti strategie práce v terénu. Důkladné seznámení s touto oblastí by vyžadovalo celou knihu a my máme k dispozici jen jednu kapitolu. Zbývají nám tu jen dvě věci: odkázat na dostupnou (a někdy méně dostupnou) literaturu a nabídnout velice stručný pohled na základní techniky sběru informací, na oblasti jejich uplatnění a zejména upozornit na to, jak a kde jsou tyto techniky nejzranitelnější. V této kapitole bychom chtěli také alespoň připomenout existenci některých méně obvyklých technik sběru.

Tak a už máme nejvyšší čas začít. Dr. Watson si už hlasitě stěžuje, že by chtěl konečně začít s konstrukcí dotazníku.

6.2 Ptát se, nebo neptat?

To je opravdu otázka. Mnohem širší a složitější, než by se na prvý pohled zdálo. Je mnoho a mnoho různých technik sběru dat. Vyjmenujme si tu alespoň několik z nich: standardizovaný rozhovor, nestandardizovaný rozhovor, skupinový rozhovor, dotazník, pozorování, telefonní výzkum, sociometrické techniky, sémantický diferenciál, sekundární analýza, zúčastněné pozorování a mnoho a mnoho jiných...

Naštěstí téměř všechny z těchto postupů jsou aplikací následujících základních technik:

- přímé pozorování
- rozhovor
- dotazník
- analýza dokumentů

Podívejme se nejdříve na definice těchto technik:

Přímé pozorování:
To je zaměřené, dobře plánované vnímání vybraných jevů. To, co bylo vnímáno, je pečlivě a systematicky zaznamenáváno.

Rozhovor:
Vyžadované informace jsou získávány v přímé interakci s respondentem. Rozhovor může být prováděn tváří v tvář nebo telefonicky.

Dotazník:
Respondent odpovídá písemně na otázky tištěného formuláře.

Analýza dokumentů:
To je analýza jakýchkoliv dokumentů, **které nebyly vytvořeny za účelem našeho výzkumu.** Záznamem mohou být právě tak dobře psané dokumenty jako jakékoliv materiální stopy lidského chování.

Podívejme se nejdříve na všechny tyto základní techniky najednou, na jejich aplikovatelnost v různých formách i přístupech, a srovnejme jejich epistemologické vlastnosti.

Přímé a nepřímé sledování

O tom jsme již hovořili v prvých dvou kapitolách; v sociologickém výzkumu většinou nesledujeme samotný studovaný jev, ale pouze jeho indikátory. Jen v technice přímého pozorování, je-li studovaným jevem chování, jsme schopni sledovat studovaný fenomén přímo.

Experimentální a neexperimentální aplikace

Ze čtyř základních technik sběru informací mohou tři být použity v experimentu. Studium dokumentů však ne. Nezapomeňme, že ve smyslu naší definice považujeme za dokument jen takové záznamy, které nebyly vytvořeny pro náš výzkum. To vylučuje možnost zachytit změny vyvolané manipulací s experimentální proměnnou.

Tabulka 6.1

	Přímé pozorování	Rozhovor	Dotazník	Dokumenty
Co je sledováno:				
jev	někdy	ne	ne	ne
indikátor	ano	ano	ano	ano
Použitelnost v experimentu				
	ano	ano	ano	ne
Forma				
standardizovaná	ano	ano	ano	ano
nestandardizovaná	ano	ano	ne	ano
Interference se zkoumaným systémem				
	někdy	ano	ano	ne

Standardizované a nestandardizované aplikace

Všechny tyto základní techniky mohou být použity ve **standardizované formě,** tři z nich i v nestandardizované; dotazník je přirozeně vždy standardizovaný. Koncept standardizace je

důležitý a je spojen s některými metodologickými problémy. Stojí za to, abychom se nad tímto konceptem na chvíli zamysleli.

> Standardizované techniky používají striktně jednotné podněty a také odpovědi jsou často omezeny na volbu z předem připraveného souboru kategorií. Pak hovoříme o „uzavřených otázkách".

Kupř. ve striktně standardizovaném rozhovoru má tazatel vlastně velice omezenou a nepříliš příjemnou funkci. Může jen přečíst otázku doslovně tak, jak je vytištěna v dotazovém archu. Když dotazovaný otázce nerozuměl, tazatel může otázku přečíst zvolna a zřetelně ještě jednou. Když ani tehdy respondent neporozumí, nezbývá tazateli nic jiného než zaškrtnout na archu kategorii „Neví/neodpověděl" a pak přejít k další otázce.

Dr. Watson:

Ale to přece nemůžete myslet vážně. Já bych určitě uměl přeformulovat otázku tak, že by každý porozuměl. Prostě bych otázku pokaždé změnil tak, že by „seděla na tělo" každého respondenta.

Nepochybujeme, že Dr. Watson by to určitě uměl stejně jako většina studentů, zejména z oblasti sociálních věd. Jenže to je důvod, proč právě tito studenti jsou zpravidla nejhoršími tazateli; snaží se většinou totiž „vylepšovat" výzkum. (Mimochodem téměř všude na světě jsou nejlepšími tazateli ženy v domácnosti, se středoškolským vzděláním a ve věkové skupině 35–45 let. Neptejte se proč. Nevím a nikdo neví, je to prostě empirický fakt.) V **kvantitativním** výzkumu si prostě nemůžeme dovolit žádné vylepšování otázek:

> Kvantitativní výzkum je testování hypotéz o skupinách, a ne o jedincích. Analýza je prováděna na kumulovaných datech o mnoha jedincích a data můžeme kumulovat jen tehdy, jsou-li totožná. Nezbytnou podmínkou pro to je, aby alespoň stimuly (například otázky) byly totožné.
>
> Nestandardizované postupy mohou získat lepší informace o jednotlivci, nabízejí lepší porozumění, ale cílem **kvantitativního** výzkumu je testování hypotéz o sociálních skupinách.

Jen si vyzkoušejte, kolika různými způsoby se můžete někoho zeptat, zda jí rád zmrzlinu. Lehce udiveně vyslovená otázka „Vy nejíte rád zmrzlinu?" určitě vyvolá jinou proporci pozitivních odpovědí než třeba otázka „Řekl byste, že máte rád zmrzlinu, nebo že ji nemáte rád?" Data získaná v nestandardizovaném rozhovoru různými tazateli mohou být nesrovnatelná. Tazateli v kvantitativním výzkumu nezbývá nic jiného než hrát obtížnou úlohu lidského magnetofonu.

Standardizace otázek (nebo jiných stimulů) je vlastně dalším typem redukce informací. Bohužel v kvantitativním výzkumu musíme aplikovat stejný typ redukce i na odpovědi zkoumaných osob. Početní operace mohou být prováděny jen na datech seskupených do relativně homogenních skupin. Chceme-li kupř. dojít k závěru, že jen 18 % populace nemá rádo zmrzlinu a že s přibývajícím věkem obliba zmrzliny klesá, nezbude nám nic jiného než standardizovat odpovědi, redukovat množství kategorií, ve kterých odpovědi budou zaznamenány. Tak kupř. odpovědi „Já bych za zmrzlinu dala život" a „Někdy zmrzlina nemusí být špatná" budou zatříděny do stejné kategorie. Horší je, že podobnou ztrátu informace musíme přijmout i tehdy, jedná-li se o problémy daleko závažnější, než je zmrzlina. Nejobvyklejším nástrojem pro standardizaci odpovědí jsou „uzavřené otázky".

> Uzavřené otázky nabízejí soubor možných alternativ, ze kterých respondent vybere vhodnou odpověď.

Ve standardizovaném rozhovoru můžeme použít místo uzavřených otázek otázky kódované respondentem. Ty mají formu normálních, „otevřených" otázek; výčet možných odpovědí **není předložen zkoumané osobě**, ale tazatel zaškrtává v záznamovém archu tu z připravených kategorií, která je nejbližší výroku respondenta.

Příprava uzavřených otázek není snadná. Jednak navržené kategorie musí představovat rozumný kompromis mezi ztrátou informace, která každou standardizaci nutně doprovází, a zvládnutelným počtem kategorií, jednak užité kategorie musí respektovat určitá formální kritéria:

1. Kategorie použité pro uzavřené otázky musí představovat soubor vyčerpávající všechny možné alternativy odpovědí.

2. Všechny kategorie se musí vzájemně vylučovat; nesmí být možné zařadit odpověď do více než do jedné z kategorií.

Vyhovět prvému požadavku může být snadné. Bohužel jen někdy. Jistě nebudeme váhat nad kategoriemi pro otázku po respondentově pohlaví. (Mimochodem tato otázka bude mít tři kategorie, musíme mít možnost sdělit počítači, že pro určitého respondenta nám odpověď na tuto otázku chybí: zkoumaná osoba zapomněla zodpovědět tuto otázku v dotazníku, tazatel opominul zaznamenat příslušnou informaci.) Ale představme si, jak by se připravovaly kategorie pro otázku na to, co dotazovaný dělá nejraději ve volném čase. Úplný výčet kategorií je zde skutečně nemožný. V takovém případě nezbývá nic jiného než zavést kategorii „jiná odpověď". V dotazníku požádáme dotazované, aby ono „jiné" vypsali, a stejně instruujeme i tazatele. To nám umožní, poté co data byla sebrána, ověřit, zda se vyskytuje v této všeshrnující kategorii nějaká odpověď velice často. Pokud by se tak stalo, můžeme pak vytvořit novou kategorii a příslušná pozorování překódovat.

Splnit druhý požadavek, vzájemnou výlučnost kategorií, je obvykle snazší. Nebezpečí zkreslení hrozí zejména tehdy, když musíme použít nějaké poměrně široké kategorie. Uveďme si alespoň jeden spíše extrémní příklad toho, co by se mohlo stát při zmíněném výzkumu volného času. Abychom se vyhnuli nepříjemnostem spojeným s dlouhým souborem kategorií, mohli bychom zavést několik málo velice obecných kategorií, jako „zábava", „poučení", „sport" atd. Kupř. luštění křížovek by někdo mohl zařadit do kategorie „zábava", někdo jiný do kategorie „poučení"; chození na fotbalové zápasy by se mohlo objevit právě tak v kategorii „sport" jako v kategorii „zábava". Jistě, spojování kategorií je možné a může být velice užitečné pro teoretické zobecnění. Seskupování původních kategorií do menšího počtu kategorií obecnějšího charakteru může být provedeno bezpečně teprve poté, co data byla sebrána.

Předběžné zavedení širokých kategorií může být velice nebezpečné: sebraná data mohou být snadno sdružována do obecnějších kategorií, ale jejich dekompozice do užších kategorií je vyloučena.

Později uvidíme, že nám statistika nabízí některé velice užitečné nástroje (jako kupř. faktorovou analýzu) pro optimalizaci tohoto procesu.

Proti pravidlu o vzájemné výlučnosti kategorií se dá hřešit ještě jiným zajímavým způsobem. V anglickém odborném žargonu existuje pro tenhle hřích pěkné jméno „double-barrel question", tedy „dvouhlavňová" otázka. Je to taková otázka, která vyžaduje jednu odpověď o dvou nebo více věcech najednou. Tento omyl potkáte často ve výzkumech, které různé instituce provádějí vlastními prostředky. Kdyby instituce byla požádána, aby provedla výzkum znečištění ovzduší v jejích objektech, najme pro tento úkol experta. Když tatáž instituce chce vědět, co si o této instituci lidé myslí, pověří kteréhokoliv zaměstnance, který umí číst a psát, úkolem zkonstruovat dotazník. Výsledky jsou pak docela neužitečné, ale většinou dost zábavné, a to se děje i v Kanadě, doslova přesycené profesionálními institucemi zabývajícími se výzkumem komerčně:

Pohádka pro odrostlejší děti 11

Jak se dělá demokracie

Náš obvod reprezentovala svého času dáma, která byla nejen velice elegantní, byla nejen ministrem ve federální vládě, ale záleželo jí moc na tom, aby každý z voličů v jejím obvodu věděl, jak si váží našeho mínění, a tak jsme dostávali asi tak třikrát ročně dotazník, krásně tištěný na těžkém křídovém papíře. Moc jsem se na to těšil. Byla to velice užitečná učební pomůcka pro mé kurzy ve výzkumné metodologii, klasický příklad toho, jak se výzkum nemá dělat. V těch dotaznících se vyskytovaly otázky tohoto typu:

Jak jste spokojen s politikou federální vlády ve vztahu k USA a se zahraniční politikou naší vlády obecně?

– velice spokojen
– dosti spokojen
– někdy spokojen, někdy ne
– nespokojen

V té době mnoho lidí bylo v podstatě spokojeno s kanadskou zahraniční politikou právě s výjimkou submisivního vztahu k USA. Tito lidé neměli možnost tuto otázku pravdivě zodpovědět.

Cvičení 6.1

Podívejme se znovu na znění otázky v naší pohádce č. 11. Víme, že se ptá na dvě věci najednou. Ale je to zkreslující otázka ještě z jednoho dalšího důvodu. Můžete vysvětlit proč?

Zatím jsme hovořili o redukci způsobené standardizací v rozhovoru a dotazníku. Nicméně všechny čtyři základní techniky mohou existovat ve standardizované formě. (Dotazník přirozeně existuje jenom ve standardizované formě, pokud neuvažujeme nějakou velice speciální situaci, ve které by bylo nutno připravit pro každého respondenta speciální, na tělo šitou verzi dotazníku.)

Standardizované pozorování je formou pozorování, která je v kvantitativním výzkumu používána téměř výhradně. Pozorování mnoha jedinců nemůže být zvládnuto jediným pozorovatelem, a tak je nutno zajistit srovnatelnost postupů celého pozorovatelského týmu. Minimálním požadavkem pro to je připravit záznamový arch s kategoriemi, ve kterých bude pozorované chování zaznamenáváno. Klasickým příkladem kategorizace tohoto typu je Balesova stupnice pro pozorování kooperativního chování v malé skupině. Zde jsou některé příklady použitých kategorií:

1. Projevuje solidaritu, pozoruje jiné, pomáhá, odměňuje.
2. Uvolňuje napětí, žertuje, směje se, ukazuje uspokojení.
3. Souhlasí, pasivně přijímá, ukazuje pochopení, podílí se, přizpůsobuje se.
4. Dává návrhy, poskytuje návod, respektuje však autonomnost ostatních.
5. Vyjadřuje mínění, oceňuje, analyzuje, vyjadřuje pocity a přání.
6. Udává orientaci, informuje, opakuje, vysvětluje, potvrzuje.
7. Vyžaduje orientaci, informace, žádá o opakování nebo o potvrzení.
8. Ptá se po mínění ostatních, po jejich postojích, hodnocení, analýze, sleduje výrazy jejich pocitů.
9. Vyžaduje pro sebe návrhy, směrnice pro postup.
10. Nesouhlasí, nesouhlas projevuje pasivně, spolupracuje jen formálně, neposkytuje pomoc ostatním.
11. Ukazuje napětí, žádá o pomoc, odtahuje se od činnosti.
12. Projevuje antagonismus, snižuje ostatní, hájí se a prosazuje se.

