Václav BŘICHÁČEK

ÚVOD DO PSYCHOLOGICKÉHO ŠKÁLOVÁNÍ

psychodiagnostické A didaktické testy
n. p. BRATISLAVA

Václav BŘICHÁČEK

ÚVOD DO PSYCHOLOGICKÉHO ŠKÁLOVÁNÍ

Psychodiagnostické a didaktické testy, n. p.,
BRATISLAVA

Název: Úvod do psychologického škálování

Autor: PhDr. Václav Břicháček

Oponenti: PhDr. Libuša Maršálová, CSc.
PhDr. Teodor Kollárik, CSc.

Tato kniha je věnována památce

mého univerzitního učitele

profesora Jana D o l e ž a l a .

O b s a h

Předmluva

Základním cílem veškeré poznávací vědecké činnosti je odkrýt a pokud možno kvantitativně formulovat objektivně existující zákony jsoucna a využít jejich poznání pro společenskou praxi. Tato obecně platná teze naráží při své konkrétní realizaci v psychologii na mnohé překážky. Psychologické procesy jsou velmi složité a poznání zákonitostí je často ovlivněno metodickou úrovní výzkumné činnosti. Vývoj psychologie je proto spoluurčován i vývojem používaných výzkumných metod. Již od vzniku psychologie jako samostatného vědního oboru se prosazovala snaha o systematické zavádění a využívání kvantitativních metod jako prostředku detailního poznávání zkoumané skutečnosti. Vytvořily se i některé zcela specifické techniky, jako jsou psychofyzické metody, psychometrika, faktorová analýza a pod. Využití kvantitativních metod by samo o sobě nebylo pokrokem, pokud by se neopíralo o racionální rozbor zkoumané problematiky, o pochopení dialektických souvislostí mezi jednotlivými procesy i o obecnou vědní teorii. Kvantitativní metody v psychologii byly mnohdy spojeny se zjednodušujícími výklady; není proto divu, že o oprávněnosti měření v psychologii se vedlo mnoho diskusí.

Tato kniha je věnována psychologickému škálování jako jedné z forem kvantifikace psychologických procesů, a to převážně takových, které se nemohou stát předmětem fyzikálního měření. Jedná se o metody, které jsou velmi rozmanité jak co do složitosti, tak co do rozsahu aplikovatelnosti. Psychologické škálování se velmi prudce rozvíjí; objevují se nové techniky - bohužel mnohdy bez přímého vztahu ke skutečným potřebám rozvoje konkrétního psychologického bádání.

Při přípravě tohoto textu jsem se opíral o vlastní zkušenosti výzkumné i o zkušenosti, které jsem nabyl při konzultačním vedení řady diplomních prací posluchačů psychologie. Text postupně krystalizoval v přednáškách, které jsem měl jednak na katedře psychologie FF UK v Praze, jednak na půdě Československé psychologické společnosti. Zájem o tyto kursy byl vlastním spouštěcím podnětem pro sepsání úvodního textu, který předkládám čtenářům. Je v prvé řadě určen psychologům v praxi, kteří při své náročné činnosti nemohou systematicky sledovat současnou literaturu o psychologickém škálování. Dále jsem měl na mysli studenty psychologie, a snad i některým pedagogům, sociologům a lékařům bude užitečný.

Text je psán pro čtenáře, který není matematicky vzdělán. Proto nejsou záměrně probírány matematicko-statistické základy jednotlivých metod, ale důraz se klade na praktické návody. Učebnice tohoto charakteru mají ovšem mnohá úskalí; nepochybuji však, že ten, kdo se bude problematikou psychologického škálování zabývat hlouběji, dokáže si potřebné teoretické informace sám získat. Snažil jsem se též, abych odstraňoval matematickou fobii, která se mnohdy mezi psychology objevuje; přál bych si, aby alespoň pro některé čtenáře byla četba korektivní zkušeností. Vlastním cílem však je přesvědčit čtenáře o užitečnosti psychologického škálování a poskytnout mu obraz o jeho základních podobách.

Při výběru škálovacích metod, které jsou zde popsány, jsem stál před řadou problémů. Volil jsem převážně metody, které považuji za relativně nejlépe propracované a zároveň též kriticky přezkoušené v praxi. Některé z nich jsou u nás publikovány po prvé. Volil jsem dále ty metody, které jsou relativně jednoduché. Pomíjím - kromě úvodní informace v kapitole XIV. - celou problematiku multidimenzionálního škálování, které je pro současnou psychologii velmi slibné. Měl bych zde slíbit, že napíši druhý díl, ve kterém by byly tyto typy škálovacích technik popsány. Ale neslibuji. Jednak se znám, jednak je tato oblast velmi obtížná a zatím není ještě dostatečně ověřena ani ve výzkumné ani v provozní praxi.

Kniha o psychologickém škálování vznikala za pomoci mnoha přátel a kolegů. V prvé řadě patří můj dík mé ženě Věře za nesmírné pochopení, které mi umožnilo, abych se této práci po dlouhou dobu věnoval. Přátelům Karlu Balcarovi a Jiřímu Koženému děkuji za stimulování a systematické ujišťování, že stojí za to, aby tato kniha byla napsána. Olze Hampejsové vděčím za velmi kritickou četbu první verze rukopisu a za mnohé komentáře a náměty, které celému textu výrazně prospěly. Oba recenzenti - Libuše Maršálová a Teodor Kollárik - mne upozornili na řadu nepřesností a pomohli mi lépe vyvážit jednotlivé kapitoly. Stáně Čermákové jsem vděčen za pečlivý a mnohde opakovaný přepis celého textu v jeho různých podobách. Ladislavu Stehlíkovi vděčím za obratné překreslení obrázků a grafů. Předem též děkuji kritickým čtenářům, od kterých očekávám sdělení všech připomínek a komentářů po stránce věcné i didaktické.

Nad Černým dolem, podzim 1975

Václav Břicháček

Kapitola I.

ÚVOD

Metodologická problematika měření a kvantifikování psychologických daností se objevila prakticky ve stejné době, kdy vzniká psychologie jako samostatná vědní disciplína. Stačí připomenout tradice psychofyziky, mentálních testů a konečně i počátky celé experimentální psychologie. Různí autoři hledali přiměřené měřící metody a stupnice, které by se svými vlastnostmi blížily stupnicím, vyvinutým v přírodních vědách. Míry, délky, váhy, teploty, času atd., se zdály ideálními nástroji, které se osvědčily v historii vědy, a řada psychologů usilovně hledala podobné nástroje pro měření takových jevů, jako je vnímání, inteligence, city, volní napětí, postoje, mínění atd.

Záhy se ukázalo, že většina psychologických vlastností a procesů nemůže být měřena tak, jako délka či váha. Tato zkušenost vedla k dvěma protichůdným postojům. Objevily se názory, že psychologické vlastnosti a subjektivní stavy nelze vůbec měřit, ale pouze kvalitativně zachytit a slovně vyjádřit. Snaha kvantifikovat pocit životního štěstí, tvůrčí schopnosti či morální hodnocení vede psychologii do slepé uličky zjednodušených konceptů, z kterých vyplývají chybné úvahy. Naproti tomu se poukazovalo, že pro měření složitých psychologických procesů nejsou přiměřené tradiční metody měření, známé z přírodních věd. Z tohoto postoje vyplynulo rozsáhlé a metodologicky plodné úsilí vytvářet postupně nové metody měření a jejich pomocí pronikat do složitých zákonitostí psychologie. Při sledování této druhé cesty bylo třeba opustit zjednodušené analogie a zamyslet se nad druhy dat, které má psychologie k dispozici. Postupně se objevily škálovací metody, psychometrika, faktorové techniky, metody sdružených měření atd.

Je jisté, že na této cestě se vyskytly různé chyby a nepřesnosti; řada vypracovaných a zdánlivě slibných metod selhala či byla dalším vývojem překonána. Metodologická problematika měření v psychologii není ani dnes definitivně uzavřena. Právě naopak - v posledních letech jsme svědky velmi bouřlivého a mnohdy i obtížně sledovatelného vývoje. Objevují se nové metody multidimenzionálního škálování, možnosti sdružených forem měření, využívá se moderní výpočtová technika atd. Tento rozvoj však s sebou nese jisté nebezpečí metodologického formalismu; nové měřící techniky se rozvíjejí bez ohledu na psychologickou realitu a na možnosti skutečné a gnoseologicky správné aplikace. Čas a praxe jistě ukáží, které metody jsou nosné a které nikoli. V této studii se zaměřujeme pouze na výklad škálovacích metod, které se osvědčily jak v psychologickém výzkumu, tak v praxi.

I.1. Základní pojmy

Pojmy "škála" i "škálování" se užívají často v různých významech. Slovo škála pochází z latinského "scalae", což znamená žebřík, schody /slovesně "scandere"= stoupati, vystoupiti/; v angličtině kromě tohoto významu značí "scale" též misku vah či vůbec váhy, a slovesně pak vážit /ale i vystoupit na něco, zakreslit v jiném měřítku aj./.

Již z tohoto pohledu se jeví dva významy pojmu škála: jednou jako nástroj pro měření určité veličiny, po druhé jako výsledek měření nějaké veličiny. V prvém případě hovoříme např. o škále následných intervalů, zdánlivě stejných intervalů, o škálách Coombsových či Guttmanových atd. Rozumíme tím metodu, kterou vypracoval Coombs či Guttman, nebo techniku, při které používáme zdánlivě stejné intervaly pro nalezení vztahů mezi posuzovanými jevy a pod. Při tom nezáleží příliš na tom, jakou veličinu chceme měřit; příslušnou metodiku je možno užít pro meření různých procesů. V druhém případě škálou rozumíme stupnici, vytvořenou podle určitých a většinou pevně stanovených pravidel a algoritmů, jako např. Binetova škála inteligence, škála pro měření úzkosti dle Taylorové, škály pro měření empatie, sebehodnocení, rizikového rozhodování atd. Jedná se tedy o určitý konkrétní a již dříve vypracovaný nástroj, jimž lze měřit určitý psychologický jev či některou jeho stránku.

V této studii se budeme zabývat škálami v prvním smyslu, t.j. určitými obecnými metodami či technikami, které lze užít pro měření různých jevů či procesů. O škálách v druhém smyslu /jako o konkrétních nástrojích pro měření jednoho jevu/ budeme uvažovat jen vyjímečně /např.v souvislosti s posuzovacími škálami/. Škál tohoto typu je dnes nesmírné množství, které je prakticky nepřehlédnutelné. Extenzívní informace o jejich vývoji a o zkušenostech s nimi vycházejí pravidelně v Burosových ročenkách mentálního měření /Mental Measurement Yearbooks/.

V minimálně dvojím významu se objevuje i pojem "škálování" /scaling/. Za prvé se jím rozumí užívání různých technik, kterými přiřazujeme číselné hodnoty k takovým jevům, které nemůžeme přímo měřit na intervalové či poměrové stupnici. V tomto smyslu budeme též tohoto pojmu užívat. Druhý význam, který se pojmu škálování přikládá, značí soubor různých postupů, kterými je možno vytvářet škály; zde bude na místě hovořit spíše o souboru škálovacích technik či procedur.

Škálování, jak jsme je právě popsali, bývá někdy ztotožňováno s pojmem měření. Ve shodě s Berkou /1972/ považujeme za užitečné tyto pojmy obsahově rozlišovat a zavést ještě třetí pojem, a to klasifikování. Mezi těmito třemi kategoriemi vyjadřování dat jsou zhruba tyto vztahy:

Klasifikováním /případně kategorizováním/ rozumíme rozdělení množiny jevů do podmnožin /kategorií/ na základě určitých a předem stanovených pravidel. Jedná se tedy o třídění a základní pravidla jsou tato: jednotlivé kategorie musí být vyčerpávající, musí se vzájemně vylučovat a být na sobě nezávislé. Přitom kategorie bývají určovány ve shodě s účelem výzkumu /či se sledovaným problémem/, a mohou se případ od případu měnit.

Škálováním budeme rozumět numerické zpracování pořadových dat, t.j. takových dat, ve kterých se opíráme o srovnávání řady jevů a jejich sestavování do pořadí /větší - menší, velmi potřebný - potřebný - neutrální - nepotřebný - zcela nepotřebný a pod./. Přitom se snažíme zpracovat tato data tak, abychom získali kvantitativní údaje, umožňující případné další metrické postupy.

Pomocí škálování se tedy snažíme převést ordinální škálové hodnoty na hodnoty kardinální.

O měření budeme hovořit tehdy, půjde-li o vyhodnocování či zpracování dat numerického typu, t.j. takových, která získáváme při práci s intervalovými či poměrovými škálami.

Škálování je tedy jakýmsi předstupněm měření; tvoří přechod od pojmů kvalitativních k pojmům kvantitativním a umožňuje kvantifikovat pořadová data. Proto budeme škálování a měření souhrnně nazývat kvantifikací.

Otázky kvantifikace tvoří složitý komplex jak z hlediska obecné vědní metodologie, tak z hlediska jednotlivých vědních disciplín. Toto předivo problémů zasahuje do obecných gnoseologických východisek vědeckého bádání a přes specifické výzkumné i aplikační postupy, vypracované v dané vědní oblasti, dochází až k detailním technickým otázkám konkretních měřících nástrojů a postupů.

Problematika kvantifikace v psychologii je stále častěji předmětem pozornosti, což souvisí s nutností vyjasňovat obecná metodologická východiska jak psychologického výzkumu, tak psychologické praxe. Velmi důkladné studie sovětských psychologů Itelsona /1964/, Kyverjalga /1971/ a Kulla /1974/, řada východoněmeckých studií Gutjahra /1968, 1972/, Clausse/1968/, Essera /1970/ či Krasemannové /1966/, stejně tak, jako polská studie Walentova /1971/, svědčí o plodném využívání obecné marxistické metodologie při řešení zásadních otázek kvantifikace v psychologii. Pro československého čtenáře má svou důležitost i široká diskuse o měření ve společenských vědách, která probíhala na stránkách Sociologického časopisu v letech 1971 až 1973 /viz souhrn, 1973/ a studie Berkovy /1971, 1972/.

I.2. Základní třídění škál v psychologii

V současné psychologii existuje celá řada třídění škál, s kterými se pracuje; v úvodním přehledu budeme v podstatě vycházet z hierarchického systému Stevensova /1946, 1951/, který rozlišil 4 druhy škál: nominální, pořadové, intervalové a poměrové. Tohoto třídění se užívá velmi často /např. Gutjahr, 1972/; bývá také základem pro systémy podobné /např. Torgerson, 1958/. Není sice nesporné /viz Berka 1972, který uvádí řadu jiných klasifikací škálových typů/, ale pro psychologické data je považuji za orientačně celkem přiměřené. Jednotlivé škály se od sebe liší svou úrovní od nejjednodušší a nejobecnější škály nominální po velmi specifické škály poměrové. Jako kritéria pro rozdělení škál Stevens užívá: a/ empirické operace použité při přiřazování čísel sledovaným jevům, b/ formální či matematické vlastnosti jednotlivých škál, c/ statistické techniky použitelné pro další zpracování získaných dat.

I.2.a. Nominální škály

Nominální škála je nejjednodušší a přináší nejméně informací. Jedná se o jednoduché třídění dat do vzájemně se vylučujících kategorií; každý prvek může být zařazen pouze do jediné kategorie a všechny prvky je možno zařadit. Je tedy nezbytné najít kritéria, podle kterých se jednotlivé údaje začleňují do určité kategorie, a zároveň je třeba jednotlivé kategorie řádně definovat. Jsou pak označovány čísly, která můžeme přiřazovat jakýmkoliv způsobem; můžeme je případně nahradit i jinými symboly /písmeny, znaky, slovním vyjádřením a pod./.

Nominální škála může mít dvě formy: a/ Každá kategorie obsahuje jen jeden jediný prvek /např. čísla na dresech sportovců, kdy v jednom družstvu se nemohou vyskytnout dva hráči se stejným číslem/. b/ V každé kategorii se může vyskytnout více prvků a jednotlivé kategorie se číslují podle stanoveného pravidla /např. 1. muži, 2. ženy; nebo 1. všechny osoby, které pracují v profesi A, 2. osoby profese B, 3. osoby profese C atd./.

