

Cvičení 2.7.

Konstrukci hypotetických výsledků můžeme provést na základě následujícího souboru otázek:

- Ve kterých ze čtyř skupin se projeví vliv filmu?
- Ve kterých vliv "historie"?
- Ve kterých vliv "testování"?

Odpověď na tyto otázky je snadná. "Historie" zasáhla všechny skupiny; předpokládáme, že vojáci ve všech skupinách se dozvěděli o útoku na Tel Aviv. "Testování" ovlivnilo jenom skupiny ve kterých bylo použito předběžné měření, a ovšem, film mohl mít vliv jen na obě experimentální skupiny.

Nejvyšší musí tedy být výsledky následného měření v první skupině. Je zde kumulován vliv všech faktorů: filmu, "testování" i "historie". Nejnižší bude skóre v druhé kontrolní skupině, ale i to bude vyšší, než bylo skóre prvních dvou skupin v předběžném měření. I tato skupina byla ovlivněna "historií". Výsledky ostatních skupin budou někde mezi těmito oběma extrémy. Jejich pořadí podle velikosti bude záviset na relativní síle obou zkreslujících faktorů. Tabulka odpovídající těmto podmínkám by mohla vypadat třeba takto:

	Čas 1	Film	Čas 2
Experimentální skupina 1.	50	ANO	85
Kontrolní skupina 1.	50	NE	65
Experimentální skupina 2.	--	ANO	80
Kontrolní skupina 2.	--	NE	60

A jistě si dovedete sami vypočítat, že "historie" odpovídá za 10 bodů přírůstku ve všech skupinách (rozdíl mezi "následným" měřením ve druhé kontrolní skupině a předběžným měřením v kterékoliv z prvních dvou skupin). Rozdíl mezi experimentální skupinou jedna a dvě v následném měření (stejně jako rozdíl mezi oběma kontrolními skupinami) nám dá sílu efektu "testování": 5 bodů. A konečně, čistý efekt filmu, 20 bodů, dostaneme porovnáním výsledků kterékoliv experimentální skupiny s odpovídající kontrolní skupinou a ještě několika dalšími způsoby. Víte jakými?

Pozorování nepozorovatelného

transformace informací
zkreslení v interview
validita
kauzalita
kontrola validity
členství ve známé skupině
prediktivní validita

konstruovaná validita
mínění skupiny soudců
výčet obsahu
zjevná validita

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 nepřijemností 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 kout 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 vysvětlíte, 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 nepřímou. 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 zvětšovací 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 naši otázku 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.

Cvičení 2.2.

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. Pít 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 žvete 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 nepřijemné, 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 šedesátých 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 pumpřdentnosti je:

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

Dotazník byl pak aplikován na malý a nereprezentativní vzorek posluchačů prvního 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 prvních výzkumů, asi v roce 1966, jsme dostali v malých vesnicích na moravskoslovenském pomezí dostali 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."

Zkreslení může vyvolat i představa pozitivní sankce. Respondent volí nepravdivou odpověď, aby se ukázal v nejlepší světlo, 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 15 minutových intervalech, co dělali od probuzení do usnutí. Tazatelé 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 nepříjemné **vyslovit** určitou nepřijemnou 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 jeho 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 povážlivým 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.

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 opožděný koberec 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 neúspěšnější byl vyhodnocen exponát věnovaný úspěchům Čtyřletého plánu hospodářské výstavby, skříně s uprášenými diagramy a plakáty, věcmi, o které v té době předstíraly zdvořilý zájem jen reprezentanti nejvyšších mocenských orgánů.

Výzkum byl publikován, ředitel dostal vyznamenání a muzeu podstatně zvýšený rozpočet. Jenom několik set zasvěcených vědělo, že kolem neúspěšnější 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 být měření, která 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, třeba 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 následujících výčetů technik pro kontrolu validity není snadná. Zdaleka ne všechny techniky, zejména ne ty nejučinnější, 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 předpokládajících, 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. P.I. 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 nepřijemnou 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, str.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 prvním 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 nepřijemnou připomínku: Neexistuje nic takového jako univerzálně validní měření určitého konceptu. **Validita prokázána 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 chzení do kina v kategoriích "často", "na 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 "necestuj" 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: Readers Digest není v Kanadě příliš populární, chzení 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 (Dalas 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 zavržená. 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 popsáný 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 gheta, jehož negramotní rodiče žijí na veřejné podpoře, nebo na rómskou 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 zamyslet 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 čistá 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 odpoví 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 chůze 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?"

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. Nařkali 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 před tím, 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 okouzlení. 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 je 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 klikatá. 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 věci, 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, str.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". Začneme 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, str.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 ztrácí 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 frekvencí 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, uvedeně 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 filosofickými 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 těžce 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 být neplatné v jiné, stejně demokratické. Prostě se tyto země liší v politické kultuře, různé politické akty mohou mít různý význam a nebo 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,

nebezpečně 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, str. 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.

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 rozhlasu
- 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 prvních 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 prvních třech typech "gladiátorské" participace působilo 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 vypořádávat 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 první 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 informacích 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 mnoho 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 neznamenají 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 ovlivňujících participaci, stejně dobře jako vliv participace na expozici stimulům. Ale to všechno už dobře známe z diskuse kauzality a experimentu v první 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 zkresení, 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 nepřijemné, 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 první 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 třista 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 "začneme 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 časném 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 prvních 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 trávniku 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.