Podkladem pro konstrukci těchto tříd byl Balesův teoretický model, sledující v podstatě dvě dimenze. Prvá sleduje sociální emoce od kladné oblasti (body 1 až 4) až do oblasti negativní (body 10 až 12). Dále tato škála sleduje komponenty chování při řešení kolektivního úkolu. Tak kategorie 6 a 7 jsou zaměřeny na nutnou orientaci, kategorie 4 a 9 na problémy kontroly, 2 a 11 na vyrovnávání napětí atd. Pozorovatel kóduje pozorované chování na základě **vcítění** se do role člena skupiny, představuje si, jak by pozorované chování hodnotil, kdyby se opravdu zúčastnil jako člen na řešení úkolu skupiny. Balesova stupnice je výsledkem nesmírně rozsáhlé teoretické a zejména empirické analýzy. Škála je doprovázena podrobnou instrukcí, definující daleko zevrubněji použité kategorie, a uvádí příklady. I tak je nutné, aby pozorovatelé, kteří mají tuto stupnici používat, prošli intenzivním tréninkem.

Příklad Balesovy stupnice jasně ukazuje nejzranitelnější bod standardizovaného pozorování: problém reliability. Reliabilitu jsme již dříve definovali jako takové měření, které při opakované aplikaci – pokud se ovšem pozorovaný objekt nezměnil – poskytne shodné výsledky. Pro situaci standardizovaného pozorování si můžeme definici reliability přeformulovat asi takto:

> Reliabilní je taková metoda pozorování, při jejíž aplikaci skupina pozorovatelů pozorujících nezávisle stejný jev dospěje k shodným závěrům.

Jistě, většinou budeme pozorovat jednodušší jevy než R. F. Bales. Ale i tak je důsledná kontrola reliability nezbytnou podmínkou pro aplikaci standardizovaného pozorování.

Konečně i **analýza dokumentů** může být prováděna ve standardizované formě. Jistě už slyšíte protestovat Dr. Watsona, že to není možné, že dokumenty ve smyslu naší definice jsou záznamy, které nevznikly pro náš výzkum, takže jsme nemohli ovlivnit jejich standardizaci. To je pravda, ale můžeme standardizovat jejich analýzu. Máme pro to dokonce speciální metodu, která se nazývá **obsahová analýza** a bývá definována jako objektivní, kvantitativní analýza sdělení jakéhokoliv druhu. Budeme se jí podrobněji zabývat později.

> Nesporně standardizace, stejně jako každá jiná redukce informací, reprezentuje povážlivou ztrátu informací: bohatý, životaplný, mnohotvárný a mnohodimenzionální charakter skutečného světa je proměněn do krajně zjednodušeného schématu několika málo proměnných a do velice omezeného počtu kategorií. Bohužel kvantitativní výzkum je příliš neohrabaným nástrojem k porozumění. Rozumět, to je úkolem kvalitativního výzkumu. Kvantitativní výzkum je jen spíše brutální a primitivní nástroj k testování hypotéz. Má však jednu nesmírně důležitou vlastnost: je nám schopen říci, jaká je pravděpodobnost, že jsme na omylu. To však vyžaduje srovnávání údajů o mnoha jedincích. Srovnatelná data mohou být vyprodukována jedině standardizovanými postupy. **Nestandardizované techniky sběru dat jsou tedy v kvantitativním výzkumu prakticky nepoužitelné.**

Interference se zkoumaným systémem

To je další z opravdu velkých bolestí empirického výzkumu. Nejsme v tom sami, i naši kolegové v exaktní fyzice mají své problémy, když měřicí aparát ovlivňuje měřený systém. Již prvý pohled na naše čtyři základní techniky nám naznačí, že se rozpadají do dvou skupin s ohledem na to, jak silně mohou ovlivnit chování systému, který chceme studovat.

Webb, Campbell, Schwartz a Sechrest (1969) navrhli zajímavou klasifikaci zkreslení vyvolaného výzkumnými stimuly:

(1) **Efekt morčete:** zkreslení vzniklé tím, že zkoumané osoby jsou si vědomy toho, že jsou zkoumány.
(2) **Výběr role:** zkoumaná osoba, která si je vědoma, že je zkoumána, nebude reagovat přirozeně, ale tak, aby se jevila v co nejlepším možném světle.
(3) **Měření jako zdroj změny:** výzkumný proces může vyvolat ve zkoumaných osobách postoje, které předtím neexistovaly.
(4) **Stereotyp ve volbě odpovědí:** volba alternativních odpovědí je ovlivněna jinými faktory než míněním respondenta.

Zmínění autoři uvádějí ještě dvě další kategorie („interviewer effect" a „změny ve výzkumném nástroji"), které jsou podle jejich mínění vyvolány výzkumníkem. My se však domníváme, že jsou to kategorie podřazené prvým čtyřem a nebudeme je na tomto místě používat. Je zřejmé, že naše základní techniky se z hlediska prvé kategorie, zvané efekt morčete, rozpadají do dvou skupin. V případě **rozhovoru a dotazníku** zkoumaná osoba **vždycky** ví, že je předmětem výzkumu. Jsou to takzvané **reaktivní metody** (obtrusive methods), kde v procesu sběru informací vyvoláváme reakci, kterou zamýšlíme analyzovat. Naproti tomu pozorování a studium dokumentů je nereaktivní neagresivní metoda. Sběr informací zde nevyvolává, nebo by alespoň neměl vyvolávat žádnou reakci. (Ovšem některé z dokumentů mohly být vytvořeny reaktivními technikami při jejich vzniku, mimo rámec našeho výzkumu.)

To ovšem neznamená, že pozorování je vždy nereaktivní: pozná-li zkoumaná osoba, že je pozorována, stává se tato technika nutně reaktivní technikou. To se velmi často děje při experimentální manipulaci. Klasický je tzv. Hawthornský efekt. Roethlisberger a Dickson (1939) studovali produktivitu dělníků Western Electric v Chicagu a jejich spokojenost

s pracovními podmínkami. Zlepšené osvětlení pracoviště se odrazilo ve zvýšené produktivitě a vyšší spokojenosti dělníků. Další zlepšení osvětlení vedlo opět k ještě vyšší produktivitě. Pak – jako experimentální kontrola – světlo bylo opět ztlumeno, a kupodivu, produktivita opět vzrostla. Rozhodujícím činitelem nebyla reakce na fyzické změny, ale reakce dělníků na pozornost, která jim byla věnována. Ale my už víme, že použití kontrolní skupiny ve skutečném experimentu může pro tento typ zkreslení vyloučit.

Pod hlavičkou „výběr role" je skryt zdroj zkreslení, který je velice typický jak pro dotazník, tak i pro rozhovor. Zkoumané osoby mají přirozenou tendenci ukázat se v odpovědích v co nejlepším světle. V případě rozhovoru k tomu přistupuje ještě nový faktor, vyvolaný osobou tazatele; respondent může volit odpovědi tak, aby udělal co nejlepší dojem na konkrétní osobu, na tazatele.

Pohádka pro odrostlejší děti 12

Historie téměř beze slov aneb O dvou tazatelích

Tazatel 1:
„Čtete rád poezii?"

Respondent Pepík:
„Zbláznil jste se?"

Tazatel 2:
„Čtete rád poezii?

Respondent Pepík:
„Ó, já miluji poezii!"

Tomuto zkreslení, nazývanému v anglicky psané literatuře „interviewer bias", je obtížné předejít. Není vyvoláno jenom tím, jaký tazatel je, nejen tím, jak se chová, ale hlavně tím, jak je respondentem vnímán. Chování tazatele může nechtě respondentovi nabídnout klíče ke konstrukci falešné či správné představy o tazateli, a tato představa ho pak vede k odpovědi, o které se domnívá, že na tazatele učiní co největší dojem. Neutrální chování tazatele je nelehký úkol, zejména je-li pro něj obsah výzkumu nějak významný. Autor výzkumu může být proto velmi špatným tazatelem. Většina profesionálních výzkumných organizací používá služby tazatelů, kteří mají své vlastní hlavní povolání, a je osvědčenou praktikou, že tazatelé, kteří jsou svým povoláním spojeni s předmětem určitého výzkumu, jsou z účasti na tomto výzkumu vyloučeni. Pojišťovací agent není použit jako tazatel pro výzkum zabývající se pojištěním; těžko by zachoval „poker face" při nehorázné odpovědi respondenta, kterou já, jako osoba zcela naivní v tomto oboru, klidně zaznamenám.

Podobnou roli může hrát tzv. **efekt záhlaví**, to jest zkreslující efekt představy zkoumaných osob o instituci, která výzkum pořádá nebo se kterou respondent daný výzkum asociuje. Následující pohádka není vtip, ale fakt z historie naší, či spíše východoevropské sociologie.

Zkreslení výběrem role se ovšem neomezuje jen na rozhovor, ale je vlastní i dotazníku. Respondent se neidealizuje vzhledem ke konkrétnímu tazateli, ale vzhledem k nějak zobecněným druhým lidem nebo vzhledem k instituci, kterou vidí či tuší za výzkumem. Naštěstí máme k dispozici některé technické postupy, které zvýší pravděpodobnost pravdivé odpovědi i na nepříjemné otázky. Tyto postupy nejsou bez problémů a budeme je probírat na následujících stránkách.

Horší je to se zkreslujícím vlivem měření jakožto zdroje změny. Krásný příklad jsme uvedli v naší pohádce č. 3. Bohužel takové benigní případy jsou spíše výjimkou. Tazatelů tak efektivních jako Jirka není mnoho a tak obětavé a hodné babičky snad existují už jenom v pohádkách. Mnohem zhoubnější je vliv měření na postoje respondenta, zejména na postoje, které ještě nejsou pevně zakotveny v jeho vědomí. Naše výzkumy jsou často zaměřeny na

problémy, o kterých někteří respondenti nikdy nepřemýšleli, které pro ně nejsou skutečnými problémy a nemají k nim vyhraněný postoj. Teprve naše **otázka tento postoj vytvoří.** Odpověď dotazované osoby není nepravdivá, ale vyjadřuje postoj, který je nový a asi značně nestabilní. Naší hlavní obranou je předvýzkum: jen ten je nám schopen říci, zda požadované informace skutečně v cílové populaci existují. Dále nám mohou pomoci **filtrační otázky**. To jsou otázky verifikující, jak dalece je respondent kvalifikován poskytnout žádanou informaci. Zjišťují třeba zkušenost respondenta z dané oblasti, jeho znalost atd.

Pohádka pro odrostlejší děti 13

O dvou výzkumech

To bylo vlastně dost dávno, kdy sociologie byla stále ještě nemravné slovo, které bylo opisováno mnoha různými způsoby, třeba jako vědecko-historický výzkum o přítomnosti. A přece v té době Erika K. dokázala připravit znamenitý výzkum o náboženství v některých oblastech Čech a Moravy. Použila velice moudrou techniku, která téměř dokonale zajišťovala anonymitu. Dotazníky byly osobně rozdávány a po několika hodinách se tazatel vrátil, dal respondentům frankovanou obálku a respondenti sami vyplněný dotazník vrátili poštou. Zajímavé je, že některé vrácené obálky měly razítka poštovních úřadů vzdálených mnoho desítek kilometrů od míst, ve kterých se výzkum konal. Erika tak docílila nejen neobyčejně vysoké návratnosti, ale získala i velice upřímné odpovědi. Její výsledky byly u mocných tehdejšího světa velice nepopulární.

V té době navštívil Prahu jeden sociální vědec z jednoho nevelkého spřáteleného socialistického státu. Uviděl výzkum a řekl Erice: „To je velice zajímavé. Já ten výzkum zopakuji. Ale udělám ho lepší!" A taky udělal. Když se asi o rok později vrátil do Prahy, prohlásil vítězně: „Neříkal jsem to, já mám mnohem lepší výsledky než Erika. Měl jsem stoprocentní návratnost, a náboženské předsudky u nás už vůbec neexistují." Možné je všechno, ale o perfektní návratnosti ještě nikdo nikdy neslyšel, a tak nám to ten soudruh profesor vysvětlil: „Já jsem vybral mnohem lepší tazatele, než měla Erika. Já jsem vybral ty osoby, které o obci nejvíce vědí a kterým každý důvěřuje: **místní milicionáře!**"

Stereotyp ve volbě odpovědi je v literatuře znám pod jménem *response set*. Je to systematické seskupování odpovědí do určitých kategorií a toto seskupování nemá nic společného s obsahem otázky ani s postojem, míněním respondentů. Nicméně toto seskupování není

náhodné. Jeden typ tohoto zkreslení je znám pod jménem **acquiescence** (podvolení, přitakání). Mnoho respondentů má tendenci s předloženým tvrzením souhlasit spíše než vyjádřit nesouhlas. To je zejména časté, když vyžadujeme odpověď ve formě souhlasu nebo nesouhlasu s předloženým tvrzením:

Většině lidí je možno důvěřovat.
- silně souhlasím
- souhlasím
- nevím
- nesouhlasím
- silně nesouhlasím

V rozsáhlejších výzkumech můžeme toto zkreslení minimalizovat tak, že na různých místech dotazníku nebo rozhovoru předložíme obsahově shodné tvrzení, jednou v pozitivní a jednou v negativní formě. Tak předchozí otázka by mohla být opakována takto:

Lepší je nedůvěřovat nikomu.
- silně souhlasím
- souhlasím
- nevím
- nesouhlasím
- silně nesouhlasím

Dalším zdrojem zkreslení je tendence mnoha zkoumaných osob volit střední, neutrální kategorii v uzavřených otázkách. Řešením tu může být technika **nucené volby**. Střední, neutrální kategorie („nevím" v našem příkladu) není prostě v uzavřené otázce použita a respondent se musí rozhodnout mezi kategoriemi popisujícími různé stupně souhlasu či nesouhlasu.

Někdy „response set" může nabýt velice zajímavé formy:

Pohádka pro odrostlejší děti 14

O zkreslení vpravdě tvořivém aneb Pohádka s mravním ponaučením

Kdysi, kdesi, ale v podstatě po celém civilizovaném světě, byly pro poradenství ve volbě povolání používány různé dotazníky. Jeden z nich se snažil o co nejúplnější inventuru kandidátových zájmů. Byl to velmi dlouhý dotazník, obsahoval několik set položek. Každá z nich obsahovala stručný popis nějaké činnosti nebo prostředí. Vypadalo to asi takto:

112. Pracovat se zvířaty A a ? n N
113. Řešit složité problémy A a ? n N

Respondenti byli instruováni zakroužkovat „A", když zmíněná činnost byla pro ně velice zajímavá, „a", když byla poněkud zajímavá, „?" ani zajímavá, ani nudná, „n" reprezentovalo nudnou a „N" velice nudnou činnost.

Viděli jsme vyplněné dotazníky z výzkumu na populaci učňů, který byl proveden v prvých poválečných letech, a bylo to velmi zajímavé čtení. Prvé stránky dotazníků byly vyplněny většinou pečlivě. Ale postupně přibývalo přeskočených otázek, zdaleka ne všichni respondenti se dostali až ke konci dotazníku. Ne všichni však podlehli v tomto zápase s nudou bez boje. Zaškrtnuté odpovědi začaly dostávat formu geometrických obrazců. Viděli jsme kolmice. Šikmé linky v nejrůznějších kombinacích a profesor Hyhlík pečlivě střežil dotazník vyplněný jakýmsi tvořivým mladým mužem, jehož odpovědi na posledních stránkách dotazníku vytvářely opakovaně populární nemravný obrázek.