Základním požadavkem pro vytvoření nominální škály je jasné definování jednotlivých tříd a stanovení znaků, podle kterých tyto třídy od sebe dělíme. Množství tříd a jemnost jejich rozlišení je většinou určeno cílem sledování. Můžeme často volit třídy kombinováním většího počtu dimenzí /např. 1. muži s diagnózou A a s profesí x, 2. ženy s diagnózou A a s profesí y, 3. muži s diagnózou B a s profesí x atd./. V mnoha případech je třeba počítat s tím, že do třídění pronikají automaticky i další dimenze, které jsou s třídícím znakem více či méně těsně svázány /např. při třídění osob do kategorií: aktivní sportovec - nesportující, můžeme čekat, že v první kategorii budou v průměru osoby mladší než v kategorii druhé/.

Nominální škálu je možno vytvářet i na základě údajů, které získáváme při posuzování či popisu jednání určité osoby. Tak třeba při popisu pracovní činnosti známe normativní postup /třeba daný určitým předpisem/ a zjišťujeme, zda pracovník všechny úkony provedl nebo ne, které byly častější atd. Podobně při sledování sociální interakce při rozhovoru můžeme zavést kategorie jako např. "klade otázku", "vysvětluje", "přerušuje kontakt" atd. a sledovat, jak často se vyskytne jednání, patřící do jednotlivých kategorií.

Čísla, která přiřadíme k jednotlivým třídám, nereprezentují skutečnou hodnotu použitých čísel, ale jsou v podstatě pouhým označením skupin. Můžeme je libovolně měnit a stupnice zůstane invariantní. Rozdělíme-li skupinu osob na muže a ženy, je zcela jedno, zda skupinu mužů označíme číslem 1 a skupinu žen číslem 2, či zda je očíslujeme opačně /či zda je označíme písmeny či jinými symboly/.

Při statistickém zpracování dat, založených na nominálních stupnicích, se opíráme pouze o četnost údajů v jednotlivých třídách; někdy se hovoří o statistice kvalitativních dat. Je zřejmé, že statisticky lze zpracovávat pouze tu formu nominální stupnice, kterou jsme označili jako formu b/. Pro formu a/ - každá kategorie má jen jediný prvek - nepřicházejí statistické techniky v úvahu. Pro popis získaných dat uvádíme prosté počty pozorování, které můžeme vyjádřit histogramem; uvádí se modus /třída nejpočetněji zastoupená; v podstatě údaj, který je možno očekávat s největší pravděpodobností/. Existuje řada neparametrických statistických technik, které se dobře hodí pro data tříděná do nominálních stupnic /viz např. Lienert, 1962, kap.V./.

I.2.b. P o ř a d o v é   š k á l y

Základním znakem pořadové stupnice je to, že mezi sledovanými jevy můžeme stanovit nejen ekvivalenci, ale i pořadový vztah, a na jeho základě konstatovat, že jev A je větší /dokonalejší, silnější, vhodnější atd./ než jev B či naopak. Operace, kterou toto srovnávání provádíme, umožňuje konstatovat, že je-li A větší než B, B větší než C, pak také A musí být větší než C. Při zjištění, že A je větší než B, nelze současně tvrdit opak. Jednotlivým jevům či pořadovým kategoriím pak přiřazujeme čísla, mezi kterými existují též pořadové vztahy; výsledkem je číselná řada, odpovídající pořadí tříděných či posuzovaných jevů.

Podobně jako u nominální stupnice i u stupnic pořadových se mohou vyskytnout dvě formy:

a/ každý jev je zařazen tak, že dostane určité místo v pořadí, a toto číslo již žádný jiný dostat nemůže. Takto se řadí např. děti v tělocviku podle velikosti, či mužstva v tabulkách sportovních soutěží;

b/ častěji se objevuje pořadová stupnice, kdy v každém jejím stupni může být řada prvků /např. děti ve třídě rozdělíme do 5 skupin podle píle nebo podle stupně zájmu o určitý předmět atd./.

Pořadové škály se v psychologii vyskytují velmi často, a to tam, kde nemůžeme získávaná data přesně měřit. Užívají se při hodnocení osob, kvality sportovních či pracovních výkonů, při posuzování řady rysů sociálního jednání, při hodnocení postojů, odhadech zlepšení či zhoršení stavu nemocných. Intervaly mezi jednotlivými body většinou nejsou stejné a často ani nemůžeme jejich velikost odhadnout. Ve sportu se uvádí pořadí závodníků, které není nikterak ovlivněno tím, jaký je rozdíl mezi prvním a druhým, resp. mezi druhým a třetím v pořadí. Jestliže třeba rozdíl mezi prvními dvěma je velmi těsný, a přitom oba zvítězí nad třetím závodníkem s velkým náskokem, pak tuto skutečnost z pouhé pořadové škály nemůžeme vyčíst.

Při zpracování dat z pořadových škál se můžeme opřít o celou řadu statistických technik. Jako míru centrální tendence určujeme medián, rozložení výsledků vyjadřujeme kvartily či kvartilovou odchylkou. Pro další zpracování pořadových dat existuje řada statistických metod, které jsou popsány jinde /Lienert, 1962, Břicháček - Hampejsová, 1961 aj./. Pořadovou škálu je možno podle potřeby transformovat jakýmkoliv způsobem, který zachová původní pořadí sledovaných jevů.

Torgerson /1958/ zavedl ještě další kategorii, a to pořadovou škálu s přirozeným počátkem. Rozumí tím takovou stupnici, kterou můžeme zakotvit v nulovém bodě. Tak např. v psychofyzice je nulovým bodem absolutní práh; při měření postojů, preferencí či hodnot můžeme za počátek považovat neutrální bod, od kterého se jedním směrem objevují podněty hodnocené pozitivně a v opačném směru hodnocené negativně. Čísla, která jednotlivým podnětům přiřazujeme, jsou pak buď pozitivní nebo negativní.

I.2.c. I n t e r v a l o v é   š k á l y

Tento typ tvoří první skutečně metrické škály. Pro jejich vytvoření potřebujeme sestavit data tak, aby mezi jednotlivými body byla stále stejná vzdálenost /t.j. aby rozdíl sousedních bodů na stupnici byl stále stejný/. Přitom není rozhodující, zda se všechny body stupnice v reálném materiálu skutečně vyskytnou. Je přitom třeba najít experimentální operaci, která umožňuje přiřadit čísla k podnětům tak, aby dvěma stejným vzdálenostem odpovídaly dva stejné číselné rozdíly /tedy $X_B - X_A = X_C - X_B = X_D - X_C \ldots$/. Jestliže takovou experimentální operaci můžeme stanovit, pak můžeme též vytvořit intervalovou škálu.

Příkladem intervalové škály, která se užívá v denním životě, je běžné měření teploty teploměrem v Celsiových stupních. Tato stupnice je zakotvena ve dvou bodech: 0^o je definováno jako teplota, kdy voda mrzne, a 100^o jako teplota, kdy voda přechází do varu /za předpokladu, že jsou dodrženy další podmínky, jako je přesně definovaný tlak atd./. Oba tyto body jsou voleny konvenčně /na Fahrenheitově stupnici odpovídá 0^o dle Celsia hodnota $+32^oF$ a hodnotě 100^oC odpovídá $+212^oF$/, stejně tak počet prvků stupnice je dán dohodou /Réaumurova stupnice dělí vzdálenost mezi tuhnutím a varem vody na 80 stupňů/. Konec konců i směr stupnice je konvenční /původně Celsius sám zakotvil stupnici obráceně: 0^o odpovídalo varu vody a 100^o tuhnutí. Teprve později byla stupnice převrácena/.

Jinými příklady může být měření času ve 24 hodinových cyklech, či kalendář, ve kterém je nulový bod letopočtu konvenčně zaveden a v různých kulturách je umístěn do jiného historického bodu.

V psychologii je definice rovnosti dvou vzdáleností mezi jevy obtížná. Jednou z možných cest je škálování, které budeme popisovat v následujících kapitolách. Budeme např. probírat škálu se zdánlivě stejnými intervaly, při které se snažíme rozdělit určité podněty /třeba určité formulace postojů, popisy chování a pod./ do kategorií, o kterých se nám zdá, že odpovídají bodům, ležícím na stupnici ve stále stejné vzdálenosti od sebe. Řadu jiných postupů vypracovala fechnerovská psychofyzika. Často se jako příklad intervalové škály uvádí měření inteligence vyjadřované v jednotkách IQ; tato hodnota je standardizována tak, aby byla normálně distribuována s průměrem 100 a směrodatnou odchylkou 15.

Všechny uvedené příklady naznačují další příznačnou charakteristiku intervalových škál; nemají totiž přesně definovaný absolutní nulový bod. Nelze stanovit "nulovou" inteligenci ani nulový bod, od kterého bychom počítali historický čas; pouze bylo konvenčně zvoleno určité datum. Stejně tak u měření teploty byla nula zvolena konvenčně jako teplota, při které voda tuhne. Právě proto můžeme srovnávat pouze rozdíly mezi jednotlivými hodnotami na stupnici /stejné rozdíly mezi různými hodnotami odpovídají stejným vzdálenostem mezi nimi/, ale nemůžeme uvažovat o jejich podílu /nelze říci, že den, kdy jsme naměřili teplotu 20^{o}C, byl dvakrát tak teplý jak den, kdy průměrná teplota byla 10^{o}C/.

Při zpracování dat z intervalových škál můžeme používat většinu parametrických statistických metod, které jsou běžně uváděny v učebnicích statistiky /průměry, standardní odchylky, parametrické testy rozdílů, Pearsonův koeficient korelace; lze počítat regrese atd./. Vlastnosti čísel na intervalové škále zůstávají zachovány i po jakékoliv lineární transformaci typu $y = ax + b$.

I.2.d. P o m ě r o v é   š k á l y

O poměrových škálách hovoříme tehdy, když poměr intervalů mezi dvěma sousedními body škály je stejný jako poměr mezi kterýmikoliv jinými dvěma sousedními body. V takovém případě můžeme těmto dvojicím podnětů přiřadit čísla, která mají také stejný poměr. Vytvořit stejné poměry lze experimentálně tak, že se určí následné intervaly, počínaje od nulové hodnoty škály, která je dána absolutně /např. u teploty absolutní nulou, t.j. $-273^{o}C$ nebo $-460^{o}F$/, i když mnohdy nemůže být v praxi dosažena. Mezi poměrové škály patří například délka /nezáleží, zda měříme v metrech či v yardech/, váha, elektrický odpor a pod. Pro měření teploty lze užít škály Kelvinovy, která je vlastně lineárním posunutím Celsiovy škály tak, že má nulový bod shodný s absolutní nulou /t.j. v $-273^{o}C$/; bod tuhnutí vody je vyjádřen hodnotou $273^{o}K$ a bod varu vody má hodnotu $373^{o}K$. U poměrových škál můžeme uvažovat o násobcích a podílech; jsme např. plně oprávněni tvrdit, že $2 \times 150^{o}K = 300^{o}K$ a pod. Řada stupnic ve fyzice má tento charakter; v psychologii se však podobné škály vyskytují jen řídce.

Poměrové škály můžeme měnit libovolnou transformací typu $y = ax$. I z tohoto pravidla je zřejmé, že poměrové stupnice musí mít pevný počátek. Volba měrné jednotky je zde libovolná. Hodnoty na škále mají pouze kladná znaménka /absolutní hodnoty/.

Do kategorie poměrových škál patří i čísla, kterými určujeme počet prvků v množinách, t.j. počet objektů v určitém souboru. Často se zjišťuje počet osob, které v rámci většího /přesně definovaného/ souboru vyřeší určitý úkol, dopustí se chyby, mají nehodu, zlepší se během terapie atd. Jednotkou zde mohou být jednotlivé osoby, ale i dvojice, trojice, tucty či jakékoliv jiné n-tice.

Po stránce statistické se při zpracování dat z poměrových škál mohou užívat všechny metody jako u škál intervalových; přibývají k nim některé další, jako je geometrický průměr, variační koeficient /omylem se v praxi užívá i u dat na intervalové stupnici/ a jiné.

I.3. Význam škálování v psychologii

Snahou několika generací experimentálních psychologů bylo přiblížit se získávanými daty co nejvíce ke stupnicím kvantitativním, t.j. intervalovým či poměrovým. Tato snaha naráží často na značné překážky. Mnohdy máme k dispozici pouze data pořadová či dokonce se musíme spokojit s pouhou klasifikací. Vlastní škálovací techniky se vyvinuly postupně proto, aby bylo možno dokonaleji zpracovat pořadová data tak, abychom je po určitých transformacích mohli chápat jako data na intervalové škále. Jde tedy o to upravit získaná data a získat z nich více informací, než které vyplývají z pouhého řazení. Zároveň je možno pomocí škálování do jisté míry překonávat subjektivní pohled v posuzování různých jevů, jako jsou rysy osobnosti, postoje či názory a pod. Nalezením vhodného měrného nástroje se vytvářejí předpoklady pro objektivnější zachycení a v optimálním případě i pro měření zmíněných jevů.

Při používaní škálovacích metod v psychologii existuje celá řada metodologických i filozofických problémů. Jde o určitý druh poznávácího procesu, kterým poznáváme kvantitativní stránky velmi komplexních procesů. Při škálování zjišťujeme numerické vlastnosti sledovaných jevů poměrně složitými procedurami. Opakovaná šetření vedou často k více či méně odlišným výsledkům, což se sice dá částečně odstranit /lepší instrukce, lepší způsob předkládání úkolů i vyhodnocování dat, citlivější postupy aj./, ale přece jen nutně zůstává určitý rozdíl mezi objektivní skutečnou hodnotou sledovaného jevu a hodnotou, kterou jsme naměřili, resp. vypočítali z výchozích dat. Byl by jistě chybný agnostický názor, že skutečná hodnota měřeného postoje, názoru či jiného jevu je nepoznatelná. Na základě získaných dat ji můžeme odhadnout a tento odhad představuje relativně pravdivý poznatek, který lze dalším zjemňováním a zdokonalováním škálovacích metod zpřesňovat.

Uznání existence objektivní hodnoty jevu, který měříme, je základním znakem materialistického chápání měření. Kvantita nezávisí na subjektu, na našich znalostech, na tom zda a jak ji měříme či pozorujeme, ale je objektivní vlastností skutečnosti. Škálování sledovaného psychologického jevu závisí jednak objektivně na jeho vlastnostech, jednak subjektivně na vstupní přesnosti a dokonalosti stávajících škálovacích postupů. Tento dialektický charakter všech měřících postupů vyjádřil již ve svých matematických rukopisech K. Marx: "Číslo je nejčistší kvantitativní určení, jaké známe. Je však plné kvalitativních rozdílů" /Marx, Engels, Sebrané spisy XX., 1966, str. 533/.

Při škálování v psychologii se můžeme ovšem dopustit řady gnoseologických chyb, formalistických postupů či dokonce hříček. Musí být předem jasná teoretická východiska jednotlivých metod, jejich předpoklady i druhy dat, pro která jsou použitelné. Bez těchto znalostí nelze škálovací metody doporučit jako vhodný prostředek poznání. V mnoha případech jsou výsledky škálování značně závislé na volbě pokusných osob, jejich věku, životních zkušenostech atd. Proto je otázka generalizace výsledků velmi obtížná a i z tohoto hlediska je na místě jistá opatrnost, a to hlavně tehdy, pracujeme-li s relativně homogenními skupinami osob. Řada nepřesností plyne i ze složitosti psychologické problematiky, která může být nepřiměřeně zjednodušována volbou nevhodných škálovacích postupů.