A tady je to slíbené mravní naučení: **Je snadné předpovědět, že výzkumný proces, který je pro respondenty nudný, vyvolá nutně zkreslení. Třeba i ve formě *response setu*. Ale je vskutku nemožné odhadnout, jaký bude obsah tohoto zkreslení.**

Tak a teď už víme něco o základních nástrojích sběru informací. Ale nevíme ještě, jak a kdy je používat. Alespoň něco si o tom povíme v příští kapitole.

Kapitola 7

Kulhavý poutník

> Absolutní znalost neexistuje. Každá informace je neúplná. Musíme k ní přistupovat pokorně.
>
> J. Bronowski, *The ascent of man*

A teď je konečně čas prodrat se alespoň některými malými úseky džungle technických aspektů sběru informací. Je to pěkně komplikovaná džungle a její spolehlivé zmapování by vyžadovalo celou knihu. Na našich několika stránkách můžeme ukázat jen některé z nejnebezpečnějších močálů, upozornit na některé důležité nebo zajímavé součásti naší výzbroje a výstroje a podívat se alespoň stručně na to, jak tyto nástroje aplikovat. Tato kapitola není návodem, který by vypočítával zevrubně všechny techniky, jejich aplikovatelnost a definoval pravidla jejich použití. Naše kapitola je spíš něco jako prodejní katalog velkého obchodního domu, který chce inzerovat něco z toho, co má na skladě. Účelem není prodat, ale přimět vážnějšího zájemce, aby se patřičně vyzbrojil na pořádnou expedici. Tato výzbroj není příliš složitá, chce to jenom číst.

7.1 Dotazník, nebo rozhovor?

Dr. Watson:
To je přece zbytečná otázka. Každý přece používá dotazník!

Tak se na to podívejme zblízka. Tabulka 7.1 porovnává některé důležité vlastnosti obou technik, standardizovaného rozhovoru a dotazníku.

Tabulka 7.1

Rozhovor	Dotazník
Velice pracná a nákladná technika sběru informací.	**Vysoce efektivní technika, která může postihnout veliký počet jedinců při relativně malých nákladech.**
Rozhovor je časově velice náročný. Získat informace v rámci určitého časového limitu může být velice nákladné a často i nemožné.	**Dotazník umožňuje poměrně snadno získat informace od velkého počtu jedinců v poměrně krátkém čase a s poměrně malým nákladem.**
Rozhovor vyžaduje spolupráci dosti velikého počtu alespoň částečně vyškolených tazatelů v terénu.	**Spolupracovníci v terénu jsou nezbytní jen někdy (při použití osobně rozdělovaných a sbíraných dotazníků). Požadavky na jejich zaškolení jsou nízké.**
Výzkum na prostorově rozptýleném vzorku je nákladný.	**Náklady šetření na rozptýleném vzorku jsou relativně nízké.**
Anonymita výzkumu je pro respondenty málo přesvědčivá.	**Anonymita je relativně přesvědčivá.**
Rozdíly mezi tazateli a rozdíly v jejich chování mohou vyvolat „interviewer bias".	**Formální shodnost podnětové situace je vysoká; „interviewer bias" je prakticky vyloučen.**
Rozhovor klade menší nároky na iniciativu respondenta, pro respondenta je obtížnější vynechat odpovědi na některé otázky.	Dotazník klade vysoké nároky na ochotu dotazovaného, je snadné „přeskočit" otázky nebo neodpovědět vůbec.
V rozhovoru je téměř jisté, že dotazovaná osoba je ta, která byla vybrána do vzorku.	U dotazníku je možné, že otázky byly zodpovězeny jiným členem rodiny, či, a to nejčastěji, celým rodinným týmem.
Proporce úspěšně dokončených rozhovorů je podstatně vyšší než návratnost dotazníku.	Návratnost je velice nízká. S výjimkou některých speciálních případů je tak nízká, že jakákoliv reprezentativnost vzorku je ztracena.

Že by měl Dr. Watson pravdu? Z devíti kritérií zmíněných v naší tabulce jich šest jednoznačně svědčí ve prospěch dotazníku. Některé důvody proti použití rozhovoru jsou závažné. Je to především existence zkreslení vyvolaného tazatelem a zejména vysoká nákladnost tohoto postupu. Náklady na rozhovor jsou ve srovnání s dotazníkem mnohonásobné. Nicméně jedna jediná vlastnost dotazníku naprosto znehodnocuje jeho dobré stránky: nesmírně nízká návratnost. To platí zejména o velice ekonomických, poštou distribuovaných dotaznících. I když použijeme všechny důmyslné triky zkonstruované ke zvýšení návratnosti (o některých z nich se zmíníme v následujících odstavcích), návratnost 50 % je považována v mnoha průmyslových zemích za slušný úspěch. Nejvážnější na tom je, že taková

dramatická redukce velikosti vzorku má **systematický charakter**, že výsledky založené na takto zkomoleném vzorku se mohou velmi podstatně lišit od výsledků, které bychom dostali ve vzorku skutečně reprezentujícím zkoumanou populaci. Žádné zvýšení velikosti vzorku toto zkreslení neodstraní, a tak nám nezbývá než se smířit s tím, že dotazník, alespoň v této formě, **je pro seriózní výzkum většinou nepoužitelný**.

Dr. Watson
(tiše vzlyká)

Tak to musíme našeho přítele alespoň trochu utěšit. Existují totiž i jiné techniky distribuce dotazníku než pošta. Dotazníky mohou být osobně distribuovány a sbírány, a to zaručuje návratnost srovnatelnou s rozhovorem; ovšem finanční a časové náklady jsou také srovnatelné. Někdy můžeme náklady spojené s osobní distribucí dotazníků minimalizovat. Je-li určitá populace prostorově koncentrována, můžeme použít techniku, pro kterou mají Němci krásné dlouhé jméno: Klassenzimmerfragenbogen. Respondenti jsou shromážděni v místnosti, obdrží dotazník, který na místě vyplní. Tato technika funguje velmi dobře na některých víceméně bezbranných populacích, jako jsou studenti, vojáci nebo zaměstnanci. Velice chytrá technika byla použita svého času studijním oddělením Čs. rozhlasu: dotazníky byly rozdělovány a sbírány poštovními doručovateli, a protože listonoši byli odměňováni podle počtu vyplněných dotazníků, návratnost byla vysoká. V tomto případě doručovatelé vybírali respondenty sami, podle jednoduchých kritérií kvótního výběru. Snad by bylo možné použít tuto techniku i pro pravděpodobnostní vzorek, přičemž doručovatelé by distribuovali a sbírali dotazníky členům vzorku zkonstruovaného technikami náhodného výběru.

Ale i poštou distribuovaný dotazník může být někdy relativně úspěšný, zejména je-li pro naši populaci předmět výzkumu důležitý nebo pokud se jedná o nějaký citlivý problém, kdy validita výsledků by mohla být podstatně ohrožena tím, že kupř. v rozhovoru není dostatečně zajištěna jejich anonymita.

V každém případě musíme použít všechny prostředky, které mohou návratnost dotazníku zvýšit. Především musíme respondentovi co nejvíce ulehčit zodpovězení a vrácení dotazníku. Zanedbání některých základních kroků může vést k opravdu katastrofální návratnosti. Frankovaná obálka, slušná kvalita papíru, dostatek místa pro odpovědi, to jsou základní nezbytnosti. Důležitou úlohu hraje délka dotazníku: se vzrůstající délkou dotazníku návratnost klesá. Je zajímavé, že dosáhne-li dotazník délky asi deseti stránek, pak další růst délky nad tuto hranici zřejmě již žádný měřitelný vliv na návratnost nemá. Vliv na návratnost může mít i prestiž sponzora, pro kterého je výzkum prováděn. Ovšem jméno některých institucí může vyvolat systematické zkreslení. Např. uvedeme-li jako sponzora výzkumu o postoji k potratům jméno organizace dobře známé svým negativním postojem k potratům, pak ovšem tento fakt výběrově zvýší návratnost u těch respondentů, kteří sdílejí postoje takové organizace. Dobrou pomůckou je i doprovodný dopis, který vysvětlí účel výzkumu, jeho užitečnost pro respondenty, vysvětlí, jak byl respondent vybrán a jak je zajištěna jeho anonymita. Z dopisu by mělo být pro respondenta zřejmé, jak významné jsou pro nás jeho odpovědi. Z tohoto důvodu se používají stále častěji personalizované doprovodné dopisy. Prakticky všechny modernější programy pro automatické zpracování textu nám umožní připravit dopis, do kterého program automaticky vloží z připravené databáze adresu, oslovení atd.

Jedním z nejefektivnějších nástrojů ke zvýšení návratnosti je odměna za vyplnění dotazníku. Obvykle nebývá příliš vysoká, jde zde spíše o její symbolický význam. V angličtině existuje běžně užívané rčení „Penny for your thought", což bychom mohli přeložit přibližně jako „Peťák za to, co si myslíš". Lístek s touto frází a zbrusu nová jednocentová mince – nejnižší hodnota existující v měně USA – byly připojeny k dotazníku. Tato jednoduchá a nenákladná akce zvýšila významně návratnost dotazníku. Jindy byl stručný dotazník vytištěn na rub šeku na malý obnos a vyinkasované šeky-dotazníky byly vráceny bankou výzkumníkovi. Častou odměnou za vyplnění dotazníku jsou losy asi v hodnotě 1 $, jindy jsou vyplněné dotazníky zařazeny do slosování o větší ceny. Ale pozor, nabídnutá odměna může někdy zkreslit naše výsledky, jak je to ilustrováno v naší absurdní pohádce č. 15:

Pohádka pro odrostlejší děti 15

Všeho moc škodí aneb Pohádka o absurdní návratnosti

Byla jednou jedna země, ze které mohli cestovat do zahraničí jen Ti Nejvybranější, a těch nebylo mnoho. Jednoho dne se Ti Nejvybranější rozhodli, že potřebují spolehlivou informaci o postojích zbytku obyvatelstva. Rozhodli se, že použijí techniku dotazníku, a protože se doslechli, že návratnost je důležitá, nabídli velkorysou odměnu: všechny vyplněné dotazníky budou zařazeny do slosování. První cena byla opravdu velkorysá: dvoutýdenní pobyt pro dva v jedné krásné a teplé cizí zemi. Tu cenu mohl získat každý, i ten, kdo nepatřil mezi Ty Nejvybranější. Výsledek byl vskutku epochální. Z 2000 dotazníků se jich vrátilo 2400, a tak poprvé v dějinách návratnost překročila 100 %. To je pozoruhodné proto, že v té zemi byly všechny tiskařské stroje přísně kontrolovány.

Ovšem nejúčinnější cesta ke zvýšení návratnosti rozesílaného dotazníku je velice pracná a prozaická. Jedná se o to, čemu se v anglicky psané literatuře říká **follow-ups**. Po určitém počtu dnů, typicky po dvou týdnech, respondenti, kteří nevrátili dotazník, dostanou „upomínku", obvykle pěkně formulovaný text na korespondenčním lístku. Asi dva týdny poté můžeme těm, kteří ještě neodpověděli, poslat nový exemplář dotazníku a jsme-li opravdu vytrvalí, můžeme po určité době těm zbývajícím poslat doporučeně další exemplář dotazníku. Ve skutečnosti to není tak jednoduché, jak to na první pohled vypadá. Tento postup vyžaduje, abychom měli přesnou evidenci o tom, kdo odpověděl a kdo ne, přitom však nesmíme porušit anonymitu výzkumu. Řešení tohoto problému není těžké: ke každému dotazníku připojíme nejen frankovanou odpovědní obálku, ale i korespondenční lístek s respondentovou adresou a požádáme respondenty, aby tuto dopisnici odeslali, když odesílají dotazník. To všechno je ovšem velmi pracné a nákladné, zejména při výzkumech na velkém vzorku; osobní počítače nám i zde mohou práci podstatně ulehčit. Jak účinná je tato metoda, svědčí výsledky studie provedené v USA: návratnost na původně rozeslané dotazníky byla 23,8 %. Pak byly aplikovány všechny kroky popsané výše, a po sedmi týdnech návratnost dosáhla úctyhodných 72,4 % (Dilman, Christensen, Carpenter & Brooks 1974).

Kdykoli je návratnost neuspokojující, měl by výzkumník vědět, jak se jeho vzorek liší od populace. Nejjednodušší je srovnat distribuci dat v dotazníku se známými daty o populaci. Prakticky to budou téměř vždycky jen demografická data. Nesmíme však zapomenout, že i kdyby byla značná shoda mezi oběma skupinami dat, nezaručuje to, že ti, kteří zodpověděli dotazník, se neliší od těch, kteří neodpověděli v jiných, pro náš výzkum důležitějších dimenzích. Jsou-li naše výsledky opravdu důležité, mají-li třeba být podkladem pro nějakou

sociální intervenci, může být prozíravé zjistit alespoň směr zkreslení v našem nedokonalém vzorku. K tomu opět potřebujeme mít evidenci o adresách těch, kteří neodpověděli. Pak můžeme na vzorku vybraném z těch, kteří neodpověděli, provést výzkum rozhovorem. To nám alespoň umožní odhadnout, kterým směrem se odchylují naše výsledky od těch, jež bychom obdrželi na nezkresleném vzorku.

Tak doufejme, že se nám podařilo Dr. Watsona utěšit. A když už jsme v takové smířlivé náladě, podívejme se na to, zda můžeme něco udělat s nejzranitelnějšími body v rozhovoru, s jeho nákladností a se zkreslením vyvolaným přítomností tazatele (interviewer bias).

Nejslibnějším řešením se zdá být telefonní interview. Pro nás je to řešení prozatím jenom perspektivní; jak již samo jméno naznačuje, tento postup je omezen jen na populaci těch respondentů, kteří mají telefon, a tato populace se může podstatně lišit od ostatního obyvatelstva. (Pamatujete na blamáž, kterou utrpěl v roce 1936 Literary Digest? Ta byla vyvolána právě zkreslením tohoto typu. Možná že by bylo užitečné podívat se znovu na naši Pohádku č. 8.) V Severní Americe se telefonní interview stává v poslední době takřka rutinní metodou terénního sběru informací: v roce 1987 (U. S. Bureau of Census 1987) mělo více než 97 % amerických domácností telefon. Tento typ sběru informací má četné výhody: Je sice dražší než rozesílaný dotazník, ale mnohem levnější než rozhovor tváří v tvář. Kupříkladu v USA se náklady na dobrý telefonní výzkum pohybují kolem 50 % nákladů na rozhovor. Návratnost je mnohem lepší než u dotazníku. Nezodpovědět dotazník znamená pro potenciálního respondenta zůstat pasivní. Odmítnout telefonní interview vyžaduje akci; výslovně odmítnout nebo vymyslet si nějakou výmluvu. Třeba: „Někdo je u dveří, musím bohužel končit." Zvláště krásnou výmluvu cituje Babbie (1989, s. 252): „Ježíši, prasata žerou moje VOLVO." To se však stává většinou tehdy, je-li náš interview příliš dlouhý nebo když překročíme dobu trvání rozhovoru, kterou jsme na začátku zkoumané osobě slíbili. Telefonní interview prakticky zamezí nepříjemnému fenoménu, velmi častému u dotazníku – vynechání odpovědi na některé otázky. Snižuje také pravděpodobnost, že otázky budou zodpovězeny jiným členem rodiny než tím, který byl vybrán do vzorku, nebo – což se v případě dotazníku stává velice často – celou rodinnou radou.

Telefonní interview má také některé výhody ve srovnání s rozhovorem tváří v tvář: anonymita je v telefonním rozhovoru mnohem přesvědčivější. Kontakt mezi tazatelem a respondentem je omezen jen na hlasovou komunikaci, a tak „interviewer bias" je mnohem slabší. Telefonní

interview však zejména umožňuje menšímu počtu vyškolených tazatelů zvládnout velký, prostorově rozptýlený vzorek v poměrně krátkém čase a bez cestovních nákladů. Další výhoda je pro nás naštěstí snad zatím jenom hypotetická. S dramaticky rostoucí kriminalitou ve velkých severoamerických městech stále vyšší a vyšší procento lidí odmítá vpustit cizince – tazatele – do svého bytu a jsou zaznamenány i případy, že tazatel byl fyzicky napaden. Výzkumná agentura pak nese riziko žaloby o odškodné. Není proto překvapivé, že se telefonní interview zvolna stává základním nástrojem rutinního výzkumu. Bohužel rozmach této techniky je nyní brzděn kontaminací terénu „telemarketingem", obchodem po telefonu. Průměrný občan dostává alespoň tak šestkrát měsíčně telefonní nabídky, aby si nechal opravit střechu za neuvěřitelně výhodných podmínek, je mu gratulováno, že byl vylosován mezi několik šťastných, kteří mohou zakoupit za poloviční cenu kondominium na Floridě, nebo že může za pouhých $ 150,– a během jednoho měsíce získat doktorát teologie od Divine University v Moose Jaw. Tyto nabídky se často maskují jako výzkum, a to pak zvyšuje pravděpodobnost odmítnutí i pro výzkum zcela legitimní. Můžeme to kompenzovat tím, že na telefonní výzkum upozorníme dopisem adresovaným vybranému tazateli. To však vylučuje použití neobyčejně slibné varianty telefonního interview: CATI.

Dnes již existuje celá řada programů pro **telefonní interview s pomocí počítače** (Computer Assisted Telephone Interview). Ty pracují asi takto. Tazatel čte otázky z obrazovky a ukládá ihned odpovědi respondentů přímo do ústřední počítačové paměti, ve které jsou kumulovány odpovědi shromážděné všemi tazateli. Po skončení kteréhokoliv interview jsou data okamžitě v takové formě, že mohou být bez jakýchkoliv dalších operací podrobena statistické analýze, a někdy jsou určité části této analýzy prováděny průběžně. V kterékoliv fázi sběru informací můžeme třeba zjistit, kolik mužů a kolik žen jsme do této chvíle interviewovali, jaký je průměrný věk respondentů, od nichž jsme až do této chvíle shromáždili odpovědi atd. Kódování a vstup dat jako samostatné operace odpadají. Tyto operace jsou velice nudné a jsou proto zdrojem častých chyb. Odpadá také nutnost tisknout tisíce stránek záznamových archů a manipulovat s nimi. Tím však výhody této metody nekončí. CATI program automatizuje aplikaci strukturovaných, větvených otázek. Tazatel už nemusí sledovat instrukci: „Když odpověď na otázku č. 136 byla „NE", přejdi ihned k otázce č. 142." O to se teď postará počítač sám. Když tazatel zaznamenal „NE" jako odpověď na otázku 136, jako další otázka se na obrazovce automaticky objeví otázka 142. CATI programy jsou opravdu docela chytré, kontrolují tazatele. Upozorní kupř., že v určité odpovědi byl použit „ilegální" kód: v záznamu odpovědi na otázku, která má jen tři kategorie kódované 0, 1 a 2, se tazatel překlep

a zaznamenal kód 3. Program je také schopen zachytit chybné odpovědi, které jsou zřejmě nemožné nebo velice nepravděpodobné, jako např. datum narození 29. února v roce, který není přestupný, nebo dokončené vysokoškolské vzdělání u respondenta, který udal, že jeho věk je dvanáct let. Počítač upozorní ihned tazatele na tyto nesrovnalosti. Tazatel se k těmto problémům vrátí a na místě je opraví. Program je také schopen zabránit náhodnému vynechání otázek: další otázka se neobjeví, dokud tazatel nezaznamenal odpověď na předchozí otázku. Chceme-li, můžeme program instruovat, aby upozornil tazatele – zejména používáme-li kontrolní otázky – na nesrovnalosti v respondentových odpovědích. CATI program může zabránit zkreslení vyvolanému pořadím otázek. V rozhovoru jakákoliv otázka může ovlivnit odpověď na kteroukoliv z následujících otázek. CATI program nám umožní měnit **náhodně** pořadí otázek, nebo lépe bloků otázek, a tak minimalizovat toto zkreslení. (Viz Frey 1989 a Lavrakas 1987.)

Ale snad nejdůležitějším důvodem pro rychlý růst popularity CATI programů je možnost aplikace **náhodné volby telefonních čísel**. Při tomto postupu se konstrukce vzorku, spolu s rozhovorem, kódováním a „čištěním" dat (jak jsme to diskutovali v předchozím odstavci) stávají jedinou, souběžnou operací. Logika, na které spočívá tato technika tvorby náhodného vzorku, je totožná s technikou vícestupňového náhodného výběru, o němž jsme mluvili v páté kapitole. Náhodná volba telefonních čísel má všechny výhody vícestupňového výběru a odstraňuje i jeho nákladnost a pracnost. Z geografické oblasti dané definicí populace jsou náhodně vybrány územní jednotky definované telefonními ústřednami a pak iniciativu přebírá počítač. Generuje náhodně potřebný počet číslic, a když volané číslo odpoví, předá slovo tazateli. Ten, poté co zjistil, že volané číslo je domácí, a ne obchodní nebo úřední telefon, začne otázkou: „Můžete mi říci, prosím, kolik osob žije ve vaší domácnosti," a pak pokračuje s náhodným výběrem uvnitř domácnosti stejně, jak jsme to popsali v souvislosti s vícestupňovým náhodným výběrem. Náhodná volba telefonních čísel je velice účinnou technikou konstrukce vzorku. Klečka a Tuchfarber (1978) ji aplikovali ve výzkumu o postojích ke zločinu, jeho obětem a k policii a porovnali výsledky se studií provedenou v roce 1974 rozhovorem tváří v tvář (U. S. Bureau of Census). Výsledky obou studií byly velice podobné jak vzhledem k demografické skladbě vzorků, tak vzhledem k distribuci zkoumaných postojů.

7.2 Otázky o otázkách

Konstrukce otázek, to je opravdu problém, hodně komplikovaný a riskantní. Některé základní gnozeologické problémy spojené s transformací informace jsme již probrali v kapitole 3. Technické stránky stavby otázek – to je sama o sobě dost složitá záležitost, o které byla napsána celá řada knih. V naší knížce máme na to jen málo místa, jen tolik, abychom uvedli několik nejdůležitějších, víceméně technických námětů, které mohou být užitečné při konstrukci otázek. Naprostá většina toho, o čem tu budeme mluvit, platí právě tak pro rozhovor jako pro dotazník.

Takže jste-li svojí povahou náchylní k dobrodružství a dáváte-li (vzhledem k zdravotnímu stavu nebo k přeplněnosti lezeckých oblastí) přednost empirickému sociologickému výzkumu před horolezectvím, je zde řada otázek, které si musíte zodpovědět o každé otázce, kterou hodláte použít pro svůj standardizovaný rozhovor nebo dotazník.

(1) Je tato otázka opravdu nezbytná?

Pro kterou z našich hypotéz tuto otázku potřebujeme? Důvod, že by určitá otázka mohla poskytnout zajímavou informaci, naprosto není důvodem k jejímu zahrnutí do výzkumu; na to je empirický výzkum opravdu trochu drahou hračkou.

(2) Měří tato otázka opravdu to, co chceme měřit?

Není nezbytné aplikovat na ni některý z testů validity, o kterých jsme již hovořili? Zdravá skepse je neocenitelnou vlastností výzkumníka. Rozhodně je lepší ztratit čas „zbytečným", to jest úspěšným testem validity než stavět své závěry na nevalidních datech.

(3) Bude tato otázka srozumitelná opravdu každému členu našeho vzorku? Budou jí všichni respondenti rozumět stejným způsobem?

Každý z nás má svůj kulturní bias a je téměř nemožné odhadnout, jak bude otázka vnímána příslušníky jiné sociální třídy, lidmi s jiným vzděláním atd. Jedinou jistotu nám může dát

vyzkoušení otázky ve skutečném předvýzkumu, tedy na malém vzorku naší populace, a ne na kolegovi, jehož porozumění otázce je zkresleno stejně jako naše.

(4) Je respondent vůbec s to poskytnout nám žádanou informaci?

To je opět velmi obtížné odhadnout. Řešením tu bude tentokrát pilotní studie, kupř. nestandardizovaný rozhovor na malém vzorku opět z cílové populace.

(5) Neptá se otázka na dvě různé věci najednou?

Ten problém „dvouhlavňové" otázky už známe.

Máte rád mléko a mléčné výrobky?
– ano
– ani rád, ani nerad
– ne

Jak mám odpovědět, když mám rád jogurt, miluji sýry, ale nesnáším mléko?

(6) Není naše otázka sugestivní?

Sugestivní otázka má takovou formulaci, která zvyšuje neúměrně výskyt určité varianty odpovědi. Sugestivnost otázky může být docela subtilní. Nebezpečí vyprodukovat vulgárně sugestivní otázku typu *„Doufám, že máte rád sýry, že ano?"* je opravdu minimální. Sugestivnost otázky může být vyvolána tím, že otázka výslovně zmiňuje jen jednu alternativu. Tak otázka *„Řekl byste, že máte rád sýr?"* je sugestivní a měla by znít asi takto: *„Řekl byste, že máte rád sýr, nebo byste spíše řekl, že jej nemáte rád?"* Sugestivnost otázky může být způsobena i nesymetrickým rozložením alternativních kategorií pro uzavřené otázky, jak jsme to ilustrovali ve cvičení 6.1. Pravděpodobně největší problémy jsou spojeny s používáním emotivně laděných pojmů. Nezáleží na tom, je-li citový náboj pozitivní, tento fakt rozhodne jen o směru zkreslení. Vhodná formulace nám někdy pomůže toto zkreslení překonat. Otázka *„Bijete někdy své děti?"* určitě vyvolá méně pozitivních odpovědí než otázka formulovaná třeba takto: *„Někdy se zdá opravdu nezbytné použít při výchově dětí místo domluvy mírný tělesný trest. Řekl byste, že souhlasíte, nebo spíše nesouhlasíte s tímto názorem?"* Velice silné zkreslení může vyvolat dovolávání se populárních autorit. Otázka uvedená *„Souhlasíte*

s názorem pana prezidenta, že... " by určitě vyvolala jinou distribuci odpovědí než otázka uvedená prostě *„Souhlasíte s názorem, že...*"

(7) Je výčet kategorií pro odpovědi na uzavřenou otázku úplný? Vylučují se vzájemně použité kategorie?

Splnění těchto podmínek by nemělo být obvykle příliš obtížné. Nejsme-li si jisti úplností výčtu našich otázek, přidáme prostě kategorii *JINÉ: Vypište prosím odpověď*. Nejlepším prostředkem, jak předejít překrývajícím se kategoriím, je použití kategorií tak specifických, jak je jen možno.

(8) Je použití otevřené otázky opravdu nezbytné?

Zkušenost ukazuje, že jenom zlomek otevřených otázek použitých ve standardizovaném kvantitativním výzkumu je opravdu použit pro finální analýzu. Je to jen přirozené. Největší silou otevřených otázek je jejich schopnost minimalizovat nezbytnou redukci informací. To je ovšem v protikladu s podmínkami kvantitativního výzkumu. Ten je schopen testovat hypotézy jen na omezeném počtu proměnných a tyto proměnné mohou mít jen velice omezený počet formálně dobře definovaných kategorií. Počet kategorií je omezen velikostí vzorku. Kdybychom měli příliš mnoho kategorií, v každé z nich by bylo jen velice málo jedinců. A představme si tabulku popisující souvislost mezi dvěma proměnnými s mnoha kategoriemi. Mnoho polí v takové tabulce by bylo prázdných nebo by obsahovala jen velice nízké četnosti. Statistika takových dat je pak nutně silně zkreslená.

Otevřené otázky v kvantitativním výzkumu musí být tedy zakódovány do omezeného počtu kategorií. Nejenže přitom ztratíme značnou část zaznamenané informace, ale proces kódování otevřených otázek je namáhavý a časově velmi náročný. Je to operace docela zajímavá, ale jen na začátku, tak pro prvních padesát dotazníků. Pro zbývajících řekněme 1450 je to pěkná otročina. Ale to už poznáte sami, až to budete jednou dělat. Zkrátka otevřené otázky bychom měli v kvantitativním výzkumu použít jen tehdy, kdy je to naprosto nezbytné, když opravdu nejsme schopni připravit předem uspokojující seznam možných kategorií. Potřebujeme-li příliš mnoho otevřených otázek, pak je to jen důkazem toho, že naše problematika není ještě zralá pro kvantitativní výzkum. V takové situaci bychom měli

uvažovat o aplikaci některých postupů **kvalitativního** výzkumu, který nám umožní generovat testovatelné hypotézy. A ještě jednu poznámku k tomuto bodu. To, co jsme zde řekli, neplatí o tazatelem kódovaných otázkách. Někdy, abychom zabránili sugestivnímu vlivu alternativních kategorií, je otázka předložena našim respondentům jako otevřená, ale tazatel zaznamená odpověď zaškrtnutím jedné z předem připravených kategorií.

(9) Nevyžaduje otázka nějaké příliš zobecněné údaje?

Otázky, které vyžadují nějaké „průměrné" údaje, mohou být pro respondenta velice obtížné. Mnohem lepší je přeformulovat tyto otázky do nějaké konkrétní formy. Kupř. na otázku „Kolik knih přečtete v průměru za měsíc?" bude respondent schopen správně odpovědět jenom výjimečně. Mnohem snazší bude odpovědět na otázku „Kolik knih jste přečetl za poslední čtyři týdny?" a ještě lepší by byla tato formulace: „**Které** knihy jste četl v posledních čtyřech týdnech?" Nejde nám přirozeně o to, zjistit skutečně tituly těchto knih. Kupř. odpověď „To byla taková historie o princovi, co potkal ducha a ten mu řekl, aby zabil strejčka" je stejně přijatelná jako odpověď obsahující přesné bibliografické údaje. Zamyšlení nad tituly dovede respondenty k daleko konkrétnější a realističtější odpovědi. K velice nekonkrétním a nespolehlivým odpovědím vedou otázky typu „Proč...?" Poznat skutečnou motivaci našich postojů a rozhodnutí je pro nás většinou nemožné. Objevit pravdivou motivaci je pro výzkumníka obtížný úkol a úkol rozhodně nemožný pro zkoumané osoby. Otázka „proč" může v respondentovi jenom zahájit určitý myšlenkový proces a patří spíše do domény kvalitativního výzkumu.

(10) Mají odpovědi všech respondentů stejnou váhu?