Měření a škálování je dnes v psychologii samozřejmostí. Společenská praxe - ať již výzkumná či aplikační - neustále rozšiřuje sféru procesů, které je možno měřit, a zároveň zvyšuje platnost i spolehlivost zavedených škálovacích a měřících procedur. Přitom však je třeba neustále mít na zřeteli, že se jedná o metodické prostředky, které jsou užitečné jen tehdy, když jsou zapojeny do širších teoretických souvislostí. Bez teoreticky a věcně fundovaných koncepcí, bez jasné formulace nosných a závažných problémů a bez přiměřené znalosti dříve dosažených výsledků je jakékoliv uplatňování kvantitativních metod v psychologii neplodné a ve svých důsledcích formalistické a gnoseologicky chybné. Obratné užívání škálovacích technik nemůže nahradit racionální rozbor tématiky, metodicky čisté shromažďování dat a začlenění nálezů do širších teoretických souvislostí. V tomto smyslu vidím ve škálování užitečný prostředek, který nesporně patří do výzbroje současné psychologie, ale sám o sobě není zárukou ani dalšího rozvoje vědeckého poznání ani správné aplikace psychologických poznatků v socialistické společenské praxi.

Kapitola II.

OBECNÉ OTÁZKY PSYCHOLOGICKÉHO ŠKÁLOVÁNÍ

II.1. Výběr indikátorů

Při projektování výzkumu, ve kterém chceme použít některou /případně i několik/ ze škálovacích metod, je třeba řádně a předem promyslit, jaké parametry chceme sledovat, a zvážit, zda a do jaké míry či za jakých předpokladů jsou škálovatelné. V další fázi se stanoví takové indikátory, které umožní zachytit sledovaný parametr v plné šíři a na celém kontinuu, ve kterém se může vyskytovat. Znamená to, že uvažujeme o celé šíři možných reakcí - od jasně kladných až po zcela záporné. Správný výběr indikátorů, které by objektivně odrážely sledovaný parametr, je rozhodující otázkou pro volbu škálovacích postupů. Vytvoření platných a spolehlivých škál má dva předpoklady: a/ detailní znalosti jednotlivých škálovacích technik, jejich gnoseologických východisek a možností aplikace na různá data, b/ sestavení souboru vhodných položek pro škálování sledovaného jevu. Oba předpoklady jsou stejně důležité; v praxi se však většinou klade důraz na první z nich, zatímco druhý bývá neprávem podceňován a považován za triviální.

Položkami se zde rozumí prvky, z kterých se vytváří škála. Vyjadřuje se jimi činnost osoby, její názory, mínění, postoje či jiné procesy, které chceme škálovat. Jednotlivé položky mohou mít různou podobu: mohou to být přímé otázky, kterými se pokusné osoby ptáme na její názory či postoje, může to být vyhodnocení činnosti člověka jinými lidmi a konečně se může použít rozbor výsledků činnosti na základě objektivních dat. Existují dva typy položek: a/ subjektivní, které se opírají o výpovědi člověka o sobě či o jiných osobách; b/ objektivní, které vycházejí z objektivně pozorovatelných interakcí mezi činnostmi člověka a jeho okolím.

Pro výběr položek uvádí Krasemannová /1966/ čtyři předpoklady, které je třeba dodržet, chceme-li škálováním objektivně zachytit či posoudit činnosti člověka:

a/ zvolené položky musí zachycovat podstatné znaky jevů, které jsou posuzovány,

b/ zvolené položky musí jednoznačně zachycovat sledovaný jev,

c/ zvolené položky musí s dostatečnou jemností rozlišovat mezi různými úrovněmi posuzovaných jevů,

d/ zvolené položky musí zahrnovat všechny podstatné formy vztahů ke sledovanému jevu /od extrémně pozitivních po extrémně negativní/.

První požadavek, t.j. zachycení podstatných znaků posuzovaných činností či škálovaných jevů, je věcí racionálního rozboru problematiky a nelze stanovit obecná formální kritéria. Z rozboru vyplyne, zda jednotlivé položky jsou přiměřené, zda nejsou jednostranně zaměřeny jen na jednu stránku sledovaného jevu, či zda nejsou formulovány tak, že v sobě zahrnují různé problémy, které nemusí jednoznačně souviset se zkoumaným jevem. Jako příklad chyby posledního typu může sloužit jedna z prvních empirických studií o měření postojů /Thurstone, Chave, 1929/, zaměřená na sledování postoje k církvi. Při výběru položek došlo ke smíšení postojů k církvi s postoji k náboženství vůbec; tím se obsahová stránka výzkumu zamlžila.

Pokud není předem proveden teoretický rozbor, z kterého je možno odvodit jednotlivé položky, pak se při jejich výběru postupuje většinou zcela intuitivně a může dojít k mnoha chybám. Jsou sice k dispozici přesná pravidla jak stanovit rozlišovací sílu jednotlivých položek, jak zjistit, zda jsou jednoznačně formulovány atd., ale využití těchto prostředků je bez věcného rozboru tématiky nepřiměřené či formální. Na tento fakt je třeba velmi důrazně upozornit, protože v literatuře existují stovky výzkumů /hlavně při výzkumu postojů/, ve kterých se používají nejrůznější škálovací postupy, ale výběru položek je věnována jen minimální pozornost. Marně se hledá zdůvodnění, proč byly použity právě určité položky a proč ne jiné. Stává se pak často, že se v literatuře objevují protichůdné či rozporné nálezy, které bývají působeny právě tím, že pojmově stejný jev je sledován různými škálami, jejichž položky se od sebe po obsahové stránce značně liší. Tyto obtíže vidím hlavně v pozitivisticky laděných studiích amerických, které často předem rezignují na teoretický rozbor problematiky a postupují čistě empirickým způsobem.

Kromě toho je třeba počítat i se značnou společensko-historickou podmíněností řady výzkumů. Položky, které mohou být přiměřené v amerických společenských podmínkách, mohou být sporné či dokonce chybné v našich podmínkách. Mnohé položky též stárnou a postojový dotazník, který mohl být přiměřený před 15-20 lety, může být sporný v současné době, pakliže se jedná o problematiku, která je silně ovlivněna dynamikou společenského vývoje.

Druhým předpokladem pro sestavení vhodné škály je požadavek, aby jednotlivé položky byly jednoznačně, přesně a srozumitelně formulovány. Bylo by možno s jistou analogií opakovat požadavky, které se uvádějí pro formulaci psychologických dotazníků. Každá položka musí být jednoznačně pochopena pokusnou osobou, a to zcela v tom smyslu, který jí dal experimentátor. Stejně tak odpovědi pokusných osob musí být jednoznačné a jejich interpretace vyhodnocovatelem se musí shodovat s tím, co pokusná osoba chtěla vyjádřit. Tak např. v postojové škále by byla pochybná položka, která by zněla "Chtěl bych podávat lepší pracovní výkon". Je těžko tuto položku jednoznačně interpretovat. Rozumí se tím kvalita výkonu, množství výkonu, jiné pracovní činnosti atd., anebo jakési smíšení všech těchto aspektů? Jednoznačnější by asi byly položky: "Chtěl bych se podílet na vývoji nové techniky", "Chtěl bych mít samostatnější práci", "Chtěl bych mít práci s větší odpovědností", "Chtěl bych pracovat kvalitněji" atd. Obecně lze říci, že čím je položka konkrétnější, tím je také jednoznačnější.

Používá-li se škála, jejíž jednotlivé položky nejsou jednoznačné, pak interpretace získaných dat vede nutně ke snížení validity celého výzkumu. Byly vyvinuty numerické techniky, které umožňují vyloučit položky, které nejsou jednoznačně interpretovatelné, resp. na které se objevují nápadně rozdílné názory. Uveďme příklad:

Při použití některých škálovacích technik se shromáždí nejprve řada výroků. Skupina expertů /často i několik set/ je hodnotí k církvi. Při výběru položek došlo ke smíšení postojů k církvi s postoji k náboženství vůbec; tím se obsahová stránka výzkumu zamlžila.

Pokud není předem proveden teoretický rozbor, z kterého je možno odvodit jednotlivé položky, pak se při jejich výběru postupuje většinou zcela intuitivně a může dojít k mnoha chybám. Jsou sice k dispozici přesná pravidla jak stanovit rozlišovací sílu jednotlivých položek, jak zjistit, zda jsou jednoznačně formulovány atd., ale využití těchto prostředků je bez věcného rozboru tématiky nepřiměřené či formální. Na tento fakt je třeba velmi důrazně upozornit, protože v literatuře existují stovky výzkumů /hlavně při výzkumu postojů/, ve kterých se používají nejrůznější škálovací postupy, ale výběru položek je věnována jen minimální pozornost. Marně se hledá zdůvodnění, proč byly použity právě určité položky a proč ne jiné. Stává se pak často, že se v literatuře objevují protichůdné či rozporné nálezy, které bývají působeny právě tím, že pojmově stejný jev je sledován různými škálami, jejichž položky se od sebe po obsahové stránce značně liší. Tyto obtíže vidím hlavně v pozitivisticky laděných studiích amerických, které často předem rezignují na teoretický rozbor problematiky a postupují čistě empirickým způsobem.

Kromě toho je třeba počítat i se značnou společensko-historickou podmíněností řady výzkumů. Položky, které mohou být přiměřené v amerických společenských podmínkách, mohou být sporné či dokonce chybné v našich podmínkách. Mnohé položky též stárnou a postojový dotazník, který mohl být přiměřený před 15-20 lety, může být sporný v současné době, pakliže se jedná o problematiku, která je silně ovlivněna dynamikou společenského vývoje.

Druhým předpokladem pro sestavení vhodné škály je požadavek, aby jednotlivé položky byly jednoznačně, přesně a srozumitelně formulovány. Bylo by možno s jistou analogií opakovat požadavky, které se uvádějí pro formulaci psychologických dotazníků. Každá položka musí být jednoznačně pochopena pokusnou osobou, a to zcela v tom smyslu, který jí dal experimentátor. Stejně tak odpovědi pokusných osob musí být jednoznačné a jejich interpretace vyhodnocovatelem se musí shodovat s tím, co pokusná osoba chtěla vyjádřit. Tak např. v postojové škále by byla pochybná položka, která by zněla "Chtěl bych podávat lepší pracovní výkon". Je těžko tuto položku jednoznačně interpretovat. Rozumí se tím kvalita výkonu, množství výkonu, jiné pracovní činnosti atd., anebo jakési smíšení všech těchto aspektů? Jednoznačnější by asi byly položky: "Chtěl bych se podílet na vývoji nové techniky", "Chtěl bych mít samostatnější práci", "Chtěl bych mít práci s větší odpovědností", "Chtěl bych pracovat kvalitněji" atd. Obecně lze říci, že čím je položka konkrétnější, tím je také jednoznačnější.

Používá-li se škála, jejíž jednotlivé položky nejsou jednoznačné, pak interpretace získaných dat vede nutně ke snížení validity celého výzkumu. Byly vyvinuty numerické techniky, které umožňují vyloučit položky, které nejsou jednoznačně interpretovatelné, resp. na které se objevují nápadně rozdílné názory. Uveďme příklad:

Při použití některých škálovacích technik se shromáždí nejprve řada výroků. Skupina expertů /často i několik set/ je hodnotí na stupnici /třeba s 11 body/ podle toho, zda uvedený výrok reprezentuje postoj velmi kladný či neutrální nebo jen mírně záporný atd. Pokud se posuzovatelé zhruba shodnou, je možno uvedenou položku považovat za jednoznačnou, pokud se neshodnou, je zřejmě její interpretace víceznačná. Pro posouzení shody či neshody bylo zavedeno Thurstonem kritérium, opírající se o interkvartilovou odchylku.

Postup vyplyne z následujícího příkladu: mějme výsledky od 300 posuzovatelů při posouzení 2 položek na škále s 11 body. Výsledky jsou zachyceny v Tab.II.1.

Z uvedených dat se vypočte pro obě položky první a třetí kvartil, tedy hodnota 25. a 75. centilu /tj. bodu, který odděluje první čtvrtinu dat od ostatních, čili pod kterým leží 25% případů, resp. bodu, který odděluje čtvrtou čtvrtinu dat od ostatních, čili pod kterým leží 75% případů/. V rozmezí těchto dvou hodnot leží 50% případů /případů, které leží "uprostřed" stupnice/. Z běžné elementární statistické literatury je známo, že hodnoty 1. kvartilu /Q_1/ a 3. kvartilu /Q_3/ se vypočtou dle vzorce

$$Q_1 = U + \left(\frac{0{,}25 - \sum p_u}{p_i} \right) . w \quad \text{resp.} \quad Q_3 = U + \left(\frac{0{,}75 - \sum p_u}{p_i} \right) . w$$

kde U = spodní hranice intervalu, ve kterém je hodnota Q_1 / resp. Q_3/

$\sum p_u$ = kumulativní součet relativných frekvencí, které jsou v intervalech nižších, než je interval, do kterého spadá Q_1 /resp. Q_3/

p_i = relativní frekvence případů, které leží v intervalu, ve kterém je hodnota Q_1 /resp. Q_3/

w = šíře intervalů /v našem příkladě w = 1/

Pro položku X platí:

$$Q_1 = 4 + \left(\frac{0,25 - 0,09}{0,17} \right) = \qquad Q_3 = 8 + \left(\frac{0,75 - 0,70}{0,14} \right) =$$

$$= 4 + 0,94 = 4,94 \qquad = 8 + 0,36 = 8,36$$

Pro položku Y platí:

$$Q_1 = 1 + \left(\frac{0,25 - 0,17}{0,39} \right) = \qquad Q_3 = 2 + \left(\frac{0,75 - 0,56}{0,24} \right) =$$

$$= 1,21 \qquad = 2,79$$

Mezikvartilová odchylka Q je dána poloviční diferencí mezi hodnotami Q_1 a Q_3, tedy:

$$Q = \frac{Q_3 - Q_1}{2}$$

U položky X je tedy $Q = \dfrac{8,36 - 4,94}{2} = 1,71$

U položky Y platí $Q = \dfrac{2,79 - 1,21}{2} = 0,79$

Tab. II.1.

Rozložení posudků expertů /N=300/ u 2 položek postojové škály

Postoj	1	2	3	4	5	6	7	8	9	10	11
Položka X											
Počet posudků f	0	2	5	18	52	46	49	39	42	33	14
relativní frekvence f/N	0,00	0,01	0,02	0,06	0,17	0,15	0,16	0,13	0,14	0,11	0,05
kumulativní součet relativních frekvencí	0,00	0,01	0,03	0,09	0,26	0,41	0,57	0,70	0,84	0,95	1,00
Položka Y											
Počet posudků f	52	118	73	30	14	11	2	0	0	0	0
relativní frekvence f/N	0,17	0,39	0,24	0,10	0,05	0,04	0,01	0,00	0,00	0,00	0,00
kumulativní součet relativních frekvencí	0,17	0,56	0,80	0,90	0,95	0,99	1,00	1,00	1,00	1,00	1,00

Je zřejmé, že položka Y je mnohem jednoznačněji interpretována a posuzována než položka X. Při sestavování škály se obvykle vylučují ty položky, u nichž $Q > 1,5$.

Tato technická pomůcka pro stanovení formálního kritéria jednoznačnosti položky je užitečná, ale nemůže být užívána nekriticky a bez vztahu k obsahové stránce. Jednoznačná položka, která však obsahově neodpovídá cíli výzkumu, vede k chybám v konečné interpretaci dat /viz první požadavek na volbu položek/.

Třetím požadavkem na volbu položek je jejich rozlišovací schopnost. Položka, která je podstatná a jednoznačná, musí též umožňovat rozlišení, zda sledovaný jev je pozitivní či negativní /případně v jakém stupni je pozitivní či negativní/. V prvé řadě se zjišťuje, zda určitá položka má vůbec nějakou diferenciační schopnost. Jestliže je všemi osobami jednoznačně pozitivně přijata, nebo naopak je všem lhostejná či ji nedokáží zodpovědět, pak se jedná o položku, která nediferencuje a nepřináší rozlišovací informace o jednotlivých osobách. Jiné kritérium může být v tom, že zjišťujeme, zda osoby s výrazně pozitivním postojem a naopak s výrazně negativním postojem ke sledovanému jevu se od sebe liší v tom, jak na danou položku reagují. Jestliže velké percento osob s jasně kladným postojem na danou otázku odpovídá kladně, očekáváme, že zhruba stejně vysoké procento osob s opačným postojem na ni odpoví záporně. Pokud by tomu tak nebylo, diferenciační schopnost této položky by byla opět malá.