Někteří z našich respondentů mohou být o určitém problému lépe informováni než ostatní a tak jejich odpovědi mohou být závažnější. V takové situaci je vhodné použít tzv. filtrační otázky, které nám pomohou odlišit lépe informované respondenty od ostatních. Řekněme, že bychom studovali postoje obyvatelstva ke zdražení pohonných hmot. Zde by bylo důležité odlišit odpovědi těch, kdo vlastní auto, od odpovědí ostatních. Pravděpodobně bychom potřebovali získat od majitelů aut i nějaké další informace. Odpověď na filtrační otázku je pak jakousi

výhybkou. Negativní odpověď na otázku o vlastnictví auta vede tazatele k instrukci, aby přeskočil sérii otázek irelevantních pro daného respondenta.

(11) Není tato otázka nepříjemná, znepokojující, nemůže se respondent cítit ohrožen pravdivou odpovědí?

Takových otázek je mnoho; pro sociologický výzkum jsou často důležité informace, které většina lidí v denním sociálním styku neinzeruje, informace, které mají intimní charakter, informace o chování, které není zcela legální nebo není zcela etické, ale kterého se přece jen většina z nás alespoň občas dopouští. Získání takových odpovědí je zpravidla obtížné. Nicméně v empirickém výzkumu existuje řada postupů, které mohou alespoň zvýšit pravděpodobnost pravdivé odpovědi i na tyto obtížné otázky. Jejich podrobnější popis by vyžadoval celou obsáhlou kapitolu. Zde si můžeme tento problém jenom ilustrovat několika jednoduchými příklady různých technik, kterým věnujeme alespoň celý následující odstavec.

7.3 Jak se ptát na to, o čem se nemluví

Velmi důležitou roli tu hraje anonymita výzkumu, nebo přesněji přesvědčení zkoumaných osob o tom, že výzkum je skutečně anonymní. Respektování anonymity je základní požadavek výzkumné etiky; nicméně to samo o sobě nestačí. Pamatujeme si dobře, jak si v jednom z našich prvých výzkumů zkoumané osoby prohlížely dotazníky proti světlu, aby zjistily, zda archy nejsou identifikovány vodotiskem. Nicméně nedůvěra v anonymitu je problémem i v relativně velice svobodných zemích s dlouhou a nepřerušovanou tradicí výzkumu veřejného mínění.

Cvičení 7.1

Viděli jsme docela nedávno záznamový arch pro komerční výzkum v Ontariu. V záhlaví tohoto dokumentu, ze kterého tazatel četl otázky, byl vytištěn velkými červenými písmeny tento nápis: ***„Nezaznamenávejte, prosím, respondentovo jméno!"*** *Zamyslete se nad tím, jaká byla funkce tohoto nápisu.*

Další jednoduchý, ale rozhodně ne snadný trik je zamyslet se nad stylizací otázky a jejím slovníkem. Použití vhodného eufemismu může být velmi účinné. Místo „trestání dětí" se

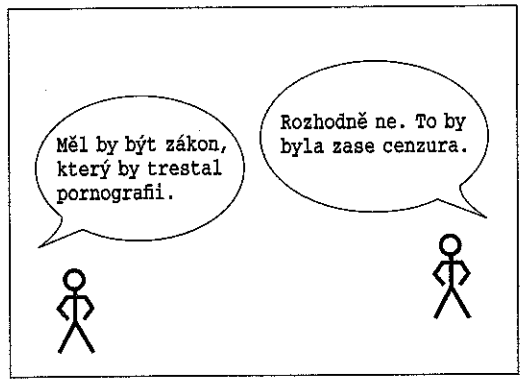
můžeme ptát na „výchovné metody", místo otázky o „rozkrádání" se můžeme zeptat, zda „lidé někdy odnesou něco z továrny". Otázka „Falšuje někdo z vašich spolupracovníků úkolové listy?" přinese jinou distribuci odpovědí než otázka „Vyplňují všichni vaši spolupracovníci úkolové listy docela správně?" Někdy můžeme použít záměrně sugestivní otázku, abychom kupř. vyvážili vliv konceptů spojených se silnou citovou náloží nebo překonali přirozenou tendenci volit „líbivější" odpověď. Dnes již je klasický příklad takové otázky, uvedený Maccobym a Maccobym (1956): „Domníváte se ve vaší rodině, že je třeba dávat si najevo vzájemnou náklonnost, nebo patříte spíše mezi lidi ukázněnější?" Ptáme-li se po postojích nebo jednáních spojených s určitým sociálním stigmatem, může být vhodnou následující strategie: Nežádáme respondenta, aby se „přiznal" kupř. k určitým aktivitám, které může považovat za závadné. Místo toho formulujeme otázku tak, aby navrhovala, že takové „problematické" chování je zcela běžné. Zde je několik příkladů: „Kdy jste naposledy kouřil marihuanu?"; „Kolik asi korun měsíčně věnujete na sázky?"

Pro dotazované osoby může být nepříjemné, že neznají odpověď na určitou otázku. Než by se k tomu přiznali, raději kupř. vyjádří postoj k něčemu, o čem nemají nejmenší tušení. Zde se vyplatí nabídnout respondentovi možnost čestného ústupu: „Naše televize uvedla nedávno nový vzdělávací pořad XV. Neviděl jste jej náhodou?" Na podobném mechanismu je založen postup, který se nám dobře osvědčil v několika zemích. Nepříjemná otázka je uvedena takto: „Nyní jinou otázku. Není příjemná, a když nechcete, nemusíte ji zodpovídat." Pak teprve následuje vlastní otázka, kupř. kolikrát byl respondent v minulém měsíci opilý. Nabídka možnosti neodpovědět vedla paradoxně ke snížení proporce osob, které odmítly odpovědět.

V podstatě vše, co jsme diskutovali v tomto paragrafu, by se dalo zahrnout pod záhlaví nepřímých otázek. Za nejúčinnější mezi nepřímými otázkami bývala považována kategorie projekčních otázek. V projekčních otázkách se neptáme zkoumané osoby na její vlastní názory, ale zdánlivě na postoje, mínění, postoje jiných osob, fiktivních postav, lidí obecně. Věříme však, že ve skutečnosti respondent promítne do této odpovědi své vlastní názory. Zřejmě je pro respondenta snazší vložit své vlastní předsudky třeba vůči rasovým menšinám do úst fiktivního „hrdiny" než je vyslovit sám za sebe. Projekční otázky mohou mít mnoho různých forem. Může to být třeba forma testu nedokončených vět. Úkolem dotazované osoby je doplnit větu, třeba: „Když si vedle Tomáše sedne v tramvaji Vietnamec, tak Tomáš vždycky..." Anekdotické otázky mohou mít formu krátkého příběhu. Zkoumaná osoba má

za úkol zvolit jedno z nabízených řešení příběhu, rozhodnout, který z fiktivních hrdinů má pravdu atd.

Projekční otázky jsou často předkládány ve formě jednoduché ilustrace. Má se za to, že ilustrace vyvolá ve zkoumaných osobách pocit nereálnosti, hry a nezávaznosti, a tím zvýší pravděpodobnost upřímné odpovědi. Kresby osob v těchto ilustracích musejí být co nejjednodušší. V procesu projekce vlastně vyžadujeme, aby se odpovídající ztotožnil s kterýmkoliv z hrdinů, aby jediným rozhodujícím faktorem byl názor, hodnoty, postoje, které tato fiktivní postava představuje. Kdyby kresba obsahovala nějaké zřetelné individuální charakteristiky (vousy, stavbu těla, věk), volba odpovědi by mohla být ovlivněna těmito faktory, a ne obsahem alternativy.



TY DVĚ OSOBY NA OBRÁZKU SE ZŘEJMĚ NEPOHODLY. KOMU BYSTE DAL ZA PRAVDU VY? TÉ OSOBĚ NALEVO, NEBO TÉ NAPRAVO?

Použití ilustrace v otázkách má ještě další význam: Dlouhý dotazník může být pro respondenty únavný a nudný. Každé zpestření, třeba ilustrace (ale i změna formy otázek), může oživit jejich pozornost.

Dr. Watson:
Tak vidíte, jak jste zaujatí proti dotazníku! Ilustraci můžete použít jenom v dotazníku, ale vůbec ne v rozhovoru.

Doktor Watson za to nemůže. Stvořili jsme si ho v naší knížce k tomu, aby se mýlil. Řešení je tu opravdu elementární. Tazatel může předložit vytištěný obrázek, ale i výčet alternativních

otázek (jak o tom budeme mluvit v příštím paragrafu) v průběhu rozhovoru. Ale teď se ještě na chvíli vraťme k projekčním otázkám. Nejsou bez problémů. Nemůžeme si totiž být jisti, že mechanismus projekce opravdu nastane. Již v roce 1956 Maccoby a Maccoby publikovali pěkný příklad toho, co se může stát. Mladá žena byla dotazována na postoje k zaměstnání. Mimo jiné dostala i tuto otázku: „Jak smýšlejí děvčata ve vašem úřadě o šéfovi?" Odpověď zněla: „Myslí, že je ohromný. Udělala by pro něho cokoliv." Ale obezřetný tazatel se zeptal (a mohl se zeptat, nejednalo se o standardizovaný rozhovor): „A jak se k němu stavíte vy sama?", a tazatel dostal tuto odpověď: „Já jím velmi pohrdám. Pokouším se kvůli němu změnit zaměstnání."

Validita projekčních otázek není tedy zaručena a je třeba si ji ověřovat i poté, co data byla sebrána. Jedním z možných postupů je technika kontrolních otázek. U důležitých proměnných bychom neměli spoléhat na jejich operacionalizaci do jediné projekční otázky, ale použít pro měření téže proměnné dvě nebo více různých operacionalizací, ať již ve formě projekčních nebo přímých otázek. Je rozumné umístit tyto otázky na taková místa dotazníku nebo rozhovoru, aby vzájemná interference mezi nimi byla minimální. Používáme-li techniku rozhovoru, je někdy vhodné použít postup nazývaný „random probe" (Schuman 1966). Každému z tazatelů je náhodně přidělena jedna nebo více ze souboru otázek, jejichž validita může být problematická. Po skončení interview se pak tazatel vrátí k přidělené otázce, třeba takto: „Před chvílí jste mi řekl, že rodina, která vlastní auto, je šťastnější než druhá rodina zmíněná v naší otázce. Co jste tím měl vlastně na mysli?" Takové spojení standardizovaných a nestandardizovaných postupů může někdy odhalit, že naše transformace konceptu do projekční otázky byla neplatná.

Projekční otázky byly alespoň po dvě desetiletí velikou módou. V poslední době jsou již méně populární. Jejich validita je již dlouho brána v potaz. (Viz kupř. Kidder a Campbell 1970.) Příprava projekčních otázek není snadná, kontrola jejich validity je časově náročná. Nicméně projekční otázky aplikované střízlivě a kriticky mohou pro nás být stále užitečné. Nepodceňujme ani jejich sekundární vlastnost: mohou podstatně oživit dlouhý a jednotvárný dotazník nebo interview.

7.4 Sociolog jako dramaturg

Žádná z otázek neexistuje sama o sobě. V dotazníku i v interview existuje v symbióze s ostatními otázkami. Otázka, která je, posuzována sama o sobě, bezvadná, může mít v konkrétním kontextu s ostatními otázkami opravdu zhoubný efekt. Podívejme se tedy alespoň stručně na některé problémy spojené s celkem otázek a s jejich pořadím.

(1) Zjišťují naše otázky opravdu všechny informace nutné pro testování pracovních hypotéz?

Někdy je opravdu snadné přehlédnout, že byla vynechána právě některá z velice jednoduchých, ale naprosto nezbytných proměnných. Opravdu se vyplatí strávit čas tímhle trochu nezábavným cvičením, a pokud se vám do toho nechce, možná že by bylo zdravé přečíst si znovu pohádku č. 7.

(2) Nemohou předchozí otázky zkreslit odpovědi na následující otázky?

Představme si, že bychom studovali problém, jakou roli hraje četba ve volném čase obyvatelstva. Řekněme, že bychom zahájili interview otázkami o četbě, ptali se třeba na oblíbeného autora, na preferovaný typ literatury, výdaje za knihy atd. Pak by následovaly otázky týkající se náplně volného času obecně. Mohli bychom se třeba ptát: „Vyjmenujte mi, prosím, tři věci, které děláte nejraději ve svém volném čase." Je vysoce pravděpodobné, že četba by byla zmíněna významně častěji, než kdybychom kladli otázky o volném čase nejdříve a otázky o četbě teprve potom. Takovému zkreslení se říká *haló efekt:* Každá otázka může ovlivnit odpovědi na všechny otázky následující. V dotazníku je situace ještě složitější: zde každá otázka – za předpokladu, že zkoumaná osoba přečte celý dotazník dříve, než jej počne zodpovídat – může ovlivnit odpovědi na **kteroukoliv** z ostatních otázek. Teoreticky by bylo snadné zabránit zkreslení haló efektem – alespoň v případě rozhovoru – velice jednoduše. Stačilo by předkládat otázky v náhodném pořadí.

Cvičení 7.2

Pro předkládání otázek v náhodném pořadí nepotřebujeme ani počítač. Stačí napsat každou otázku na tuhý kartotéční lístek a požádat tazatele, aby lístky před každým rozhovorem pečlivě zamíchal tak, jako by chtěl hrát mariáš s velice nedůvěřivými přáteli. A teď se prosím zamyslete, zda bychom tuhle techniku mohli skutečně použít.

Jak jste už jistě objevili, tato technika není skutečně použitelná pro změnu pořadí otázek. Ale je dobře použitelná a velmi užitečná, pracujeme-li s dlouhými výčtovými otázkami. Chtěli bychom třeba vědět, co zkoumané osoby dělají alespoň někdy ve svém volném čase. Taková otázka by měla pravděpodobně několik desítek položek. Vzhledem k oscilaci pozornosti zkoumaných osob prvé položky seznamu budou zkoumanými osobami posuzovány daleko pečlivěji než aktivity zmíněné uprostřed nebo ke konci seznamu. Zde je opravdu na místě použít naši techniku kartotéčních štítků. Na každém lístku bude vytištěno jméno jedné aktivity. Tazatel předá dobře zamíchané lístky zkoumané osobě s instrukcí, aby na jednu hromádku dávala lístky s aktivitami, které nikdy nedělá, a na druhou hromádku ty aktivity, které dělá alespoň někdy.

(3) Jsou otázky předkládány v nějakém logickém sledu?

Mnoho z nás má tendenci předkládat otázky v takovém pořadí, které je pro respondenty nějakým způsobem logické, a není matoucí a zneklidňující, jako by bylo třeba předkládání otázek v náhodném sledu. Pro většinu zkoumané populace nemusí být řazení otázek sledující logiku testování našich hypotéz relevantní. Někdy by přímé odhalení našich hypotéz mohlo vést ke zkreslení, naši respondenti by mohli vědomě nebo nevědomě volit takové odpovědi, které by záměrně podporovaly (nebo popíraly) naše hypotézy, místo aby vypovídali o svých názorech, aktivitách atd. Zejména hodláme-li použít hodně projekčních nebo jiných nepřímých otázek, je vhodné položit otázky zřetelně prozrazující náš výzkumný cíl až ke konci, kdy už tato informace nemůže ovlivnit odpovědi na projekční otázky.