Z toho, co bylo řečeno, plyne, že úvahy o diskriminační schopnosti jednotlivých položek je třeba vést dvojím směrem:

a/ Je užitečné orientovat se v předpokuse, zda určitá položka je pro pokusné osoby pochopitelná a zda na ni vůbec dokáží rozumně reagovat. Jestliže velké procento osob uvede, že na danou otázku nedokáže odpovědět, nemá názor, či nemůže se rozhodnout, pak tato položka zřejmě není vhodná pro další používání a zpracování. I zde existují určitá pravidla, jak je možno tento jev numericky vyjádřit /viz Krasemannová, 1966, str. 264-268/.

b/ Vlastní diskriminační charakter jednotlivých položek bývá měřen tak, že se srovnávají výsledky dvou skupin osob s odlišnými postoji. Jestliže odpovědi těchto skupin se v určité položce od sebe statisticky neliší, pak se tato položka nehodí pro škálování jevu, podle kterého byly osoby do skupiny rozděleny. Při numerickém zpracování se používají běžné statistické techniky pro srovnávání 2 výběrových průměrů /obvykle t-test/. Krasemannová uvádí, že pro běžné účely postačuje, když $t \geq 1{,}75$. Položky, u kterých bylo této hodnoty dosaženo, je možno používat. /Tato úvaha se však opírá o srovnání skupin, které jsou sestaveny tak, že z původního výběru se volí vždy 25% osob s extrémně pozitivním a 25% osob s extrémně negativním postojem; "prostředních" 50% osob není pro výpočet hodnoty t bráno v úvahu. Pokud se takto nepostupuje, je třeba volit jiný postup pro určení kritické hodnoty t-testu/.

Konečně čtvrtý požadavek pro výběr položek stanoví, že má být možnost, aby se prosadily všechny očekávané způsoby chování či postoje. Kdyby tomu tak nebylo, získávaly by se nutně jen dílčí údaje, které by neumožnily zachytit sledovaný jev v plné šíři.

Jako příklad můžeme uvést jeden aspekt z výzkumu sovětských autorů Zdravomyslova a Jagova /1964/, kteří v široce zaměřeném výzkumu postojů k práci v různých profesích použili kromě jiného též dotazník, ve kterém bylo úkolem pokusné osoby označit, co se jí na vlastní práci líbí, resp. nelíbí. Dotazník měl celkem 16 položek zhruba tohoto typu:

1/ Proměnlivá práce	1/ Monotónní práce
2/ Práce vyžaduje "fortel"	2/ Práce nepodněcuje myšlení
3/ Dává možnost zvýšit si kvalifikaci	3/ Nedává možnost zvýšit si kvalifikaci
4/ Dobré pracovní nástroje	4/ Špatné pracovní nástroje
5/ Pravidelný pracovní rytmus	5/ Nepravidelný pracovní rytmus
6/ Dobrá organizace práce	6/ Špatná organizace práce
7/ Dobré poměry mezi spolupracovníky	7/ Špatné poměry mezi spolupracovníky
8/ Pozorné jednání mistrů s pracovníky	8/ Nesprávné jednání mistrů s pracovníky

atd.

V jiném dotazníku se na jiné položky odpovídalo na několikabodové škále. Již z této dílčí ukázky je zřejmé, že se autoři snažili zachytit postoj k práci z nejrůznejších hledisek a dospět tak k dostatečně širokému pohledu, který zahrnuje celou řadu podstatných aspektů sledovaného jevu.

Z toho, co bylo řečeno, je snad zřejmé, že prvotní výběr položek při sestavování škál není jednoduchou záležitostí. Zjednodušený postup, opírající se pouze o intuitivní odhad, které položky asi jsou vhodné, musí nutně vést k chybám. Přípravná fáze a její pečlivé zhodnocení je nezbytným předpokladem pro to, abychom vůbec byli oprávnení škálovací šetření provádět. Obratné užití různých škálovacích technik nemůže překrýt nedostatky, vznikající z nesprávného výběru či formulace jednotlivých položek.

II.2. Platnost a spolehlivost škálovacích metod

Tak jako k jiným výzkumným metodám patří i k psychologickému škálování zcela nezbytně úvaha o jeho platnosti /validitě/ a spolehlivosti /reliabilitě/. V současné době je mnohem více známo o spolehlivosti jednotlivých škálovacích technik. Podobně jako v psychometrii se i při škálování uplatňují 3 techniky pro odhad spolehlivosti: a/ opakování zkoušky, b/ půlení zkoušky, c/ použití paralelních či vyrovnaných forem zkoušky. Při opakovaném měření se stejná škála užije dvakrát či vícekrát a výsledky se vzájemně srovnávají. Záskáváme tak informaci o stálosti výsledků v čase. Při metodě půlení předložíme osobám celou škálu, ale výsledky zpracujeme tak, že je rozdělíme na 2 části /první a druhá polovina nebo sudé a liché položky a pod./ a zjišťujeme míru shody; získáváme tak informace o vnitřní konzistenci škály. Mezi těmito dvěma metodami leží technika paralelných škál, která předpokládá, že máme k dispozici dvě stejně hodnotné verze téže škály. Jestliže obě formy předložíme současně, pak získáváme údaje blízké metodě půlení; předložíme-li je s delším časovým odstupem, blížíme se metodě opakované zkoušky.

Proti psychometrickým kritériím spolehlivosti jsou kritéria používaná při psychologickém škálování volnější. Jedná se většinou o jevy mnohem subtilnější a snáze ovlivnitelné, a proto též striktní požadavky na vysokou reliabilitu jsou hůře splnitelné. Při opakovaném měření může v průběhu času dojít ke značným posunům /např. změní se názor na posuzované osoby či jevy, změní se postoje v důsledku nových zkušeností atd./ a v důsledku toho se jen zdánlivě měří opětovně tentýž jev. Půlení škál je často neproveditelné, protože mnohé škály záměrně pracují s malým počtem položek a půlení může vést k velké náhodné chybě, která činí dojem malé spolehlivosti. Sestavení paralelních škál je velmi obtížné; právě proto se může snadno stát, že vznikne nesprávně dojem nespolehlivosti škálovacího postupu jen proto, že nebylo dostatečně přesně ověřeno, zda dvě formy škály jsou skutečně stejnomocné.

Celkově je však nutno konstatovat, že v literatuře se zatím neobjevilo příliš mnoho pokusů, ve kterých by se systematicky sledovala spolehlivost používaných škál a stanovila se pro ni pevná kritéria.

Neméně obtížná je situace z hlediska odhadu platnosti škálovacích technik. Často se uvádí, že platnost nějaké výzkumné metody je ověřena tím, že umožní na základě získaných či známých dat predikovat budoucí výsledky. Zde se však často zaměňuje metoda s určitou teorií a tím je celá situace zamlžena. Nabízí se určitá varianta: srovnávat data získaná škálováním u náhodně vybraných osob s hodnotami získanými stejnou metodou u přesně definovaného souboru osob, u kterých předpokládáme vyhraněné výsledky. Např. při měření určitého postoje zjišťujeme, zda výsledky diferencují dostatečně citlivě osoby s vyhraněným postojem od běžně vybraných osob /např. dobrovolní dárci krve budou mít asi kladnější vztahy ke svým bližním, k nemocným osobám a pod. než jiné skupiny pokusných osob/.

Jiná technika, která se někdy užívá, je posuzování, zda výsledky, získané při použití určité škálovací metody, jsou ve shodě s výsledky, které byly získány jinými metodami, či s předpoklady, které plynou z věcného rozboru problematiky. Tak např. můžeme očekávat, že data, získaná škálováním rizikového rozhodování, se budou shodovat s reálným rizikovým jednáním, které můžeme sledovat v životních podmínkách. Optimální je situace tehdy, když můžeme užít dvě různé metody pro sledování určitého jevu; přitom metoda A se považuje za prokázaně platnou a vhodnou, metoda B se s ní srovnává; z výsledků se pak odhaduje, zda a do jaké míry obě metody měří tentýž jev. Při sledování platnosti škálovácích procedur se tento postup užívá teprve v posledních letech, a to hlavně při hodnocení posuzovacích škál či některých škál pro zjišťování a měření postojů.

Platnost škálovacích technik bude stát zřejmě v popředí budoucího výzkumu. Existuje totiž reálné nebezpečí, že v současné konjunktuře různých - lépe či hůře vyvinutých - škál bude tato problematika zanedbána, což by jistě neprospělo ani dalšímu rozvoji škálovací metodologie, ani využívaní psychologického škálování ve společenské praxi.

II.3. Třídění škálovacích technik

Škálovacích technik, které se dnes v psychologii užívají, je celá řada a ani zdaleka neexistuje jednotné třídění. Torgerson /1968/ uvažuje o třech různých přístupech ke škálování:

a/ Přístup zaměřený na jednotlivé osoby. Zjištěné systematické rozdíly v reakcích osob na určité podněty či jevy jsou působeny individuálními rozdíly mezi lidmi. V podstatě se tedy škálují pokusné osoby, kterým se přiřazují určitá čísla /např. hodnocení pracovníků, posouzení žáků ve škole atd./. Tyto škálovací metody jsou velmi důkladně propracovány v oblasti mentálních testů, ale v ostatních oblastech psychologie se používaly jen sporadicky.

b/ Přístup zaměřený na podněty. Zjišťují se rozdíly v reakcích osob na určité podněty. Účelem je škálovat podněty a zjišťovat vztahy mezi nimi. Tyto druhy škál jsou v psychologickém výzkumu velmi časté a jsou též nejlépe propracovány.

c/ Přístup zaměřený na reakce či činnosti. Rozdíly v praktických a pozorovaných činnostech jsou připisovány interakci rozdílů mezi lidmi a mezi podněty. Účelem je škálovat buď osoby nebo podněty, anebo obé současně. V této oblasti existuje též řada metod, ale jedná se vesměs o složitější škály /většinou o různé formy multidimenzionálních škál/.

Existuje řada dalších třídění škálovacích technik; o některých se zmíníme později /Coombsovo třídění - viz Kap. XI., třídění multidimenzionálních škál- viz Kap. XIV./. Kritický rozbor různých principů třídění škál i škálovacích technik uvádí Berka /1972/.

II.4. Volba škálovací techniky

Závěrem ještě několik poznámek k výběru škálovacích technik. Současná literatura neposkytuje jednoznačné návody, kdy a za jakých okolností se může určitá metoda optimálně užívat. Proto jsou při práci se škálami kladeny na psychologa značné nároky. V prvé řadě se předpokládá dostatečně široká znalost škálovacích technik, které jsou k dispozici. Tato znalost znamená jednak zvládnutí gnoseologických a racionálních východisek jednotlivých postupů, jednak porozumění tomu, pro jaký typ otázky a pro jaká data se jednotlivé škálovací techniky dobře hodí. Z toho pak vyplývá, jaká metoda bude zvolena. V mnoha případech je volba omezena i možnostmi výpočetní techniky. V poslední době se vyvíjejí škálovací postupy, které předpokládají použití samočinných počítačů /většinou středních/. Mají-li být racionálně využity, pak musí psycholog být schopen nejen tyto škálovací techniky myšlenkově zvládnout, ale zároveň musí spolupracovat s dobrým programátorem. Většinou se jedná o programy, které byly vyvinuty pouze pro potřebu psychologie a nepatří ke standardním programům, které jsou běžně uváděny.

Z obecných pravidel bychom mohli uvést tato:

a/ Nepoužívej škálovací techniky, o které jsi se pouze doslechl či dočetl, aniž bys ji blíže poznal; snadno se dojde k falešným závěrům, plynoucím z neznalosti.

b/ Neužívej komplikovaných škálovacích technik tam, kde lze volit jednodušší postupy. Pořadová data pro řadu výzkumných účelů jsou dostatečně citlivá a jejich transformace do intervalových škál nemusí být však užitečná a nezbytná.

c/ Pracuj raději s menším počtem dobře vybraných a promyšlených položek než s rozsáhlou škálou, jejíž položky nejsou dostatečně prověřeny.

d/ Škálování není jednoduchou záležitostí a v celkové výzkumné strategii přichází v úvahu až v pozdějších fázích výzkumu. Vytvoření dobré škály předpokládá již dostatečné množství předcházejících informací a dat, a případně i pochopení vztahů mezi různými poznatky.

e/ Perfektní škálovací technika nemůže nahradit nejasně formulovaný výzkumný cíl a hypotézy a nemůže ani vylepšit nepřesně či dokonce chybně získaná data.

f/ Ač se škálování stává módní záležitostí, přece jen se valná většina psychologických výzkumů obejde bez něho. Což však neznamená, že základní informace o škálování by nepatřily do metodické výzbroje psychologa v druhé polovině 70. let.

Kapitola III.

METODA PÁROVÉHO SROVNÁVÁNÍ PODNĚTŮ

III.1. Ú v o d

Vytváření škál pomocí metody párového srovnávání určitých jevů patří v psychologii k nejstarším výzkumným metodám. Uvádí ji již Fechner v Elementech psychofyziky /1860/, ale za vlastního zakladatele je považován Thurstone /1927/. Po stránce technické a statistické byla tato metoda důkladně a detailně propracována /viz Guilford, 1954, Krasemannová, 1966, aj./. Její velkou výhodou je to, že ji lze použít pro nejrůznější účely, jako je třeba pro škálování vlastností osobnosti, motivů jednání, různých projevů chování, různých dimenzí předmětů vnějšího světa, i pro řadu jevů v sociální psychologii. Můžeme ji však rozumně užívat pouze tehdy, když počet podnětů, které srovnáváme, není příliš velký.

Základní myšlenkou metody párového srovnávání je postupné srovnávání každého podnětu se všemi ostatními. Vytvoří se všechny možné páry a zjišťuje se, jak často je podnět i posuzován jako lepší /větší, výhodnější, krásnější atd./ než podnět j. Odpověď "stejný" se nepřipouští; posuzovatel se musí vždy jednoznačně rozhodnout. Počet možných párů je určen vztahem $n(n-1)/2$, kde n značí počet podnětů. Tedy např. při 5 podnětech máme 10 párů, při 10 podnětech 45 párů, při 50 podnětech by bylo nutno porovnávat 1225 párů, což by zřejmě bylo experimentálně neúnosné. /O některých možnostech redukce počtu párů se zmíníme v části III.6./.

Teoretická východiska a matematicko-statistické základy metody párového srovnávání zde neuvádím. Psycholog, který se o ně speciálně zajímá, je najde nejlépe popsány v jiných učebnicích /Guilford, 1954, Torgerson, 1958/.

Data, na která je možno aplikovat metody párového srovnávání, můžeme získávat zhruba trojím způsobem:

a/ jedna osoba posuzuje všechny vytvořené páry podnětů mnohokrát,
b/ od mnoha osob žádáme posouzení všech párů jedenkrát,
c/ několik osob posuzuje všechny páry několikrát.

Volba mezi těmito postupy závisí na účelu experimentu, na povaze posuzovaných jevů, na očekávané velikosti individuálních rozdílů mezi posuzovateli, i na některých technických omezeních /např. daných tím, že nechceme pokus s jednou osobou příliš protahovat, apod./.

Jestliže chceme získat pro určitého jedince škálu, která odráží jeho osobní standard /případně chceme-li několik málo osob navzájem srovnat/, pak volíme variantu a/. Chceme-li získat jakousi průměrnou škálu pro celou populaci, pak volíme variantu b/ - ovšem za předpokladu, že máme reprezentativní výběr posuzovatelů. Jestliže nás individuální diference nezajímají a považujeme je za zanedbatelné, pak můžeme užít kteroukoli variantu; řídíme se spíše úvahou, zda by opakované posouzení stejných podnětových párů neovlivňovalo výsledky /zapamatování dříve uvedené preference!/, či zda by experiment netrval příliš dlouho a neobjevily se pak chyby plynoucí z únavy, napozornosti apod.