Někdy takové pořadí otázek, které sleduje logiku našich hypotéz, není ani možné. Výzkumné organizace spojují do jediné terénní akce několik výzkumných problémů pro několik různých zákazníků. Tyto sdružené výzkumy mají často společnou jenom cílovou populaci. Tyto omnibusové výzkumy se stávají s rostoucí nákladností výzkumné práce stále obvyklejší. Navléknout několik sérií takových „piggyback questions", otázek, které hlavní výzkum nese „na koni" jako malé dítě, na jednotící logickou nit není tak obtížné, jak by se na prvý pohled zdálo. V podstatě totiž stačí, abychom v úvodu výzkumu (nebo ještě lépe v dopise, který informuje členy vzorku o výzkumu) naznačili účel výzkumu a aby otázky sledovaly jakoukoliv linii, bez nečekaných skoků a zákrutů, aby se tematika otázek neměnila náhle a bez jakéhokoli varování.

Obvykle stačí, aby nová tematická série otázek byla uvedena třeba takto: „A teď bych vám rád položil několik otázek o něčem jiném..." Jindy při změně tématu, zejména může-li se

dotazovaný cítit následující otázkou nějak ohrožen, je výhodné použít tzv. výhybkovou otázku. Je to jakýsi nárazník, který má předejít šoku z následující kritické otázky, nebo ho alespoň zmírnit. Kdysi v pravěku televize, kdy televizor byl ještě docela novou a nákladnou hračkou, Institut pro demoskopii v Allensbachu provedl rozsáhlý výzkum motivace vedoucí k nákupu televizoru. Z tohoto výzkumu jsme viděli jen nedatovaný tazatelský arch. Ten byl znamenitý, představoval skutečně mistrovskou aplikaci nepřímých otázek. Mezi nimi jsme našli jednu, která zněla asi takto: „V poslední době bylo tady v okolí mnoho vloupání. Slyšel(a) jste o tom také?" V kontextu výzkumu na celostátním vzorku je to otázka svým obsahem nesmyslná. Jejím smyslem zřejmě nebylo získat informaci, zda respondent slyšel o vloupáních, ale připravit tázaného na následující otázku: „Je váš manžel(ka) často mimo dům?" To už byla zcela funkční otázka, ale mohla by být mylně vykládána v situaci, kdy s tazatelem byl v místnosti přítomen jen respondent opačného pohlaví.

(4) Podporuje dramaturgie sledu otázek podmínky pro úspěšné a úplné dokončení rozhovoru/dotazníku?

Prvé otázky rozhovoru musí být snadné, musí získat důvěru dotazovaných osob, ale měly by být pro respondenta také zajímavé. Často ještě přežívá tradice začínat otázkami o demografických charakteristikách. To je relikt nejen z předpočítačové doby, ale z pravěku před etapou děrných štítků, kdy všechna třídění dat byla prováděna ručně. Protože demografické údaje jsou používány jako nezávisle proměnná nejčastěji, bylo výhodné mít je na nejlehčeji viditelném místě, na začátku prvé stránky dotazníku. Odpovědět na tyto otázky je pro respondenta snadné, ale tyto otázky jsou pro většinu lidí nudné a výzkum se hned začne podobat nějakému úřednímu aktu, třeba daňovému přiznání. Nadto mohou tyto otázky už od samého začátku vést k pochybám o anonymitě výzkumu. Němčina má pro správné úvodní otázky pěkné jméno: Eisbrechenfragen, tedy otázky, které prolamují ledy. Tak prvá otázka ve zmíněném výzkumu motivace pro nákup televizoru zněla: „Letos bylo hodně chřipky. Byl jste také nemocen?" Co je pro mne zajímavější než já sám a moje zdraví? Náklady na vytištění jedné otázky, zejména když data z ní nebyla analyzována, jsou zanedbatelné a odměnou je potenciálně lepší postoj dotazovaných k výzkumu.

Vhodná úvodní otázka může minimalizovat iritaci dotazovaných určitým výzkumným tématem. Na někoho, kdo padá pod tíží denních úkolů a čas, jejž věnuje rozhovoru, si bude muset trpce napracovat, mohou otázky o využití volného času působit jako vyložená provokace.

Proto Hrdý a Perglerová (1967) uváděli výzkum o kulturních zájmech takto: „Času máme v dnešní době málo. Denních starostí je tolik, že nám zbývá jen velmi málo času na naše zájmy, koníčky, zábavu i odpočinek. Přesto ale, co děláte nejraději ve svém volném čase?"

Méně zajímavé otázky (jako třeba ty, které zjišťují základní osobní data o respondentech) by měly být umístěny někde ve střední části rozhovoru, kdy se již respondent dostal do tempa, ale není ještě unaven (Noelle 1963).

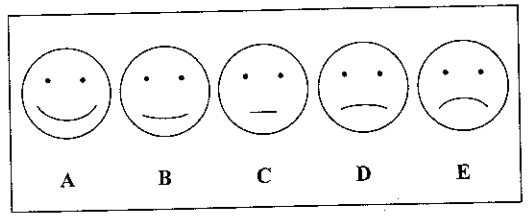
V pozdějších fázích rozhovoru bychom již také měli počítat s narůstající únavou zkoumaných osob. Tehdy je čas uvést nějaké zvláště zajímavé otázky, změnit jejich formu, použít třeba vizuální formu prezentace otázek. Řekněme, že pro jednu z našich hypotéz potřebujeme vědět, jak jsou zkoumané osoby spokojeny se životem. Mohli bychom se ptát, jak je respondent spokojen s rodinným životem, prací, bydlením, příjmem atd. Všechny otázky by měly asi tuto formu:

Když přemýšlíte o svém zaměstnání, řekl byste, že jste většinou

– velice šťastný
– spíše šťastný
– ani šťastný, ani nešťastný
– spíše nešťastný
– velice nešťastný

Stejnou informaci bychom mohli obdržet i jinou formou, která nejen může oživit respondentovu pozornost, ale má i další výhodu: nenutí respondenta verbalizovat odpověď o pocitech. Pro mnoho lidí může být nepříjemné **vyslovit**, že ve svém manželství jsou velice nešťastní.

Podívejte se, prosím, na následující obrázek. Jsou tu nakresleny tváře, které vyjadřují různé pocity. Pod každou tváří je napsáno písmeno.



Která z těchto tváří odpovídá nejlépe tomu, co pociťujete o svém zaměstnání? Je to tvář A, B, C, D nebo E?

A která tomu, co cítíte o svém manželství?

Podobně bychom mohli získat informaci i o jiných dimenzích spokojenosti se životem. A ovšem v podobné formě bychom se mohli dotazovat na mnoho jiných věcí. Samozřejmě nebudeme takové zajímavé formy otázek používat samoúčelně, jen pro oživení pozornosti. Ale máme-li důvod použít třeba anekdotické otázky, je vhodné je použít na takovém místě rozhovoru nebo dotazníku, kde cítíme, že dotazovaný už může mít našich otázek plné zuby.

Sémantický diferenciál je další technika, která podle naší zkušenosti zaujme pozornost většiny zkoumaných osob. Tato technika byla původně vyvinuta v roce 1956 Osgoodem, Sucim a Tannenbaumem. Respondent je požádán, aby popsal něco na několika škálách, obvykle s lichým počtem kategorií definovaných v protikladných párech slov; třeba *pomalý – rychlý, silný – slabý, špinavý – čistý* atd. Následující příklad je adaptací škály použité Levinem (1988) pro výzkum stereotypů o starých lidech.

Tabulka 7.2

	1	2	3	4	5	
pomalý						rychlý
vzdělaný						nevzdělaný
líný						pilný
silný						slabý
moudrý						hloupý
sobec						štědrý
přitažlivý						odpuzující
zámožný						chudý

Cvičení 7.3

Podívejte se ještě jednou na předchozí tabulku. Nebylo by výhodnější, aby každá ze stupnic byla kódována ***logicky*** *stejným směrem? To znamená, aby všechny negativní vlastnosti byly spojeny s nízkými a pozitivní vlastnosti s vysokými čísly?*

Sémantický diferenciál je velice účinným nástrojem pro zjištění toho, jaké stereotypní představy mají lidé o příslušnících jiných sociálních skupin, a tedy nástrojem k měření dimenzí předsudků, které máme. Tak bychom mohli studovat typickou představu Slováků o Češích a Čechů o Slovácích, stereotypy o příslušnících různých povolání. Mohli bychom studovat, jaká je vlastně naše dnešní představa o policii, o důstojnících, o lidech, kteří rychle

zbohatli v obchodě, o novinářích, popelářích, univerzitních profesorech, ale ne o sociolozích; nikdo neví, co to vlastně je. Mohli bychom studovat, zda a jak familiarita se studovanou skupinou může změnit naše stereotypy, a vůbec celou řadu sociologicky relevantních problémů. Použití sémantického diferenciálu je ovšem mnohem širší. Je často používán v marketingu. Studium stereotypních představ o určitém produktu může optimalizovat strategie marketingu. Zkrátka se zdá, že můžeme souhlasit s názorem Osgooda a ostatních původních autorů, že toto měření je jenom málo ovlivněné druhem objektu, který je studován, nebo typem respondentů, na které je tento nástroj aplikován.

A teď se velice krátce podívejme na závěrečnou fázi rozhovoru nebo dotazníku. V té je respondent unaven, znuděn, ale také spokojen, že konec je už blízko. To nám dává jistou svobodu, jaké otázky umístit do této fáze. Tady se také hodí dobrá rada, kterou nám kdysi dal profesor Hyhlík. Někdy musíme použít takovou otázku, která může některé respondenty rozzlobit natolik, že odmítnou v rozhovoru pokračovat. Je proto dobré umístit takovou otázku až na konci rozhovoru, tak abychom byli vyhozeni se všemi ostatními daty již zaznamenanými. Ale lepší je respondenta nerozčílit, nebo ho alespoň usmířit. Po nás může k němu jednou přijít jiný tazatel...

(5) Podporuje stavba rozhovoru nebo dotazníku představu o anonymitě dotazníku?

Nejsou v dotazníku nějaké nevysvětlené kódy, které by respondent mohl považovat za nástroj k jeho identifikaci? V době, kdy každá rozmnožovací blána musela být registrována a opatřena identifikačním číslem, jsme strávili velmi mnoho času vysvětlováním nedůvěřivým respondentům, že toto číslo slouží k identifikaci blány, a ne respondenta.

(6) Není rozhovor nebo dotazník příliš dlouhý?

Skoro určitě je! Ten náš věčně přítomný nepřítel, velikost přirozených systémů v sociálních vědách, nás nutí použít více proměnných, než jsme skutečně schopni měřit. Pokud naším výchozím bodem byly pracovní hypotézy, neměli bychom mít v našem nástroji zbytečné proměnné. Ale nebylo by možné vypustit některou z našich hypotéz a tak dotazník/rozhovor zkrátit? U rozhovoru je otázkou základní slušnosti informovat respondenta, jak dlouho bude

interview trvat, a informovat ho o tom pravdivě. Nepravdivá informace je iritující a za naši bezohlednost zaplatí všichni výzkumníci, kteří by chtěli zkoumat naše respondenty v budoucnosti. Zamoření terénu se v sociologii stává stále závažnějším problémem.

Pohádka pro odrostlejší děti 16

O indiánech, dortu „Začátek Školního Roku" a o zamořeném terénu

Máme českou přítelkyni, která se seznámila s tlupou Hopi indiánů, žijící poblíž jednoho velkého města. Měli Zuzanu velice rádi, přijali ji téměř za vlastní. Zvali ji na rodinné slavnosti a upekli jí dokonce dort, zvaný Začátek Školního Roku. Byl prý růžový a velice sladký. Snad aby osladil indiánským dětem smutek spojený s nutností odejít z rezervace do cizí, ne právě přátelské internátní školy. Jenom když se indiáni dověděli, že její manžel pracuje na univerzitě, musela slíbit, že manžela a kohokoliv, kdo má něco společného s univerzitou, nikdy do rezervace nepřivede: „Víte, ti lidi z univerzity jsou takový divný. Přijdou s blokem nebo s magnetofonem a ptají se takové nesmyslné otázky. To my nemáme vůbec rádi."

(7) Jsou odpovědi na otázky jedinou informací, kterou má tazatel zaznamenat?

Prakticky vždycky potřebujeme, aby nám tazatelé poskytli nějakou další informaci. Záznam času, kdy byl rozhovor zahájen a skončen, trvání rozhovoru, počet neúspěšných kontaktů, které musel tazatel učinit, než se mu podařilo kontaktovat vybranou osobu, to jsou nejenom informace relevantní pro mzdového účetního, ale i důležité informace pro organizaci budoucích výzkumů a mohou nám někdy říci i leccos o charakteristikách populace. Informace o tom, zda byla rozhovoru přítomna ještě jiná osoba než tazatel a dotazovaný, může být nápomocná pro vyhodnocení validity odpovědí daného respondenta. Chceme-li, aby tazatelé provedli „random probe", nesmíme zapomenout zahrnout nutnou instrukci, včetně identifikace otázek, které byly tomu kterému tazateli náhodně přiděleny.

Někdy od tazatelů také vyžadujeme, aby provedli nějaká jednoduchá pozorování. Informace o domě nebo bytu, ve kterém zkoumaná osoba žije, mohou hrát významnou úlohu pro objektivnější odhadnutí sociální třídy zkoumaných osob. V našem výzkumu postojů starých Italů, Portugalců a rodilých Kanaďanů k možné institucionalizaci (Disman & Disman 1989) tazatelé zaznamenávali, v jakém jazyce byl rozhovor proveden. Dále tazatelé ocenili na

jednoduché, krátké stupnici respondentovu schopnost komunikovat v angličtině a posouzení etnického charakteru bezprostředního okolí respondentova domu nebo bytu.

Je samozřejmé, že záznamový arch musí poskytnout příslušné rubriky pro záznam takových informací, opět pokud možno ve velice jednoduché standardizované formě.

7.5 Když se sociolog nevtírá

Přirozeně zde budeme mluvit – a opět jenom velmi stručně – o neobtrusivních metodách, o pozorování a o studiu dokumentů.

Standardizované pozorování

Akt pozorování – alespoň potenciálně – neinterferuje se studovaným systémem. To ovšem platí jen tehdy, když pozorované osoby nevědí, že jsou sledovány. Známe různé techniky, jak můžeme pozorovatele skrýt, od skryté kamery přes jednostranně průhledné zrcadlo až po sociální ukrytí pozorovatele. Pokud tyto techniky fungují, největším problémem je etičnost takové akce. Ale o tom budeme hovořit ve zvláštní kapitole. Zde jenom připomeňme, že neobratné použití technických pomůcek jako zrcadel nebo skryté kamery, i když nevede nutně k fyzickému ataku na pozorovatele (ale i to se stává), vyvolá takové zkreslení, že zcela otevřené pozorování by bylo lepší. (O zkreslení vyvolaném tím, že zkoumané osoby vědí, že jsou pozorovány, jsme již hovořili.) Pravděpodobně nejspolehlivější je sociální ukrytí tazatele. Pozorujeme-li, jak se třeba chovají lidé ve frontě v samoobsluze, na přednášce, v tramvaji, na fotbale, není obvykle obtížné, aby pozorovatel přijal roli zákazníka, posluchače, sportovního fandy atd. Je to jedna z forem zúčastněného pozorování, o kterém budeme mluvit daleko podrobněji v souvislosti s technikami kvalitativního výzkumu.