Vlastní sestavení škály závisí do jisté míry na zvoleném postupu získávání dat a na varianci získaných dat; z těchto důvodů Thurstone rozeznával 5 různých případů /blíže viz Guilford, 1954/. Pro náš účel zvolíme jako příklad takovou situaci, kdy posouzení provádí větší počet osob pouze jedenkrát; podobná situace se v praxi vyskytuje nejčastěji.

III.2. V y t v o ř e n í š k á l y

Při výzkumu rozhodování člověka jsme předložili 94 osobám /vojáci v základní službě/ k posouzení 7 jednoduchých her s pravidly shrnutými v tabulce III.1.

Posuzovatelům byly předloženy všechny možné dvojice her; tedy celkem 21 párů. Úkolem bylo určit, zda by raději hráli hru X či hru Y; vyskytovat se může výsledek X>Y či Y>X, ale nepřipouští se X = Y . Jestliže zvolené hře přiřadíme hodnotu 1 a nezvolené hře hodnotu 0, můžeme výsledky shrnout do matice D /dominanční/, která tvoří základ pro vytvoření intervalové škály. Získané výsledky jsou shrnuty v tabulce III.2.

V tabulce III.2. jsou získaná data uspořádána tak, že ve sloupcích je uveden počet posuzovatelů, kteří určitou hru považovali za výhodnější, než hru, která je uvedena v odpovídající řádce. Tedy např. 69 osob dávalo přednost hře B před hrou A a 25 osob naopak hře A před hrou B; 35 osob preferovalo hru G před hrou F a 59 osob preferovalo hru F před hrou G. Hodnoty v tabulce označíme f_{ij}. Je zřejmé, že musí vždy platit $f_{ij} + f_{ji} = N$ /počet posuzovatelů/.

Tab. III.1.

Pravidla 7 her, použitých při jejich párovém srovnávání

Hra	Pravděpodobnost výhry p	Výše výhry v	Hodnota hry p . v
A	1/8	7 Kčs	7/8
B	2/8	6 "	12/8
C	3/8	5 "	15/8
D	4/8	4 "	16/8
E	5/8	3 "	15/8
F	6/8	2 "	12/8
G	7/8	1 "	7/8

Tab. III.2.

Výchozí matice dat /D/ - párové srovnání 7 různých her. V jednotlivých polích je počet osob /f_{ij}/, které daly přednost hře i před hrou j.

Hra	A	B	C	D	E	F	G
A		69	75	76	67	64	56
B	25		67	71	58	56	44
C	19	27		67	49	44	37
D	18	23	27		51	42	32
E	27	36	45	43		42	35
F	30	38	50	52	52		35
G	38	50	57	62	59	59	

Data z tab. III.2 vyjádříme nyní v relativních hodnotách p_{ij} a stanovíme, jaká část posuzovatelů dávala přednost hře i před hrou j a naopak. Hodnoty v jednotlivých polích tab. III.2. dělíme proto počtem posuzovatelů, čili $p_{ij} = f_{ij}/N$. Kdyby jednotlivé hry byly posuzovány stejně, pak by platilo $p_{ij} \triangleq 0,5$. Skutečné výsledky jsou uvedeny v tab. III.3. Vidíme, že posouzení jednotlivých párů her se odchyluje od očekávané hodnoty /0,5/. V tomto momentě můžeme též srovnávat výsledky, získané od různě početných či různě složených skupin posuzovatelů.

Tab. III.3.

P - matice odvozené z tab. III.2. $/p_{ij}/$

Hra	A	B	C	D	E	F	G
A		0,734	0,798	0,808	0,713	0,681	0,596
B	0,266		0,713	0,755	0,617	0,596	0,468
C	0,202	0,287		0,713	0,521	0,468	0,394
D	0,192	0,245	0,287		0,543	0,447	0,340
E	0,287	0,383	0,479	0,457		0,447	0,372
F	0,319	0,404	0,532	0,553	0,553		0,372
G	0,404	0,532	0,606	0,660	0,628	0,628	

Kontrolu hodnot v tab. III.3. je možno provést ověřením vztahů $p_{ij} + p_{ji} = 1$; součet všech hodnot v tabulce se musí rovnat počtu párů /v našem příkladě 21/.

V dalším kroku převedeme matici P na matici z. Přitom vlastně určujeme odchylku zjištěných hodnot z jednotlivých polí tab. III.3.

od celkového teoretického průměru, t.j. od hodnoty p = 0,5, kterou bychom dostali, kdyby mezi podněty nebyl při posouzení žádný rozdíl /podrobný výklad viz Guilford, 1954, str. 160 a dále/. Používají se tabulky standardizovaného normálního rozložení, které najdeme ve všech učebnicích statistiky. Uvádíme je s určitou úpravou v dodatku jako tab. A. Při $p_{ij} > 0,50$ má odpovídající hodnota z_{ij} kladné znaménko; pro $p_{ij} < 0,50$ je hodnota z_{ij} záporná. Tabulka hodnot z je tedy numericky symetrická; pravá horní polovina je shodná s levou dolní polovinou; liší se pouze znaménka. Do diagonály, ve které v tab. III.2 a III.3. byla zatím prázdná políčka, zavedeme hodnoty 0,00 /vlastně se jedná o srovnání každého indikátoru sama se sebou při p = 0,5 /.

Pro lepší ilustraci postupu uvádím malé repetitorium pro transformaci hodnot p na hodnoty z : Ze základů statistiky víme, že při normálním standardizovaném rozložení jsou hodnoty na ose x vyjádřeny v jednotkách směrodatné odchylky /viz obr. III.1./. Určitou empirickou hodnotu můžeme vyjádřit jako odchýlení od průměru v jednotkách směrodatné odchylky. Platí vztah

$$z = \frac{X - \bar{X}}{\sigma}$$

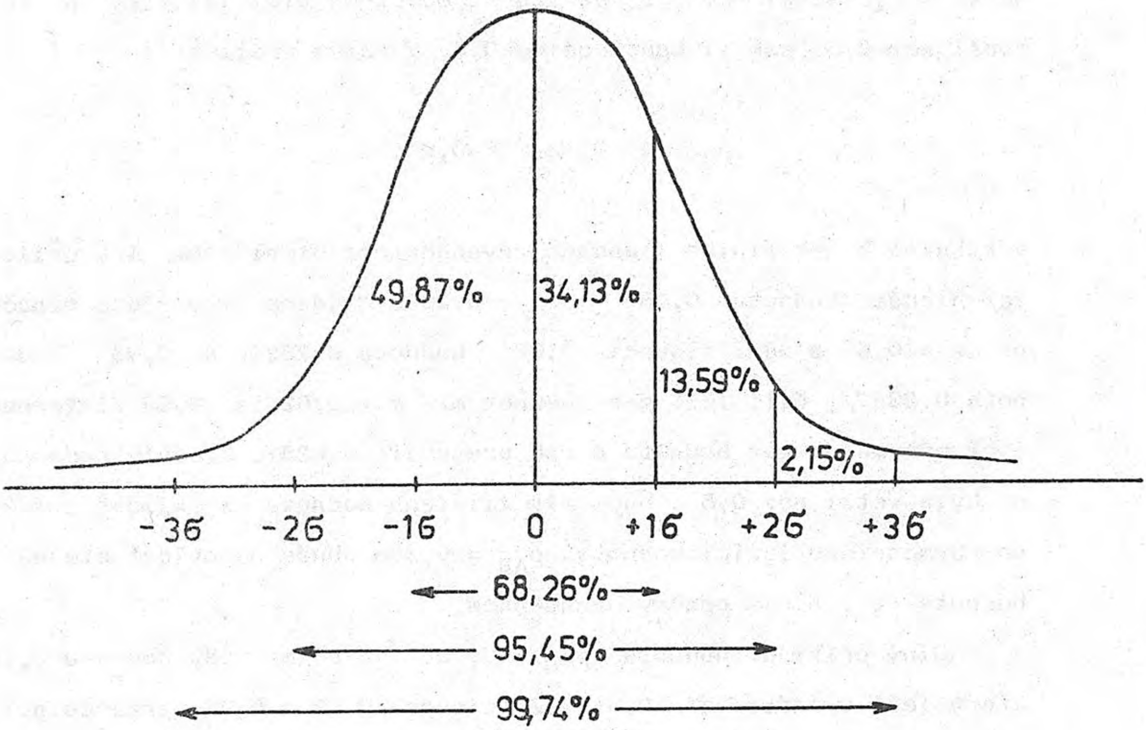
kde X = naměřená hodnota

$\bar{X}$ = průmer

σ = směrodatná odchylka

OBR. III. 1.

NORMÁLNÍ ROZLOŽENÍ.



Předpokládejme, že určitá veličina /třeba IQ/ má průměr $\bar{X} = 100$ a $\sigma = 15$. Chceme zjistit, jaké procento případů bude mít ve sledované veličině hodnotu 115 a více. Ze vztahu

$$z = \frac{X - \bar{X}}{\sigma} \quad \text{plyne} \quad \frac{115 - 100}{15} = +1 \;;$$

při hodnotě $z = +1$ leží 34,13% případů v rozmezí mezi $z = 0$ až $z = +1$; čili 15,87% má hodnotu větší než $z = +1$.

V našem případě je úloha opačná: známe hodnotu p /např. p_{BA} = 0,734/ a hledáme k ní odpovídající hodnotu z . Jestliže hodnota p je větší než 0,5, pak od ní odečteme 0,5; jestliže p je menší než 0,5, pak ji odečteme od 0,5. V našem případě

$$p_{BA} = 0,734 > 0,5 ;$$

v tabulkách normálního standardizovaného rozložení /tab. A v příloze/ hledáme hodnotu 0,734 - 0,5 = 0,234. Najdeme ji v řádce označené z = 0,6 a mezi sloupci 0,02 /hodnota 0,2324/ a 0,03 /hodnota 0,2357/, čili leží mezi hodnotami z = 0,62 a 0,63 /interpolací můžeme dostat hodnotu o řád přesněji: 0,625/. Protože hodnota p byla větší než 0,5 , bude mít zjištěná hodnota z kladné znaménko. Symetricky ležící hodnotě p_{AB} = 0,266 bude odpovídat stejná hodnota z , ale s opačným znaménkem.

Jiný příklad: hodnota p_{FG} = 0,628 ; hledáme tedy hodnotu 0,128, která leží v řádce z = 0,3 mezi sloupci 0,02 a 0,03; interpolací dostaneme hodnotu z = 0,327 .

Uvedeným postupem získáme matici z_{ij} , která je shrnuta v tabulce III.4.

Poslední krok, potřebný k sestavení škály, je připojen k tab. III.4. Hodnoty z se v jednotlivých sloupcích sečtou a vypočte se jejich průměr $\bar{X}$. Získáme tak škálové hodnoty jednotlivých podnětů ve formě odchylky od průměru všech škálových hodnot, který byl 0. V takto získané škále mají některé podněty záporné hodnoty.

Tab. III.4.

Matice z_{ij} odvozená z tab. III.3.

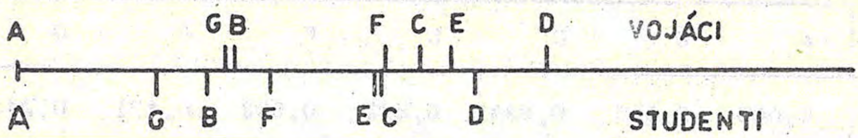
Hra	A	B	C	D	E	F	G
A	0,000	0,625	0,834	0,872	0,562	0,471	0,243
B	-0,625	0,000	0,562	0,690	0,298	0,243	-0,080
C	-0,834	-0,562	0,000	0,562	0,053	-0,080	-0,269
D	-0,872	-0,690	-0,562	0,000	0,108	-0,133	-0,413
E	-0,562	-0,298	-0,053	-0,108	0,000	-0,133	-0,327
F	-0,471	-0,243	0,080	0,133	0,133	0,000	-0,327
G	-0,243	0,080	0,269	0,413	0,327	0,327	0,000
$\sum$	-3,607	-1,088	1,130	2,562	1,481	0,695	-1,173
$\bar{x}$	-0,515	-0,155	0,161	0,366	0,212	0,099	-0,168
H	0,000	0,360	0,676	0,881	0,727	0.614	0,347

/ $H = \bar{x} + 0,515$ /

Provádí se obvykle ještě tato úprava: ke všem hodnotám přičteme konstantu, která se rovná absolutní hodnotě podnětu s největší zápornou hodnotou. Tím tento podnět získá hodnotu 0,00 a ostatní jsou vesměs pozitivní; jejich vzájemné vzdálenosti se přitom nemění. V našem případě použijeme konstantu +0,515; konečná škála je zachycena v posledním řádku tab. III.4. a graficky znázorněna na obr. III.2.

OBR. III. 2.

VÝSLEDNÁ ŠKÁLA (PÁROVÉ SROVNÁVÁNÍ 7 HER)



Ze získané škály plyne, že nejčastěji je volena hra D, na druhém místě hra E; hry B a G jsou voleny prakticky stejně často atd. Takto získanou škálu můžeme porovnat s objektivními hodnotami jednotlivých her /viz tab. III.1./.

Na okraji poznamenávám, že zhruba podobnou škálu jsem získal, když sázky hodnotila skupina pozorovatelů zhruba stejného stáří, ale jiných sociálních podmínek - 126 studentů psychologie. Jejich výsledky jsou zachyceny též na obr. III.2. /dolní škála/.

Popsaný výpočet škálových hodnot nelze provést v případě, kdy se při posouzení některého ze srovnávaných párů objeví stoprocentní shoda mezi posuzovateli. Kdyby např. všichni preferovali X před Y /a tedy žádný Y před X/, pak odpovídající hodnoty z by byly $\pm\infty$ a vznikla by nereálná škála. K podobným případům může dojít tehdy, když extrémní hodnoty škály se od sebe velmi nápadně liší /i když podněty, které leží mezi nimi, jsou vzájemně jen málo odlišné/, anebo tehdy, když si jeden posuzovatel formuluje určité posouzení, které stabilně opakuje. Většinou se uvádí, že maximální rozdíl v posouzení dvou podnětů může být 97% : 3%. Pokud se extrémní hodnoty v posouzení skutečně vyskytnou, je třeba je vyloučit; škála je pak sestavována z neúplné výchozí matice /viz Esser, 1970, str. 204 a dále, nebo Torgerson, 1958, str. 173 a dále/.

Je třeba ještě upozornit, že uvedený postup má další modifikace, a to tehdy, když variance výsledků ve výchozí matici je příliš veliká. Tyto úpravy zde již neuvádím; čtenář je najde v učebnicích, o kterých byla zmínka již dříve /např. Krassemannová 1966, Guilford, 1954, aj./.

III.3. V ý z n a m n o s t r o z d í l ů m e z i p o l o h o u j e d n o t l i v ý c h b o d ů n a š k á l e

Dunn a King /1969/ navrhli zjednodušenou formu vytváření škály s využitím metody párového srovnávání podnětů. Jejich postup má několik výhod. Umožňuje určit významnost rozdílů v umístění jednotlivých podnětů na vytvořené škále. Je možno stanovit, zda zjištěné rozdíly jsou v mezích náhodného kolísání, či zda je možno podněty kategorizovat do několika odlišných skupin, anebo zda se jednotlivé podněty svým umístěním na škále od sebe liší. Zároveň navržený postup umožňuje stanovit minimální počet posuzovatelů, který je nezbytný k tomu, aby se rozdíly v posouzení jednotlivých podnětů mohly dostatečně a citlivě prokázat. Kromě toho je výpočet škály jednodušší.

Postup obou autorů budu demonstrovat na stejných datech, která jsem použil v části III.2. /viz tab. III.2./. Jedná se tedy o párové srovnávání 7 podnětů, které bylo provedeno 94 posuzovateli.

Před vlastním pokusem se nejprve stanoví minimální počet posuzovatelů, který je nezbytný, aby bylo možno testovat při zvolené hladině významnosti α rozdíly mezi umístěním jednotlivých podnětů na škále. Tyto hodnoty byly vypočítány - způsob výpočtu pominu - a tabulovány. Pro účel tohoto výkladu předkládám upravenou a částečně též rozšířenou tabulku pro 3 - 10 posuzovaných prvků při hladině významnosti $\alpha = 0{,}05$, resp. 0,01.

Tab. III.5.

Minimální počet posuzovatelů, potřebných pro stanovení významnosti rozdílů mezi jednotlivými body na výsledné škále /upraveno dle Dunna a Kinga, 1969/

Počet podnětů	Hladina významnosti α	
	0,01	0,05
3	17	11
4	33	22
5	53	38
6	80	57
7	112	82
8	149	110
9	192	144
10	245	183

V našem příkladě srovnáváme 7 podnětů; počet posuzovatelů je 94. Tento počet je zřejmě dostatečný, jestliže se spokojíme s hladinou významnosti $\alpha = 0{,}05$; minimální počet posuzovatelů by musel být 32. Pokud bychom zvolili hladinu významnosti 0,01 , pak bychom museli mít alespoň 112 posuzovatelů.

Při vlastním sestavení škály postupujeme takto /výpočty jsou shrnuty v tab. III.6./ : Výsledky jednotlivých osob se sloučí a vytvoří se preferenční matice /viz tab. III.2./. Nyní zjistíme, kolikrát byl určitý podnět preferován před ostatními /tedy sečteme hodnoty ve sloupcích tab. III.2./. K vypočteným hodnotám přičteme vždy 94, tj. počet posuzovatelů; toto číslo by mělo být vlastně ve volných diagonálních polích tab. III.2. Výsledné součty značíme R_i . Pro kontrolu výpočtu můžeme ověřit, že

$$\sum R_i = N . I . I + 1 / 2$$

/N = počet posuzovatelů; I = počet podnětů/

Škálové hodnoty jednotlivých podnětů získáme transformací výchozích dat na škálu s rozsahem od 0,00 do 100. Postupujeme podle vzorce

$$S = \frac{R_i - R_{min}}{R_{max} - R_{min}} . 100$$

kde R_i = počet preferencí jednotlivých podnětů, zvětšený o počet posuzovatelů

R_{min} = N /hodnota, kterou bychom dostali, kdyby i-tý podnět nebyl ani jednou preferován/ = 94

R_{max} = N.I = 658 /hodnota, kterou bychom získali, kdyby i-tý podnět byl vždy preferován/

Tab. III.6.

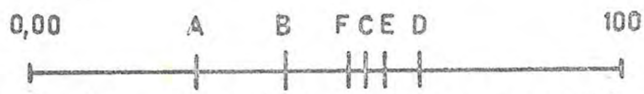
Postup výpočtu pro stanovení významnosti rozdílů mezi jednotlivými body škály /Data z tab. III.2./

Hra	A	B	C	D	E	F	G
Celkový počet preferencí	157	243	321	371	336	307	239
N	+94	+94	+94	+94	+94	+94	+94
R_i	251	337	415	465	430	401	333
S	28	43	57	66	60	54	42

Vypočtené hodnoty S jsou zapsány v poslední řádce tab. III.6. a můžeme je graficky znázornit takto:

OBR. III. 3.

VÝSLEDNÁ ŠKÁLA



Vlastní výpočet statistické významnosti rozdílů mezi posouzením jednotlivých podnětů předpokládá, že známe kritickou hodnotu, která je již statisticky významná. Vypočteme ji ze vztahu

$$CR = V \cdot Q_a$$

kde $V = \sqrt{\frac{N \cdot I \cdot (I+1)}{12}}$

Q_a = jsou hodnoty, uvedené v tab. III.7.

Tab. III.7.

Hodnoty Q_a pro výpočet statistických rozdílů mezi podněty při párovém posuzování /upraveno dle Dunn, King, 1969/

Počet podnětů	Hladina významnosti α	
	0,05	0,01
3	3,314	4,120
4	3,633	4,403
5	3,858	4,603
6	4,030	4,757
7	4,170	4,882
8	4,286	4,987
9	4,387	5,078
10	4,474	5,157

V našem příkladě při hladině významnosti α = 0,05 platí:

$CR = 4{,}170 \cdot \sqrt{94 \cdot 7 \cdot 8 / 12} = 4{,}170 \cdot 20{,}94 \doteq 87$

Podněty, které se od sebe liší alespoň hodnotou 87, jsou statisticky významně odlišné. Zvolíme-li hladinu významnosti α = 0,01, pak kritická hodnota CR = 4,882 . 20,94 ≐ 102. Pro lepší vhled si hodnoty R_i znovu vyjádříme v tab. III.8. Podněty jsou seřazeny od nejvíce preferovaného po nejméně preferovaný. Stanovíme rozdíly hodnot R_i /rozdíl mezi preferencemi hry D a E je 465 - 430 = 35, mezi D a C je 465 - 415 = 50 atd./; rozdíly vyšší než vypočtené hodnoty CR jsou statisticky významné.

Tab. III.8.

Významnost rozdílů mezi umístěním jednotlivých podnětů na výsledné škále

Hra	D	E	C	F	B	G
E	35					
C	50	15				
F	60	29	14			
B	128^{++}	93^{+}	78	64		
G	132^{++}	97^{+}	82	68	4	
A	214^{++}	179^{++}	164^{++}	150^{++}	86	82

$^{++}$... $p < 0{,}01$ $^{+}$... $p < 0{,}05$

Z výsledků plyne, že hra D je velmi významně preferována před hrami B, G, A; od ostatních se významně neliší. Podobně hra E se liší od her B,G,A a konečně hry C a F jsou statisticky velmi významně preferovány před hrou A.

Obvykle se ještě stanoví index škálovatelnosti /IŠ/; je dán podílem mezi počtem párů podnětů, které se od sebe statisticky liší, a počtem všech párů. V našem příkladě IŠ = 8/21 = 0,381. Jestliže IŠ = 1,00, pak máme škálu, ve které se všechny její body od sebe navzájem statisticky liší; jestliže IŠ = 0, pak se jednotlivé škálové body od sebe neliší.

Vyložený postup vytvoření škály má několik výhod: Vznikne škala, která má smysluplně zakotveny koncové body /minimum a maximum preferencí/. Výsledky nejsou ovlivněny extrémními rozdíly v posuzování jednotlivých podnětů; jestliže všichni posuzovatelé dávají přednost podnětu X před Y, pak je možno s tímto údajem pracovat, zatímco při postupu, popsaném v části III.2., tento výsledek znemožňoval vytvoření škály. Je možno předem racionálně stanovit velikost výběru posuzovatelů a testovat významnost rozdílů v posouzení jednotlivých podnětů.

Výsledná škála se podle srovnávacích údajů Dunna a Kinga /1969/ velmi dobře shoduje s výsledky, získanými jinými postupy. I když navržený postup se zatím objevuje v literatuře jen velmi řídce, lze jej doporučit k používání.

III.4. Zjištění stálosti posudků jednotlivých osob

Při použití metody párového srovnávání můžeme často zjistit určitou nestálost či nepravidelnost v posuzování. Při srovnávání 3 podnětů /A, B, C/ očekáváme, že posuzovatel, který prohlásí jev A za lepší než jev B, a jev B za lepší než jev C, nutně prohlásí také jev A za lepší než jev C. Ve skutečnosti se však často objeví tzv. kruhová triáda: A > B , B > C , C > A . K podobné nestálosti v posuzování může dojít z různých příčin: a/ posuzované jevy jsou velmi podobné a posuzovatel se obtížně rozhoduje, b/ posuzované jevy neleží na jediném kontinuu a mohou být posuzovány z různých hledisek, která nemusí vzájemně korelovat /např. tři jídla posuzujeme podle chuťových preferencí, ale též dle ceny; chutnější jídlo třeba odmítneme, že cena je příliš vysoká; v jiné situaci převládne gurmánství nad ekonomií/, c/ nepozornost či únava posuzovatele při příliš dlouhém pokuse, d/ různé jiné náhodné vlivy, které působí na posuzovatele.

Velký počet kruhových triád svědčí o nespolehlivém výkonu posuzovatele; proto ani výsledná škála nemůže být spolehlivá. Je proto užitečné se orientovat, zda jednotliví posuzovatelé pracovali pravidelně či nikoliv, případně vyloučit z nich ty, jejichž posudky jsou značně nepravidelné a víceznačné. K tomuto účelu vyvinul Kendall /1948/ koeficient konzistence - zeta /Coefficient of Consistency/. Vychází se z výsledků jednotlivých posuzovatelů a zjišťuje se, zda počet kruhových triád je zanedbatelný či nikoliv. Uvedený postup se tedy používá individuálně pro jednotlivé posuzovatele.

Uveďme příklad jednoho z posuzovatelů z předchozího výzkumu rozhodování mezi volbou různých her. Jeho posudky jsou uvedeny v tab. III.9.

Počet triád by bylo možno odvodit přímým nahlížením do tab. III.9. /či do její grafické transformace/, ale tento postup je zdlouhavý a nespolehlivý. V uvedeném příkladě dal posuzovatel přednost hře E před C, hře C před F, ale též hře F před E. Podobná kruhová triáda je mezi hrami C, D, E a též mezi A, B, G.

Kendall odvodil následující vzorec pro zjištění počtu kruhových triád:

$$d = \frac{1}{12} n \left(n-1\right)\left(2n-1\right) - \frac{1}{2} \left(\sum h_i \right)^2$$

kde d = počet zjištěných triád

n = počet posuzovaných podnětů

$\sum h_i$ = součet hodnot v jednotlivých sloupcích /t.j. počet preferencí i-tého podnětu/

V našem případě platí

$$d = \frac{1}{12} \cdot 7 \cdot (7-1)(14-1) - \frac{1}{2} \cdot 85 = \frac{91}{2} - \frac{85}{2} = 3$$

Tab. III.9.

Párové srovnávání 7 her, provedené posuzovatelem X.Y.

Hra	A	B	C	D	E	F	G	$\sum$
A		1	1	1	1	1	0	
B	0		1	1	1	1	1	
C	0	0		0	1	0	0	
D	0	0	1		0	0	0	
E	0	0	0	1		1	0	
F	0	0	1	1	0		0	
G	1	0	1	1	1	1		
$\sum h_i$	1	1	5	5	4	4	1	21
$(\sum h_i)^2$	1	1	25	25	16	16	1	85

Chceme určit, zda vypočtená hodnota d je v mezích náhodného kolísání posudků, a je proto zanedbatelná či nikoliv. Kendall stanovil, že při lichém počtu posuzovaných jevů může být maximálně $(n^3-n)/24$ kruhových triád; při sudém počtu pozorovaných jevů jich může být $(n^3-4n)/24$. Koeficient konzistence - zeta pak definoval takto:

$$\zeta = \begin{cases} 1 - \dfrac{24\,d}{n^3-n} & \text{při lichém } n \\[2ex] 1 - \dfrac{24\,d}{n^3-4n} & \text{při sudém } n \end{cases}$$

Jestliže se nevyskytne žádná kruhová triáda, pak $\zeta = 1$; při samých triádách /zcela nekonzistentní postup/ $\zeta = 0$. V našem případě /při n=7 a d=3/ tedy platí

$$\zeta = 1 - \frac{24\,.\,3}{7^3 - 7} = 1 - \frac{72}{336} = 1 - 0{,}214 = +0{,}786$$

Statistická významnost této hodnoty se zjišťuje přes χ^2 rozložení podle vzorce:

$$\chi^2 = \frac{8}{n-4}\left[\frac{1}{4}\binom{n}{3} - d \pm \frac{1}{2}\right] + v$$

kde n = počet posuzovaných jevů

d = počet zjištěných triád

v = počet stupňů volnosti, který se určuje ze vztahu

$$v = \frac{n\,(n-1)\,(n-2)}{(n-4)^2}$$

V našem případě $v = \frac{7.6.5}{3.3} = 23,33$, a tedy

$$\chi^2 = \frac{8}{3}\left[\frac{1}{4}\binom{7}{3} - 3 + \frac{1}{2}\right] + 23,33 \qquad \binom{n}{3} = \frac{n!}{3!\;(n-3)!}$$

$$= \frac{8}{3}\left[\frac{1}{4}\,.\,35 - 3 + \frac{1}{2}\right] + 23,33$$

$$= \frac{8}{3}\,.\,6,25 + 23,33$$

$$= 16,67 + 23,33 = 40,00$$

Z tabulek rozložení χ^2 /viz příloha, tab. B/ zjistíme, že vypočtená hodnota $\chi^2 = 40$ při 23,33 stupních volnosti je statisticky významná ($p < 0,02$) . Při interpretaci však je třeba dbát na to, že vypočtená pravděpodobnost je komplementární k hodnotám, které jsou tabelovány. Výsledky tedy interpretujeme obráceně, t.j. tam, kde získáme statisticky významnou hodnotu χ^2, považujeme počet triád za zanedbatelný. V našem případě, pokud volíme 5% hladinu významnosti, můžeme konstatovat, že sledovaný posuzovatel postupoval systematicky, a zjištěné tři triády považujeme za chyby, které se vyskytují v rámci náhodného kolísání.

Pro malý počet posuzovaných jevů vyvinul Kendall tabulku, ze které je možno při znalosti hodnoty d přímo určit, zda počet triád je zanedbatelný či nikoliv, aniž bychom počítali koeficient konzistence /Kendall, 1948, str. 190, tab. 9/. Upravenou tabulku zde uvádím jako tab. III.10.

Tab. III.10.

Tabulka kritických hodnot d pro stanovení významnosti počtu kruhových t. d

	n=4	n=5	n=6	n=7
$p < 0{,}05$		1	2	4
$p < 0{,}01$			0	1

V našem případě při n=7 jsme získali hodnotu d=3, která je nižší než kritická hodnota při $p<0{,}05$, což se shoduje s výše provedeným výpočtem. Z tabulky dále vidíme, že teprve tehdy, kdybychom u určitého posuzovatele zjistili pět či více triád, byl by tento počet závažný a svědčil by o zřejmě nepravidelném postupu posuzování /při 5% hladině významnosti/.

Zjistíme-li u některého posuzovatele významnou nepravidelnost v posuzování, pak jej můžeme buď vyloučit /tam, kde odhadujeme, že nepravidelnost je působena jeho nepozorností či podobnými chybami/, anebo můžeme blížc zjišťovat, na základě jaké úvahy došel ke svému řešení. Jestliže podobně postupuje většina posuzovatelů /kteří se však vcelku shodují - viz dále/, pak se zřejmě jedná o systematickou chybu, která vyplývá z povahy podnětů. Jsou zřejmě posuzovány dle různých dimenzí a tento způsob posouzení se prosazuje analogicky u většiny posuzovatelů.

III.5. Zjištění shody mezi posuzovateli

Jestliže pracujeme s větším počtem posuzovatelů, je možno se ptát, zda se jejich posudky vcelku shodují, či zda jsou mezi nimi nápadné rozdíly. Z tohoto zjištění pak můžeme odhadnout, zda jsme pracovali se skupinou, která je ve svém posuzování homogenní či nikoliv.

Pro tento účel Kendall /1948/ vyvinul koeficient shody /Coefficient of Agreement/, značený μ . Je sestaven tak, že při úplné shodě mezi posuzovateli je hodnota μ = 1 a při úplné neshodě je záporná. Její velikost závisí na počtu posuzovatelů (m) a je dána vztahem

μ = -1 / m při lichém počtu,

resp. μ = -1 / m-1 při sudém počtu.

V našem případě při úplné neshodě 94 posuzovatelů bychom získali hodnotu μ =-0,011.