Máme-li použít výsledků pozorování v kvantitativním výzkumu, standardizace pozorování je nezbytná. Bez ní bychom nemohli zaručit srovnatelnost dat od různých pozorovatelů. Řekněme, že bychom chtěli testovat hypotézu, že studenti, kteří sedí ve velkých univerzitních aulách v zadních řadách sedadel, sledují přednášky s menší pozorností než studenti, kteří sedávají vpředu. Navrhovaný postup by byl použitelný jen ve velké třídě a na začátku školního roku, když se posluchači ještě neznají. Mladí pozorovatelé by se pak nelišili od ostatních studentů. Každý pozorovatel by měl přidělenou skupinu sedadel a zaznamenával

by chování studentů, kteří by seděli na těchto místech. Záznamový arch by obsahoval plán sedadel, který byl pozorovateli přidělen, a chování studentů by bylo zaznamenáváno v předurčených časových intervalech, řekněme každých deset minut. Pro každou z těchto period by pozorovatelé obdrželi zvláštní formulář. Chování studentů v každé periodě pozorování by bylo kódováno do připravených kategorií, které by mohly vypadat asi takto:

(1) Sleduje přednášejícího, dělá si poznámky.
(2) Občas sleduje přednášejícího, poznámky dělá jen sporadicky.
(3) Občas si dělá poznámky, ale také se baví se sousedem, nebo si čte.
(4) Je pasivní.
(5) Dřímá, usnul.
(6) Odešel.

Takovéhle uspořádání výzkumu zní docela pěkně a snad i elegantně. Ale pozor, mohou zde být různé, obratně zamaskované pasti. Tak se proto podívejme na pohádku číslo 17.

Pohádka pro odrostlejší děti 17

O schůzích a křížovkách

To bylo ještě v době, kdy umění, jak přežít nekonečné schůze a zůstat duševně zdráv, patřilo k základním životním nutnostem. Zkoušeli jsme všechno. Až jednoho dne přišel náš přítel Míša Pergler s nápadem vpravdě geniálním: „Křížovky!" Ne tedy je luštit, to je příliš viditelné, ale vytvářet je. Po prvých několika slovech to není tak snadné. Musíte se zamyslet, tak se třeba zamyšleně podíváte na přednášejícího. Pak vás něco napadne, napíšete další slovo a legendu, lektor je potěšen posluchačem, který je pozorný a dokonce si dělá poznámky, a ty dvě tři hodiny schůze utečou přece jenom snesitelněji.

Asi tři dny po tomhle vynálezu jsem jel přednášet do Košic. Na jakési divné, nudné téma. V poslední řadě seděl mladý muž. Velice pozorný, dělal si poznámky, když nepsal, sledoval mě zamyšlenýma očima. To bylo přece jen trochu mnoho. Protože se při přednáškách rád procházím mezi lavicemi, zamířil jsem při nejbližší příležitosti do severozápadního rohu místnosti, kde seděl ten pozorný mladý muž. Samozřejmě tvořil křížovku, docela velkou a složitou. Tak jsem mu o přestávce pogratuloval, jak tvořivě využívá čas. Po lehkém začervenání odmítl skromně moji chválu: „Ale to dělají v Praze už dlouho..."

Poučení:
Za prvé, zkušenosti se diseminují v naší vlasti zatraceně rychle, zejména jde-li o něco, co není tak docela košer. A za druhé, interpretace chování může být někdy velice mylná.

Zatímco pozorování neruší studovaný proces, a v tom je jeho veliká síla, je spolehlivým nástrojem jen pro zkoumání takových jevů, které můžeme přímo vnímat. Spoj mezi pozorností a chováním jako jejím ukazatelem nemusí být vůbec jednoduchý. Zatímco záznam chování

může být poměrně přesný, jeho interpretace v termínech pozornosti může být neplatná. Co když ten student, jehož chování jsme zakódovali do jedné z nejnižších kategorií pozornosti, je tak zaujatý, že zavřel oči, aby se mohl lépe soustředit? Je-li však chování přímo tím, co chceme studovat, pozorování je jedinečným nástrojem a řada typů chování je sociologicky relevantní. Tohle jsou jenom náhodně vybrané příklady: respektování dopravních pravidel za různých okolností, chování ve frontě, stereotypy chování v restauraci a mnoho jiných. Pozorování je nástrojem používaným velice rutinně v experimentu, více v sociální psychologii a psychologii než v sociologii. Hmotné artefakty jsou nesporně sociálními fakty a pozorování je zde takřka jediným použitelným nástrojem. Už jsme se zmínili, že téměř vždy požadujeme od tazatelů, aby doplnili rozhovor záznamem určitých pozorování. Lidské výtvory jsou často jediným zdrojem sociální informace o minulosti. Archeolog je náš blízký bratranec. Ale zde se už dostáváme k druhé významné skupině neobtrusivních technik.

Studium dokumentů

Jak jsme si již řekli, za dokument tu považujeme jakýkoliv hmotný záznam lidské činnosti, který nevznikl za účelem našeho výzkumu. Tak za dokumenty bychom mohli považovat třeba

- úřední statistiky,
- daňové záznamy,
- nápisy na stěnách veřejných záchodků,
- míru opotřebování koberce v muzeích a galeriích,
- záznamy snů,
- osobní deníky,
- obsah nádob na smetí,
- romány,
- telefonní účty,
- plakáty a letáky,
- novinové články a stovky jiných...

Dr. Watson:

Tak teď opravdu doufám, že je to jenom hloupý vtip. Copak se může vědecká analýza spolehnout na něco tak fiktivního, jako jsou romány? A noviny říkají pravdu jenom příležitostně. Já jsem v podstatě čestný člověk, ale přece i v mých dopisech je hodně sladkých lží, a když jsem si psal deník, tak to byla Wahrheit und Dichtung...

Tak tentokrát bude mít odpověď našemu příteli poněkud dialektický charakter. Ve všem, o čem se Watson zmínil, má pravdu. Ale i zkreslené dokumenty nám mohou poskytnout velice

cennou informaci. Někdy jsme schopni odhadnout směr zkreslení, jindy je to právě zkreslení, které je předmětem našeho výzkumného zájmu.

Klasický příklad není z oblasti sociologie, ale z romantického světa výzvědné služby. Za druhé světové války britská zpravodajská služba získávala nejspolehlivější informace o stavu morálky německého obyvatelstva ne od agentů, ale analýzou notoricky zkreslené nacistické propagandy. Znalost získaná systematickou analýzou zkreslení umožnila předpovědět na základě rozboru projevů vedoucích nacistických funkcionářů, že Hitlerovy tajné zbraně nejsou pouhou propagandou, odhadnout čas, kdy budou asi použity, a dokonce odhadnout i typ těchto zbraní. O britské vědecké zpravodajské službě existuje znamenitá kniha jednoho z jejích tvůrců, R. V. Jonese, *The Wizard War: British Scientific Intelligence 1939–1945* (1978). Je to kniha, kterou budete číst se zatajeným dechem, a pro nás je to geniální úvod do analytického myšlení právě toho typu, jaké potřebuje sociolog.

Jedna z nejdůležitějších studií klasické sociologie o sociálních změnách vyvolaných emigrací je velice rozsáhlá (čtyřsvazková) studie Williama I. Thomase a Floriana Znanieckého *Polish Peasant in Europe and America*, prvý svazek publikován v New Yorku 1917. Tato studie je založena výhradně na analýze soukromých dopisů, které si vyměňovali vystěhovalci z polských vesnic do USA se svými příbuznými ve staré vlasti.

O hmotných stopách chování jsme se již zmínili. Ale opotřebení koberce je jen jeden příklad z mnohem širšího spektra. Popularita lokálních FM rádiových stanic v Chicagu byla studována za pomoci autodílen a „service stations". Ty zaznamenávaly nastavení stupnice na přijímačích vozů zákazníků. Za rok byla získána – velice levně – informace o více než 50 000 vozech. Popularita knih v knihovnách, zejména referenčních materiálů, které se nepůjčují, může být odhadnuta stupněm opotřebení jednotlivých svazků. Popularita inzerce byla dokonce měřena počtem různých souborů otisků prstů na daných stránkách (DuBois 1963). Analýza odpadků byla s úspěchem použita pro analýzu vzorců spotřeby alkoholu. Analýza nápisů ve veřejných záchodcích byla často použita ve studiu lidské sexuality, mimo jiné i v klasické Kinseyově studii (Kinsey et. al 1953).

Nejčastěji se však setkáme s analýzou psaných záznamů. Nikdo v sociálních vědách asi nepochybuje o užitečnosti sčítání lidu a statistických ročenek. Méně automaticky si uvědomíme, že máme k dispozici podobné údaje i z dávné minulosti, jako třeba berní role.

Je zřejmé, že některé úřední záznamy mohou obsahovat systematické zkreslení. Nemáme zde na mysli jen propagandistické zkreslování dat, ale i zkreslení vyvolané různou metodologií použitou pro různé záznamy a různými kulturními vlivy. Kupř. se zdá, ze poměrně vysoké číslo sebevražd uváděné ve statistikách skandinávských zemí je alespoň částečně funkcí kulturních rozdílů. Sebevražda v těchto zemích není spojena s tak silným negativním stigmatem jako v Anglii nebo Severní Americe. Zde lékaři mají silnou tendenci s ohledem na přežívající rodinu – je-li to jen trochu možné – hlásit „náhodné úmrtí", a ne sebevraždu.

V podstatě jakýkoliv psaný materiál je potenciálním zdrojem sociologické informace. Není proto divu, že byla vytvořena speciální technika pro analyzování obsahu sdělení:

Obsahová analýza

je kvantitativní, objektivní analýza sdělení jakéhokoliv druhu. Obsahová analýza se může zabývat právě tak obsahem sdělení jako jeho formou, autorem i adresátem takového sdělení.

Sféra uplatnění obsahové analýzy je nesmírně široká. Může být použita pro srovnání, jak objektivně různé masové sdělovací prostředky referují o určitých událostech. Obsahová analýza byla často použita pro identifikaci autora. V notoricky se vracející diskusi obsahová analýza spočívající na rozboru jazykových elementů znovu potvrdila, že Shakespeare byl opravdu Shakespeare, a ne Bacon. Obsahová analýza je běžně používána pro objektivní určení obtížnosti textu pro čtenáře. Může být použita nejen pro definování, ale i pro odhalení skrytých úmyslů sdělovatele. Obsahová analýza byla použita pro rozlišení mezi autentickými dopisy sebevrahů a simulovanými dopisy psanými těmi, kteří sebevraždu opravdu nezamýšleli (Ogilvie, Stone a Shneidman 1986). Obsahová analýza má i uplatnění uvnitř jiných výzkumných technik. Může být použita kupř. pro zpracování dlouhých otevřených otázek a je i důležitým nástrojem pro organizování záznamů získaných kvalitativními postupy.

Typickou doménou pro aplikaci obsahové analýzy je studium trendů na různých úrovních sociálního vývoje. Kupř. Naisbittův bestseller *Megatrends* (1982) se pokusil definovat hlavní trendové změny v současných Spojených státech. Východiskem pro tuto studii byla obsahová analýza více než šesti tisíc lokálních novin. V tomhle výčtu bychom mohli pokračovat ještě dlouho.

A jak se to dělá? Logika této analýzy je velmi přímočará. Organizačně obsahová analýza sleduje stejné logické schéma jako každý jiný kvantitativní výzkum. I zde budou východiskem pracovní hypotézy. Ty nám pomohou racionálně rozhodnout o populaci a vzorku.

A protože už slyšíme brumlat našeho Dr. Watsona někde v pozadí, podotkněme, že populace zde není soubor lidí, ale soubor sdělení. Je-li objem těchto sdělení nezvládnutelně veliký, je samozřejmé, že jednotky, které podrobíme analýze, nemohou být vybrány nazdařbůh, ale podle pravidel konstrukce vzorku, tak jak jsme to diskutovali v kapitole 5. Jenom si zde připomeňme, že musíme být obezřelí při použití systematického výběru. Řekněme, že bychom studovali noviny. Kdyby zvolený krok byl násobkem sedmi, všechny vybrané výtisky by byly z téhož dne v týdnu, a kdyby to byla neděle, naše výsledky by byly obzvláště zkreslené.

Ovšem nejdůležitější funkce pracovních hypotéz je jinde: Obsahová analýza je definována jako **objektivní** metoda analýzy sdělení. To znamená, že plán analýzy musí být takový, že budou-li dva nebo více výzkumníků nezávisle analyzovat stejný materiál, musí dojít k shodným výsledkům. Tedy operační definice musí být velice jasné, jednoznačné a někdy i dosti podrobné sdělení. Konstrukce dobrých kategorií, do kterých bude obsah sdělení kódován, je asi nejcitlivější operací v procesu obsahové analýzy.

Obsahová analýza je dále definována jako **kvantitativní** metoda. Součástí projektu pro obsahovou analýzu musí být definice kvantitativních jednotek, ve kterých bude měřen rozsah sdělení, který patří do určité kategorie. Tyto kvantitativní jednotky bývají definovány na dvou úrovních:

(1) **záznamové jednotky**, ve kterých opravdu měříme rozsah, a

(2) **kontextuální jednotky**, které potřebujeme tehdy, kdy záznamové jednotky jsou poměrně malé a izolovaný obsah záznamové jednotky by mohl být zařazen do nesprávné kategorie. Kupř. když záznamovou jednotkou je věta, ironické tvrzení („A Brutus jistě ctihodný je muž") by bylo kategorizováno jako chvála. Kontextuální jednotka je rozsahem širší než záznamová a je východiskem pro kategorizaci.

Jednotky měření existují v celé řadě forem. Podívejme se na ně podrobněji. Následující tabulka podává příklady záznamových jednotek v několika z možných přístupů.

Tabulka 7.3

Typ jednotky	Příklady
Jednotky definované rozsahem:	slovo, tvrzení, věta, řádek, odstavec, článek, kapitola, délka sloupce v centimetrech, pro televizi a filmy trvání v minutách atd.
Zdroje, autority:	citované zdroje informace jako tiskové agentury; autority, jichž se sdělení dovolává apod.
Námět, téma:	problémy zdůrazňované politickou stranou v předvolební kampani; zápletka v románu, povídce, filmu, dramatu; geografická lokalizace děje
Hrdina:	osoby objevující se ve sdělení; hrdinové fiktivních děl atd.

Tohle opravdu nezní složitě, ale praxe může být někdy komplikovaná. Chtěli bychom třeba zjistit, zda se české a slovenské noviny liší v roce 1991 v tom, jak intenzivně je v nich referováno o separatistických hnutích v Evropě a jinde ve světě. Řekněme, že nám šlo jenom o to, kolik pozornosti je věnováno tomuto tématu, ne o postoje k těmto hnutím, pokud je noviny vyjadřují. Rozhodli jsme se, že do naší populace zařadíme jenom deníky, které jsou pravidelně distribuovány alespoň v celé jedné z obou federálních republik. Za rok by mohly tyto noviny představovat takový objem informace, který bychom nemohli z časových důvodů zvládnout, a bylo by nutné vytvořit vzorek. Konstruovat pravděpodobnostní vzorek pro každé noviny zvlášť by mohlo být riskantní: pro některé noviny bychom mohli náhodně zvolit třeba den, kdy chorvatskému kandidátovi bylo zabráněno nastoupit do pozice prezidenta, a pro jiné noviny bychom vybrali den, kdy se v oblasti separatistických hnutí neudálo nic důležitého. Lepší by bylo vytvořit náhodný vzorek dat a analyzovat všechny noviny z naší populace publikované ve vybraný den.