Pro výpočet koeficientu shody vycházíme z původních dat /viz tab. III.2./. Upravíme ji však tak, že seřadíme jednotlivé hry do pořadí, které odpovídá zjištěné škále; při tom nejníže posuzovaný podnět je na prvém místě v pořadí a nejvýše posuzovaný podnět na posledním. Pracujeme pouze s polovinou výchozí tabulky. Rozdělíme ji úhlopříčně /od levého horního rohu k pravému dolnímu/ a zvolíme tu polovinu, ve které leží méně dat. V našem případě se jedná o polovinu, která je pod úhlopříčkou. Dostaneme data, shrnutá v tab. III.11.

Tab. III.11.

Data pro výpočet shody mezi posuzovateli /Úprava tab.III.2./

Hra	A	G	B	F	C	E	D	$\sum$
A								
G	38							
B	25	44						
F	30	35	38					
C	19	37	27	44				
E	27	35	36	42	45			
D	18	32	23	42	27	51		
$\sum X_{ij}$	157	183	124	128	72	51	.	715
$\sum X_{ij}^2$	4 383	6779	4000	5464	2754	2601		25.981

Koeficient shody se dle Kendalla vypočítá takto:

$$\mu = \frac{2\left[\sum X_{ij}^2 - m\sum X_{ij} + \binom{m}{2}\binom{n}{2}\right]}{\binom{m}{2}\binom{n}{2}} - 1$$

kde m = počet posuzovatelů /94/

n = počet posuzovaných podnětů /7/

$$\binom{m}{2} = \frac{m\,(m-1)}{2}$$

$$\binom{n}{2} = \frac{n\,(n-1)}{2}$$

$\sum X_{ij}$ = součet všech hodnot v tab. III.11. /715/

Po dosazení empirických dat získáváme hodnotu

$$\mu = \frac{2\,.\left[25\,981 - 94\,.\,715 + \left(\frac{94\,.\,93}{2}\right)\left(\frac{7\,.\,6}{2}\right)\right]}{\left(\frac{94\,.\,93}{2}\right)\left(\frac{7\,.\,6}{2}\right)} - 1 =$$

$$= \frac{2\,.\,(25\,981 - 67\,210 + 91\,791)}{91\,791} - 1 = \frac{2\,.\,50\,562}{91\,791} - 1 =$$

$$= \frac{101\,124}{91\,791} - 1 = 1{,}102 - 1 = +\,0{,}102$$

Významnost této hodnoty opět zjišťujeme pomocí χ^2 rozložení dle vztahu

$$\chi^2 = \left[\frac{4}{m-2}\right]\left[\sum X_{ij}^2 - m\sum X_{ij} + \binom{m}{2}\binom{n}{2} - \frac{1}{2}\binom{n}{2}\binom{m}{2}\left(\frac{m-3}{m-2}\right)\right]$$

Když dosadíme příslušné empirické hodnoty, vypočteme

$$\chi^2 = \left(\frac{4}{94-2}\right)\left[25\ 981 - 94 \cdot 715 + \frac{94 \cdot 93}{2} \cdot \frac{7 \cdot 6}{2} - \right.$$

$$\left. - \frac{1}{2}\left(\frac{7 \cdot 6}{2} \cdot \frac{94 \cdot 93}{2} \cdot \frac{94 - 3}{94 - 2}\right)\right] = 224,75$$

Počet stupňů volnosti stanovíme ze vztahu:

$$v = \binom{n}{2} \frac{m\ (m-1)}{(m-2)^2} = 21 \cdot \frac{94 \cdot 93}{92 \cdot 92} = 21,69$$

Z tabulek rozložení χ^2 zjistíme, že hodnota $\chi^2 = 224,75$ při 22 stupních volnosti je statisticky velmi významná, a proto můžeme konstatovat, že skupina 94 posuzovatelů srovnávala jednotlivé hry stejně a rozdíly mezi nimi jsou v rámci náhodného kolísání.

III.6. Metody redukce výchozí matice

V úvodu jsem upozornil, že při rostoucím počtu podnětů vzniká velké množství párů, které je experimentálně neúnosné. Uvažuje se proto, jak by bylo možno jejich počet snížit, aniž by tím utrpěla kvalita získávaných dat. Existuje několik postupů:

První možností je použít vyrovnané neúplné bloky. Postupujeme tak, že jednotlivé osoby párově srovnávají jen část podnětů /Maxwell, 1974/. Chceme např. srovnávat 6 podnětů, ze kterých se vytvoří 15 párů. Tento počet se však zdá z určitých důvodů příliš velký a považujeme za únosné použít u každého posuzovatele pouze 4 podněty. Postupujeme tak, že vytvoříme všechny možné čtveřice ze 6 prvků a jednotlivým osobám se předloží vždy jedna z nich. Po sloučení všech dat platí, že každý podnět byl srovnáván stejně často se všemi ostatními. Strukturu pokusu si znázorníme v tab. III.12.

Tab. III.12.

Příklad systému vyvážených neúplných bloků /ze 6 podnětů se užívají pouze 4/

Blok	P o d n ě t y					
1	A	B	C	D		
2	A	B	C		E	
3	A	B	C			F
4	A	B		D	E	
5	A	B		D		F
6	A	B			E	F
7	A		C	D	E	
8	A		C	D		F
9	A		C		E	F
10	A			D	E	F
11		B	C	D	E	
12		B	C	D		F
13		B	C		E	F
14		B		D	E	F
15			C	D	E	F

Z tab. III.12 vidíme, že první osoba srovnává podněty A, B, C, D, druhá osoba podněty A, B, C, E atd. Celkem se vyskytne 15 čtveřic podnětů, což znamená, že počet osob bude dán násobkem 15. Každý podnět je v této sestavě srovnáván s kterýmkoliv z ostatních podnětů celkem šestkrát. Chceme-li mít alespoň 60 porovnání všech párů, budeme potřebovat celkem 150 posuzovatelů. Při větším počtu podnětů však narůstá i počet bloků. Kdybychom třeba při 9 podnětech chtěli, aby jednotlivé osoby srovnávaly pouze 7 z nich, pak toho docílíme při 36 blocích; v této soustavě se každý podnět s každým vyskytne 22 krát. Kdybychom chtěli z 10 podnětů vybírat vždy 5 podnětů, pak bychom potřebovali celkem 252 bloků a v jejich rámci by každá dvojice podnětů byla srovnána 56 krát. Z těchto úvah vyplývá, že používání vyrovnaných neúplných bloků má řadu těžkostí. Počet srovnávaných podnětových párů sice klesne, ale počet posuzovatelů naopak narůstá; proto tento postup může být v některých případech jen těžko použitelný.

Snížení počtu pozorování můžeme docílit i tím, že část podnětů použijeme jako standardní; ostatní k nim systematicky přirovnáváme. Tak např. při srovnávání 9 podnětů /A... I/ bychom měli 36 párů podnětů. Zvolíme 4 podněty /třeba B, D, F, H/ jako standardní a s každým z nich srovnáme všechny ostatní. Tedy podnět A se srovnává pouze s B, D, F, H, podnět B se srovnává se všemi podněty, podnět C opět jen s B, D, F, H atd. Počet srovnávaných párů se sníží na 26.

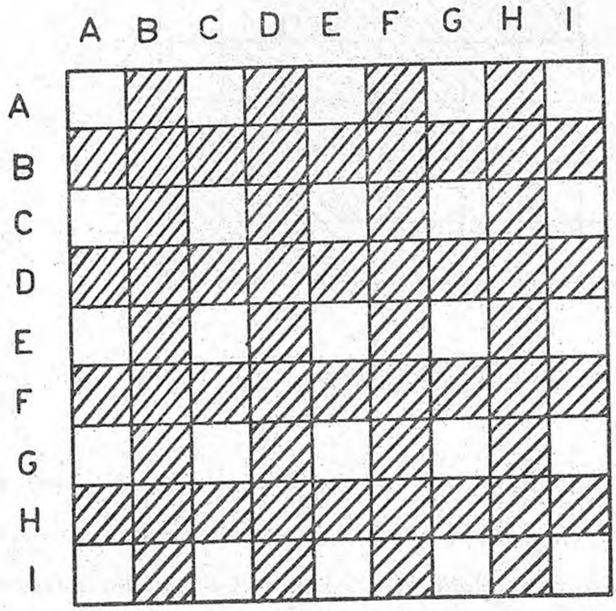
Výchozí matici máme schematicky naznačenou na obr. III.4. Počet výsledních párů zjistíme snadno ze vztahu $rs - r(r+1)/2$, kde s je počet všech podnětů (9) a r je počet vybraných standardů (4).

Jestliže známe alespoň zhruba očekávané pořadí podnětů na sledované škále, můžeme zvolit další cestu pro snížení počtu srovnávaných párů. Podněty seřadíme podle očekávaného pořadí a srovnáváme pouze ty, které leží blízko sebe /třeba vždy 5 sousedních podnětů/. Tento postup je schematicky znázorněn na obr. III.5. Vyloučí se vlastně takové páry podnětů, které se pravděpodobně značně liší a u kterých bychom asi získávali jednoznačné výsledky.

Některé další metody, jak je možno snížit počet posuzovaných párů, shrnuje Torgerson /1958, str. 191 a dále/.

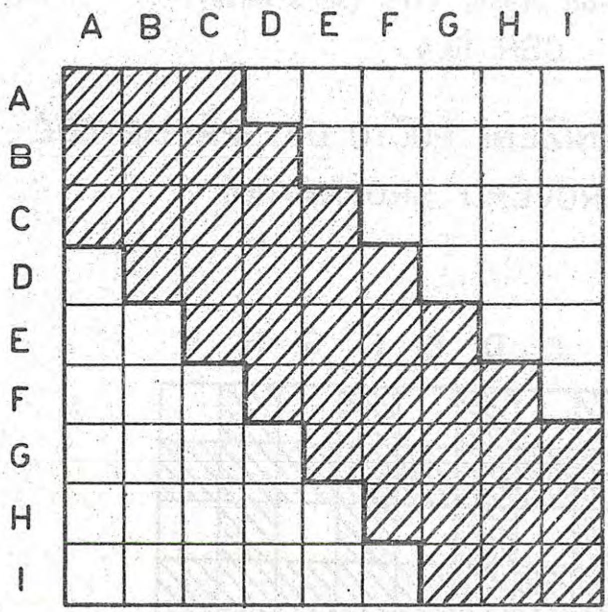
OBR. III.4.

PŘÍKLAD SNÍŽENÍ POČTU DAT PŘI METODĚ PÁROVÉHO SROVNÁVÁNÍ.



OBR. III.5.

JINÝ PŘÍKLAD SNÍŽENÍ DAT PŘI METODĚ PÁROVÉHO SROVNÁVÁNÍ.



III.7. S n í ž e n í   e x p e r i m e n t á l n í c h   c h y b

Při použití metody párového srovnávání se do výzkumu může vloudit řada chyb, které zkreslují výsledky. Při neobratném postupu může dokonce dojít ke znehodnocení celého výzkumného šetření. Časté experimentální chyby popíši; o jejich odstranění či alespoň neutralizování je třeba uvažovat již ve fázi přípravy experimentu.

Pracujeme-li s větším počtem podnětů, a tedy s velkým počtem srovnávaných dvojic, musíme počítat s časovou chybou. Objevují se nepřesnosti v začátečních fázích, kdy posuzovatelé nejsou ještě s celou situací dostatečně seznámeni a nevytvořili si určitý postup. Tento druh chyby pak většinou mizí pod vlivem nácviku. Při dlouhých sériích dochází k chybám, působeným únavou, ztrátou zájmu o pokus i celkovým snížením pozornosti. Oba tyto rušivé vlivy můžeme do jisté míry odstranit tím, že pracujeme se zacvičenými a dobře motivovanými posuzovateli, ale ani u nich je nelze zcela vyloučit. Proto je vhodnější vliv časové chyby pravidelně rozložit na všechny dvojice podnětů. Postupuje se většinou tak, že se používá různý sled podnětových dvojic u různých posuzovatelů. Někdy se posuzovatelé rozdělí na poloviny a oběma polovinám se předkládají podnětové dvojice v obráceném pořadí. Jindy se vytvoří několik skupin a pořadí dvojic se prostě posune. Tak třeba při 40 dvojicích část osob začíná od první dvojice a postupuje dle předem stanoveného pořadí; druhá skupina začne 10. dvojicí, pokračuje dále až do 40. dvojice, a pak teprve se předkládá prvních 10 dvojic. Třetí skupina začne od 20. dvojice; čtvrtá od 30. dvojice.

Druhým zdrojem možných chyb je pořadí podnětů ve srovnávaném páru. Doporučuje se měnit pořadí tak, aby určitý podnět byl zhruba v polovině párů na prvém místě a v polovině na druhém. Často se doporučuje i vyvážení v rámci každého páru, t.j. žádá se, aby polovina posuzovatelů měla dvojici podnětů předloženou v pořadí j - k /dal byste přednost podnětu j nebo podnětu k ?/ a druhá polovina v pořadí k - j /dal byste přednost podnětu k nebo podnětu j ?/.

Jiným zdrojem chyb mohou být různé konfigurace, které vznikají při nevhodném pořadí podnětových dvojic. Není vhodné, když ve dvou bezprostředně za sebou předložených dvojicích se objevuje tentýž prvek /po páru k - j následuje pár k - l /. Tam, kde známe nebo alespoň očekáváme určité vztahy mezi podněty, je vhodné, aby "správné" odpovědi byly stejnoměrně, ale přitom nesystematicky rozloženy na první i druhé místo ve dvojicích. Obvykle se též doporučuje, aby se měnila obtížnost posuzování a nekumulovaly se dvojice, mezi nimiž jsou minimální rozdíly, či naopak dvojice, u kterých lze předpokládat snadné rozhodnutí.

Kapitola IV.

METODA POŘADOVÝCH STUPNIC

IV.1. Ú v o d

Tato metoda se používá při psychologickém škálování velmi často. Má dvě velké výhody: jednak je možno ji užít při posuzování a srovnávání většího množství podnětů /což - jak jsme viděli - nelze dobře provést metodou párového srovnávání/, jednak je aplikovatelná na nejrůznější problémy počínaje od celkového hodnocení osob přes srovnávání uměleckých děl až po posouzení řady vlastností různých předmětů. Sestavení posuzovaných jevů do pořadí předpokládá, že posuzovatel je bude navzájem srovnávat; tím se získaná data blíží informacím, získaným metodou párového srovnávání podnětů. Další výhodou metody pořadových stupnic je, že můžeme snadno pracovat s řadou posuzovatelů současně,a tím poměrně rychle a s malými náklady získat rozsáhlá data. Podobně jako při metodě párového srovnávání i zde dospíváme v konečné fázi k intervalové stupnici - jedná se tedy o metodu, kdy z pořadových dat získáme stupnici vyššího řádu.

Metoda pořadových stupnic se v psychologii používala velmi záhy a i dnes se s ní pracuje poměrně často. Zpracování získaných dat však bývá většinou značně zjednodušené a nevyčerpávají se všechny možnosti detailní analýzy získaných údajů; tím se ztrácí část informací, které jsou v pořadových datech obsaženy.

IV.2. Vytvoření škály z pořadových dat

Úkolem posuzovatele je seřadit do pořadí určité množství podnětů, posuzovaných podle určitého kritéria. Přitom se nedovoluje zařadit dva podněty jako stejné; zároveň nelze žádnou pořadovou hodnotu přeskočit, i když třeba mezi dvěma jevy sousedícími v pořadí je značná mezera. Podnět S_j je přitom tedy lepší /větší, užitečnější.../ než podněty S_k, S_l ..., které jsou uvedeny v pořadí za ním, a naopak je horší než podněty S_a, S_b ... S_i, které jsou v pořadí před ním.