Nejjednodušší by bylo zvolit jako záznamovou jednotku článek. To by podstatně zvýšilo produktivitu naší analýzy, ale malý sloupeček by měl stejnou váhu jako půlstránkový článek. Bylo by proto výhodné doplnit naše měření další, podrobnější jednotkou rozsahu. Poměrně rychlé výsledky nám dá měření rozsahu článku v délce sloupců v centimetrech. Stačí připravit jednoduchou tabulku, která nám umožní kontrolovat šířku sloupce a velikost písma

a převést konkrétní výsledky na normalizované jednotky. Tak třeba 8 cm sloupce novin se rovná 6 cm normalizovaného sloupce.

Cvičení 7.4

V naší analýze bychom mohli dostat třeba tyto výsledky: Zatímco noviny A věnovaly naší problematice 360 cm sloupců, noviny B jenom 200 cm. Můžeme z toho usoudit, že noviny A skutečně věnují separatistickým hnutím více pozornosti než noviny B?

Téměř bez výhrad bychom měli sledovat následující pravidlo: Srovnáváme-li kvantitu mezi dvěma sděleními, vyjádření této kvantity nestačí a musíme je doplnit informací, jakou proporci z celého zdroje daná kvantita představuje. To je zvláště důležité při mezinárodním srovnávacím výzkumu: typické americké víkendové vydání novin může vážit asi tak půl kila.

Ale i pak srovnání může být zavádějící. Je rozdíl mezi článkem na titulním listě a článkem skrytým někde uvnitř, mezi inzeráty. Závažnost sdělení může být indikována typografickou úpravou, přítomností fotografií, v rozhlase a televizi zvukovými efekty atd. Někdy je prostě třeba doplňovat informaci o rozsahu také záznamem o existenci takových atributů.

Pro analýzu fiktivních výtvorů běžné kvantitativní jednotky často nestačí. Před časem jsme se studenty v mém kurzu vypracovali projekt obsahové analýzy, která by mohla odpovědět, zda změněná pozice žen v americké ekonomické struktuře a rapidně sílící feministické hnutí měly vliv na smýšlení širokých vrstev a zda se to také odráží na obsahu Harlequin Romances. To je typická braková literatura pro ženy, nesmírně populární. Každý rok jsou v těchto edicích publikovány stovky titulů.

Dr. Watson se horlivě hlásí:

Proč pro tenhle účel nevybrat skutečnou literaturu? Ta přece jen zrcadlí skutečnost mnohem hlouběji a pravdivěji. Proč vybrat brak, který si neklade žádná estetická kritéria?

Vy už asi víte, že odpověď je zcela elementární. Cílem výzkumu tohoto typu není sdělení samo, ale **adresát**. Edice jako Harlequin Romances si neklade jiné cíle než maximalizovat prodej jednotlivých titulů. Obchodní úspěch této velice masové edice je důkazem, že autoři a zejména editoři mají velice dobrou představu o tom, co masový čtenář vyžaduje. Čili pro výzkum adresáta je brak tohoto typu ideálním médiem.

Výzkum zaměřený na sociální změnu vyžaduje srovnání. Pro náš účel by bylo asi nejekonomičtější vybrat jeden náhodný vzorek ze svazků této edice publikovaných v roce 1981 a druhý vzorek z knih publikovaných v roce 1991. Daleko obtížnější by bylo operacionalizovat sociální změny v hodnotách. Jistě by bylo významné studovat kupř. změny ve vzorcích komunikace mezi hlavním mužským a ženským hrdinou, kategorizovat je třeba v dimenzích dominace a submisivnosti, aktivity a pasivity v tomto vztahu apod. V tom případě by bylo nutno použít motiv, téma jako záznamovou jednotku. To je pak z hlediska reliability klasifikace velice obtížné. Nebo by bylo možno z těchto hledisek analyzovat dialogy obou hrdinů. Záznamovou jednotkou by tu byla pravděpodobně věta, kontextuální jednotkou třeba odstavec. Operacionalizace by zde byla snazší a reliabilnější, ale pracná. Nejlepší by pravděpodobně bylo vytvořit jakési slovníky nebo podrobné příklady agresivních frází, submisivních frází atd. V obou zmíněných postupech by sama analýza byla časově velice náročná. Na analýze jediného svazku bychom museli strávit řadu hodin. Výzkum tohoto typu by byl jistě významný, ale po mnoha stech, ne-li tisících hodin bychom mohli zjistit, že žádné změny v Harlequin Romances v posledních deseti letech nenastaly.

Z těchto důvodů bylo třeba zvolit jiný postup. Obsah těchto knih je velice jednoduchý. ON je pilot, majitel plantáže v nějakém velice exotickém místě, vynálezce, majitel historické sklárny na benátském Muranu a držitel staletí starého tajného receptu na výrobu muranského skla nebo lékař. (V tomto případě to bude téměř vždy věhlasný chirurg a velice často holandský věhlasný chirurg. Proč holandský, nevím, ale je to empirický fakt.) Škála povolání mužských hrdinů je široká, ale vždy to jsou povolání velice romantická a prestižní. Ještě jsme neslyšeli o tom, že by hrdinou byl sociolog.

ONA je téměř vždy sirotek. Její rodiče zahynuli (a) při železničním neštěstí nebo (b) při letecké katastrofě nebo (c) při autonehodě. Její povolání spadá nejčasteji do jedné z následujících kategorií: (a) společnice staré, zámožné dámy, (b) učitelka, (c) diplomovaná sestra, (d) její povolání není zmíněno. Zápletka je dojemně prosťoučká. ONA miluje JEHO, ale ONA netuší,

že pod drsnou skořápkou JEHO chování se skrývá milující srdce. ON ji pozve, jako svoji nejzručnější asistentku na operačním sále, s sebou do Holandska a ubytuje ji u své stařičké maminky. Zde se vynoří tmavovlasá cizinka, které se ON přespříliš věnuje. Několik stran před koncem se ukáže, že tato cizinka je (a) jeho sestra nebo (b) zloduška, kterou ON už dávno odhalil. V posledním odstavci ji, chvějící se, sevře do své mužné náruče a zašeptá do jejích plavých vlasů: „Vy pošetilý blázínku, já vás miloval od prvého okamžiku."

Pokud se dosti dramatické změny v pozici žen v socioekonomické struktuře opravdu odrážejí ve vědomí průměrného ženského čtena americké nižší střední třídy, mělo by se to odrážet na socioekonomickém statusu fiktivních hrdinů tohoto druhu literatury. Kdyby byly následující pracovní hypotézy potvrzeny obsahovou analýzou, potvrdily by, že změny v sociální skutečnosti **nejsou** odráženy v obsahu Harlequin Romances:

- Povolání hrdinky se omezuje na typicky ženská povolání.
- Počet povolání ženských hrdinů je mnohem nižší než škála povolání připisovaná mužským hrdinům.
- Status povolání mužského hrdiny je vždycky významně vyšší než status povolání hrdinky.

Je zřejmé, že pro testování takovýchto hypotéz použijeme jako záznamovou jednotku **„hrdinu"** a že budeme používat jenom dvě proměnné, povolání hrdiny a povolání hrdinky. S použitím standardních škál prestiže povolání by bylo snadné transformovat povolání do číselných hodnot. Produktivita takové analýzy by byla značná. Podle naší zkušenosti bylo možné vyhodnotit po určité době zácviku 4–6 svazků za hodinu.

V našem miniaturním výzkumu provedeném na miniaturním a zcela nereprezentativním vzorku se ukázalo, že prvé dvě z hypotéz mohly být zamítnuty. Zdá se tedy, že ani harlekýnské romance nemohou zcela ignorovat sociální změny. Ale i kdyby naše výsledky byly opravdu reprezentativní, týkají se jenom malé části problému. Neřekly by nám mnoho o zápletce a v posledním odstavci by ji ON opět sevřel do své mužné náruče a zašeptal do jejích plavých vlasů: „Vy pošetilý blázínku, já vás miloval od prvého okamžiku." Jediný rozdíl by byl v tom, že by šeptal do plavých vlasů slibné teoretické fyzičky. Abychom pronikli hlouběji do problému, museli bychom použít některý z mnohem pracnějších postupů, zmíněných na začátku naší diskuse o Harlequin Romances.

Studium dokumentů obecně a obsahová analýza zvláště jsou jedinečným nástrojem analýzy a pochopení společenských změn. Jsem si jist, že zatímco píši v Torontu tyto řádky, někde u nás někdo srovnává hodnoty Pražského jara s hodnotami Sametové revoluce tak, jak se odrazily v jazyce masových médií. Někdo jiný třeba studuje, jak určité historické události a historické postavy mizely a opět se objevovaly v několika posledních desetiletích v učebnicích historie pro naše střední školy. Studium změn v profilu našich novin v posledních několika letech, citované autority v profesionálním tisku, zdroje informací o cizině dnes a před třemi lety, to jsou jen namátkově zmíněná témata z mnoha a mnoha desítek problémů z naší sociální skutečnosti, ke kterým by obsahová analýza byla nejvhodnějším klíčem.

Obsahová analýza je jednou z nejproduktivnějších mezi ostatními technikami sběru informací. Je daleko méně nákladná než kterákoliv z jiných technik, dá se často zvládnout v malém pracovním týmu a někdy i jednotlivcem. Zdaleka však ne vždy, zejména když je materiál, který má být analyzován, rozsáhlý, zvolený vzorek velký a záznamové jednotky malé. Víme o případu studenta, který musel čekat, až v Pánu zesne jeho profesor. Ten profesor byl vyhlášený teoretik, ale jeho představy o realitě empirického výzkumu byly spíše mlhavé, a naléhal tedy, aby tento student použil obsahovou analýzu takovým způsobem, kde jenom hrubá analýza by vyžadovala několikaletou práci početné výzkumné skupiny. Jistě, stalo se to ještě před epochou počítačů, ale počítač může číst data jen v určité formě. Vstup rozsáhlejších materiálů, to jest přepsání celého textu do počítače, je neúnosně nákladný. Určitou naději nám slibuje použití scannerů, optických snímačů textu a grafiky. Jejich kvalita rychle roste a ceny rychle klesají. Největší překážkou je kvalita ORC, programu, který přeloží grafické obrazy písmen zachycených snímačem do jejich reprezentace v ASCII kódu. Jenom pak mohou být tyto texty analyzovány počítačem. Kvalita těchto programů je dosti neuspokojivá, jsou schopny rozeznat jen omezený počet typů a velikosti písem a produkují mnoho chyb. Teprve v těchto týdnech (jaro 1991) odborný tisk označuje některé ohlášené ORC programy za průmyslově použitelné. A až jejich cena klesne, bude snad i jednotlivec schopen využít počítače pro obsahovou analýzu. O počítačové analýze textu si řekneme něco více v souvislosti s kvalitativním výzkumem.

Kdybyste si zatím chtěli přečíst více o obsahové analýze, můžeme vám doporučit malou, ale chytrou knížku Klause Krippendorfa (1981).

Sekundární analýza dat

Vlastně jsme vám před chvílí lhali. Existuje ještě jedna možnost, jak získat data ještě daleko levněji než obsahovou analýzou. Ne vždycky, ale častěji, než bychom očekávali, mohla být data, která potřebujeme, už někdy někde někým sebrána. Více a více institucí zřizuje archivy dat. To není jen sbírka zpráv a výzkumných dat, ale původní data, tak jak byla sebrána, zakódována a uložena ve formě, kterou počítač může přímo číst. Co více, tato data jsou většinou dostupná každému zájemci. Ten si někdy musí vyžádat svolení původního autora, ale často je archiv vydá zájemci na požádání a za cenu materiálu (disků, pásek), na kterých jsou data překopírována. S výjimkou dat shromážděných komerčními organizacemi jsou tedy tato data zadarmo.

Dr. Watson:
To je všechno krásné, ale co budu dělat s daty z nějakého jiného výzkumu, na jiné téma, a provedeném na jiné populaci. Ostatně tato data byla již zpracována původními autory.

Náš přítel může mít pravdu, přinejmenším když hovoří o populaci. Pro některé problémy je definice populace relativně flexibilní; některé tvrzení o sociálních změnách v České republice může být právě tak dobře testováno v Západočeském jako v Severočeském kraji a pro testování některých teoretických hypotéz může být volba populace do jisté míry flexibilní. Ale jedná-li se o aplikovaný výzkum, který má být podkladem pro nějakou sociální intervenci, definice je dána. Ovšem velká většina výzkumů má tendenci volit obdobné populace: dospělé obyvatelstvo hlavního města, voliče, středoškolské studenty atd. Hlavně je dobře, že se Dr. Watson mýlí v ostatních bodech, zejména v tom, že data v archivech již byla zpracována. **Každý výzkum testuje jen omezený soubor hypotéz a tyto testy představují jen omezenou množinu sociologicky relevantních kombinací sebraných proměnných. V každém výzkumu je využita jen část užitečné informace, která byla v datech nashromážděna.**

Omezuje-li se náš výzkumný problém na poměrně malý počet proměnných a máme-li trochu štěstí (a ovšem přístup do archivu dat), máme naději, že data, která potřebujeme, už někdo

pro nás shromáždil. Jsme si jisti, že nyní, kdy se počítače stávají běžnou součástí našeho života, „data banks" porostou rychle i u nás.

7.6 Na co jsme nezapomněli

V zajímavé a rozlehlé oblasti sběru dat je mnoho toho, o čem jsme se ani slovem nezmínili. Do naší malé knížky se vejde jenom zlomek. Chtěli jsme však v něm ukázat principy logických problémů, se kterými se v této oblasti setkáváme. Ty dobrodružné povahy, které by chtěly vědět více, naleznou řadu relevantních titulů v naší bibliografii.

Řešení úkolů z kapitoly 7

Cvičení 7.1

Záznamový arch je dokument určený výhradně pro tazatele, respondent nemá tento arch číst a tazatel přirozeně ví, že nemá zaznamenat respondentovo jméno, nemá pro to na formuláři žádnou rubriku. Ten velký červený nápis je poselství pro respondenta, který se přirozeně bude snažit nahlédnout do dokumentu a zahlédnutý nápis ho ujistí o anonymitě výzkumu.

Cvičení 7.2

Tento postup by ovšem plně anuloval zkreslení vyvolané haló efektem, ale působil by psychologicky na většinu respondentů. Mohli by si připadat jako při křížovém výslechu nebo by se mohli cítit jako objekty jakési absurdní hry. Otázky musí sledovat nějakou třeba i jen velice povrchní logiku.

Cvičení 7.3

Odpověď už asi znáte. Otázky kódované stejným směrem by zvýšily pravděpodobnost stereotypních, mechanických odpovědí, zvýšily by pravděpodobnost „response setu". Pokud bychom z nějakého důvodu potřebovali sumarizovat pozitivní/negativní hodnocení, je velice snadné sebraná data překódovat. Stačí kratičká instrukce počítači, a ten změní kód pro všechna pozorování ve zlomku vteřiny.

Cvičení 7.4

Noviny se mohou lišit svým rozsahem. Noviny A mohou mít 16 stránek a noviny B jen 4, a tak noviny B věnují danému tématu **relativně** daleko vyšší pozornost než noviny A, které věnovaly této tematice více řádků.