Při sestavování pořadí má posuzovatel obvykle všechny posuzované podněty simultánně předloženy /ať již ve skutečnosti nebo alespoň v nějakém symbolickém vyjádření/. Může s nimi různě manipulovat a měnit či upravovat svá rozhodnutí, dokud nesestaví definitivní pořadí. Při předkládání podnětů je výhodnější dávat je posuzovatelům na kartách; na každé z nich je napsáno jméno jediného z nich. Posuzovatel si pak může karty libovolně rozkládat a srovnávat. Pokud to není možné a seznam posuzovaných podnětů je předkládán na jednom listě, pak je třeba zvážit, zda pořadí, ve kterém jsou napsány, neovlivní úsudek posuzovatelů. Může totiž dojít k řadě chyb. Předložené pořadí působí často sugestivně; jevy, které jsou na začátku seznamu, resp. na jeho konci, bývají ostřeji a detailněji posuzovány, než ty, které leží uprostřed seznamu. Bývá proto užitečné použít několik seznamů, ve kterých je pořadí posuzovaných jevů systematicky měněno, a různým posuzovatelům /případně podskupinám posuzovatelů/ se předkládají podněty v pozměněném sledu.

Získané výsledky jsou teprve zaznamenány v tabulce, ve které sloupce reprezentují posuzované podněty a řádky odpovídají jednotlivým posuzovatelům. V každém poli tabulky je tedy zachyceno pořadí, které přiřadil určitý posuzovatel k určitému podnětu. Předpokládejme následující příklad: 25 posuzovatelů posuzuje celkem 15 literárních děl a sestaví je do pořadí podle toho, jak na ně tato díla esteticky působí. Přitom číslem 1 je označeno dílo, které je hodnoceno nejvýše, číslem 15 dílo hodnocené nejníže. Data jsou shrnuta v tabulce IV.1. str.81

V této podobě jsou však výsledky nepřehledné a tato obtíž nezbytně vzrůstá při větším počtu posuzovatelů /často i 100 a více osob/. Proto se data transformují do nové matice, ve které zachycujeme frekvenci výskytu jednotlivých podnětů na různém místě v pořadí / f_{ij} /. Takto jsou výsledky vyjádřeny v tab. IV.2. Do této tabulky zavádím i další údaje, potřebné pro sestavení intervalové škály.

Vlastní postup při výpočtu škály má několik forem. Pokud se spokojíme se stanovením celkové škály, pak sečteme pořadové hodnoty, které byly přiřazeny jednotlivým podnětům. Výsledné součty se pak seřadí do pořadí. Přiměřenější však je vyjádřit pořadové hodnoty, přiřazené jednotlivým podnětům, jejich mediánem a z nich vycházet při sestavení celkové škály /vzhledem k tomu, že máme data pořadová, může být výpočet průměrných hodnot někdy sporný/. Při použití mediánů však vznikají určité obtíže při větším počtu posuzovatelů. Obvyklý výpočet přesné hodnoty je věcně sporný, protože musíme většinou interpolovat /na jednotlivých místech v pořadí bývá často více hodnot/ a dostáváme se k číslům, která svým řádem neodpovídají původním datům. Medián, který je vyjádřen desetinnými čísly, není přiměřený, protože výchozí pořadové hodnoty jsou vyjadřovány v celých číslech.

Jestliže jsou k dispozici data, získaná od dostatečného počtu posuzovatelů, pak je možno vytvořit škálu intervalovou. Postupuje se dvěma způsoby, které se svými výsledky od sebe příliš neliší. První způsob se opírá o předpoklad normálního rozložení získaných dat, druhý způsob se opírá o srovnávání jednotlivých posudků a blíží se technice párového srovnávání. Oba postupy budeme demonstrovat na stejných datech, která jsme uvedli výše /s vědomím že počet 25 posuzovatelů není příliš velký a ve skutečných výzkumech by měl být vyšší/.

a/ Postup při předpokladu normálního rozložení dat

Jestliže seřadíme určitý počet prvků do pořadí, vytvořili jsme tím stupnici, ve které neznáme intervaly mezi sousedními členy řady. Můžeme však předpokládat, že řazené jevy jsou vybrány z populace s normální distribucí, a proto budou intervaly uprostřed škály menší a naopak budou větší v extrémních oblastech. To můžeme odvodit i z přímé zkušenosti. Rozdíly mezi prvním, druhým či třetím členem řady obvykle určíme snáze než rozdíly mezi prostředními členy, kde obvykle více váháme a nebýváme si příliš jisti, zda jsme jednotlivé podněty zařadili správně.

Tab. IV.1.

Pořadové hodnoty přisouzené 15 literárním dílům 25 různými posuzovateli

Posuzovatelé:	Literární díla:														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1.	10	15	13	7	5	1	6	12	11	3	9	8	4	2	14
2.	11	8	6	1	4	2	5	7	15	3	13	14	10	12	9
3.	6	7	11	2	4	5	14	10	8	1	12	15	13	3	9
4.	10	15	13	8	1	3	5	12	14	2	9	6	7	4	11
5.	9	10	14	6	3	1	8	15	13	5	4	7	11	2	12
6.	12	14	15	1	7	2	6	11	9	3	8	13	5	4	10
7.	9	7	10	8	4	1	15	11	14	2	3	5	13	6	12
8.	4	8	10	2	7	1	12	15	13	3	6	14	9	5	11
9.	4	9	14	1	8	3	12	15	10	6	7	13	5	2	11
10.	4	9	15	2	7	1	8	14	11	5	10	12	6	3	13
11.	6	8	11	3	5	4	12	15	14	2	10	9	7	1	13
12.	5	13	12	3	6	1	8	14	11	2	7	10	9	4	15
13.	4	11	10	3	9	2	7	15	14	5	8	13	6	1	12
14.	6	12	13	2	5	1	8	15	11	3	7	10	9	4	14
15.	4	14	11	1	5	2	8	15	13	6	9	10	7	3	12
16.	1	9	14	3	5	2	8	15	12	7	10	11	6	4	13
17.	3	10	13	5	4	1	9	14	11	2	7	12	8	6	15
18.	6	11	14	2	5	1	9	13	12	3	8	10	7	4	15
19.	4	14	11	3	7	2	10	12	13	5	8	9	6	1	15
20.	11	7	15	2	4	1	5	13	9	3	10	14	6	8	12
21.	4	12	13	1	5	2	9	15	11	6	7	10	8	3	14
22.	4	10	12	2	5	1	8	14	11	3	6	15	7	9	13

Tab. IV.1.

Pořadové hodnoty přisouzené 15 literárním dílům 25 různými posuzovateli /pokračování/

Posuzovatelé:	Literární díla:														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
23.	4	7	12	6	1	2	9	15	10	5	11	13	8	3	14
24.	4	9	13	2	7	1	11	14	10	5	8	12	6	3	15
25.	5	11	14	2	7	1	8	15	12	3	9	10	6	4	13

Z tohoto důvodu není na místě pracovat s jednotlivými pořadími jako s ekvidistantními jednotkami, ale je vhodnější aproximovat normální rozložení pořadových posudků. Guilford /1954/ navrhl užívat C-škálu, která je nezávislá na počtu řazených podnětů. Tato škála má průměr 5, standardní odchylku 2 a devět standardizovaných stupňů. Je tabelována vždy v celých číslech, což snad může znamenat určitou chybu, ale při řadě nepřesností, které se při řazení podnětů vyskytují, ji můžeme považovat za zanedbatelnou. C-škálu rozšířili Bartlett a Edgerton /1966/. Jejich upravenou tabulku uvádíme v příloze jako tab. C. V našem příkladě jsme nechali posuzovatele řadit 15 literárních děl; vyhledáme tedy patřičné hodnoty C-škály v řádku označeném číslicí 15. Zjistíme, že podnět, který má pořadovou hodnotu 15, získá 1 bod; podněty, které mají pořadí

10 - 12, získají 4 body; podněty s pořadím 4 - 6 mají 6 bodů atd. Hodnoty C-škály jsou zavedeny v tab. IV.2. ve sloupci označeném C_i.

Výpočet intervalové škály je nyní již jednoduchý. Frekvenci posudků v jednotlivých polích tabulky / f_{ij} / násobíme jim odpovídající hodnotou C_i. Data v jednotlivých sloupcích sečteme. Výsledky jsou zachyceny v řádce $f_{ij}C_i$. Vypočteme-li nyní průměr /C-škálu považujeme za intervalovou, a proto jsme k tomu oprávněni/, získáme škálové hodnoty, které jsou zachyceny v řádku M_c. Nejvýše tedy bylo ceněno dílo F., na druhém místě D, na třetím J ... , nejméně pak dílo H, které je na posledním místě v pořadí.

Kontrolu vlastního výpočtu umožňuje skutečnost, že součet hodnot v řádku $f_{ij}C_i$ je roven hodnotě 5Nn /kde N je počet posuzovatelů a n je počet řazených jednotek; hodnota 5 je průměr standardizované škály C/. Součet hodnot v řádku M_c se rovná hodnotě 5n /s možnou malou chybou, která vznikne při zaokrouhlování průměrů/.

Podíváme-lise však nyní podrobněji na získanou škálu, která je znázorněna na obrázku IV.1., vidíme zde určitý přeryv mezi 9. a 10. místem v pořadí. Je možné, že díla, která byla posuzována, tvoří dvě podskupiny, mezi kterými je relativně větší vzdálenost. Tím je poněkud zpochybněn předpoklad normality rozložení pořadí, a je proto na místě - nejen z důvodů didaktických - vyzkoušet i druhý způsob konstrukce intervalové škály.

b/ Postup opírající se o srovnávání posudků

Tato metoda vychází z toho, že z pořadových hodnot můžeme odvodit vztahy mezi všemi páry řazených jevů. Mějme 4 jevy: P, R, S, T, které byly seřazeny v uvedeném pořadí. Z toho můžeme vyvodit

6 párových vztahů: P > R , P > S , P > T , R > S , R > T , S > T . Jestliže jsou tyto jevy posuzovány vícekrát /ať již různými posuzovateli, anebo jedním posuzovatelem opakovaně/, je možno stanovit, jak často je jev i posuzován výše než jev j a naopak. Můžeme postupovat tak, že v tab. IV.1. srovnáváme párově jednotlivá díla, která jsou zachycena ve sloupcích, a zjistíme, jak často bylo třeba dílo A řazeno výše než B a naopak; dále srovnáváme A s C atd. Tento postup by byl při velkém počtu posuzovatelů, resp. při velkém počtu posuzovaných jevů, příliš zdlouhavý a těžko proveditelný.

Guilford /1954, str. 183-188/ navrhl jiný postup pro vytvoření škály; jeho zdůvodnění zde neuvádím a odkazuji na původní pramen. Máme řadu posuzovaných jevů S_a, S_b, ... S_j, ... S_n, kterým jsou přiřazeny pořadové hodnoty R_1, R_2, ... R_i, ... R_n. Kdykoliv je jevu S_j přiřazena hodnota R_i, považujeme jej za větší než všechny jevy s nižší pořadovou hodnotou. Každý jev je předložen N-krát; sečteme-li všechny případy, ve kterých S_j bylo hodnoceno výše než kterékoliv jiné S, získáme hodnotu $C_{j>CS} = \sum f_{ij}R_i - 0{,}5N$. Hodnoty R_i jsou řazeny opačně než hodnoty r_i uvedené výše v tabulce /tedy $R_i = R_n - r_i + 1$ /. Hodnota 0,5N vzniká tím, že jev S_j může být též srovnáván sám se sebou /analogicky jako tomu bylo v párovém srovnávání v tab. III.4./. Celkový počet srovnání pro každé S je Nn. /V našem příkladě tedy 25 . 15 = 375/. Chceme-li vyjádřit relativní četnost toho, že určitý jev S_j byl hodnocen výše než ostatní, použijeme vztah:

$$p_{j>CS} = \frac{\sum f_{ij}R_i - 0{,}5N}{Nn}$$

Tab. IV.2. Frekvence posudků 15 literárních děl, seřazených pořadově 25 posuzovateli f_{ij}

Pořadí r_i	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	$\sum$	P	C_i
1	1			5	2	13				1				3		25	96,7	9
2				9		8				5				3		25	90,0	8
3	1			5	1	2				9	1			6		25	83,3	7
4	10				5	1					1		1	7		25	76,7	6
5	2			1	8	1	3			6		1	2	1		25	70,7	6
6	4		1	2	1		2			3	2	1	7	2		25	63,3	6
7		4		1	6		1	1		1	5	1	5			25	56,7	5
8		3		2	1		8		1		5	1	3	1		25	50,0	5
9	2	4			1		4		2		4	2	3	1	2	25	43,3	5
10	2	3	3				1	1	3		4	6	1		1	25	36,7	4
11	2	3	4				1	2	7		1	1	1		3	25	30,0	4
12	1	2	3				3	3	3		1	3		1	5	25	23,3	4
13		1	6					2	4		1	4	2		5	25	16,7	3
14		3	5				1	5	4			3			4	25	10,0	2
15		2	3				1	11	1			2			5	25	3,3	1
$\sum f_{ij}$	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	375		

Tab. IV.2. Frekvence posudků 15 literárních děl, seřazených pořadově 25 posuzovateli – pokračování

	Literární díla:															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
$\sum f_{ij} C_i$	142	98	77	185	149	207	118	56	88	171	122	92	129	167	74	1875
M_C	5,7	3,9	3,1	7,4	6,0	8,3	4,7	2,2	3,5	6,8	4,9	3,7	5,2	6,7	3,0	75,1
Pořadí	6	10	13	2	5	1	9	15	12	3	8.	11	7	4	14	
$\sum f_{ij} R_I$	250	140	91	322	270	356	180	64	108	307	194	125	211	299	83	$\sum$3000
$N/2$	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	12,5	187,5
$\sum f_{ij} R_i - N/2$	237,5	127,5	78,5	309,5	257,5	343,5	167,5	51,5	95,5	294,5	181,5	112,5	198,5	286,5	70,5	2812,5
$p_{j>CS} = \dfrac{\sum f_{ij} R_i - N/2}{N_n}$	0,633	0,340	0,209	0,824	0,687	0,916	0,447	0,137	0,253	0,785	0,484	0,300	0,529	0,764	0,188	7,496
z	0,340	-0,412	-0,810	0,930	0,488	1,379	-0,133	-1,093	-0,665	0,789	-0,040	-0,524	0,074	0,720	-0,882	
$H = z + 1{,}093$	1,433	0,681	0,283	2,023	1,581	2,472	0,960	0,000	0,428	1,882	1,053	0,569	1,167	1,813	0,211	
Pořadí	6	10	13	2	5	1	9	15	12	3	8	11	7	4	14	

Výpočet těchto hodnot je zachycen ve spodní části tab. IV.2.

Hodnoty v řádku $p_{j>CS}$ transformujeme na hodnoty z /viz opět tab. A v příloze/ a konečnou škálu získáme zakotvením nejnižšího bodu škály v počátku. Docílíme to tím, když ke každé vypočtené škálové hodnotě přičteme konstantu +1,093, která odpovídá nejnižší zjištěné hodnotě z . Konečná forma škály je zachycena v předposlední řádce tab. IV.2. /H/.

Tak jako při předchozí metodě i při tomto výpočtu máme možnost kontroly. Součet hodnot v řádku $f_{ij}R_i$ = N R_i = 3000, kde N je počet posuzovatelů a R_i je součet pořadových hodnot posuzovaných jevů / 1 + 2 + 3 ... + 14 + 15/. V řádku 0,5Nn musí být součet 187,5; další řádek je rozdíl obou předchozích, čili 3000 - 187,5=2812,5. Konečně v dalším řádku dostaneme tuto hodnotu dělenou Nn, což musí být 7,5 /opět s možnou malou chybou při zaokrouhlování/.

Srovnáme-li obě popsané metody, vidíme, že pořadí posuzovaných jevů zůstalo stejné, a obě metody vedly tedy k prakticky stejnému výsledku. I při druhém postupu zůstala opět vzdálenost mezi 9. a 10. bodem škály relativně velká, což potvrzuje domněnku, že původní soubor 15 podnětů je pravděpodobně tvořen dvěma podskupinami.

IV.3. Neúplné pořadové škály

Někdy dochází k tomu, že z nejrůznejších důvodů nechceme, či ani nemůžeme pracovat s kompletním pořadovým uspořádáním všech možných podnětů. Tak např. při poradenství ve volbě povolání se někdy z velkého množství profesí žádá, aby pokusná osoba označila podskupinu profesí, které by se jí nejvíce líbily, a tyto seřadila do pořadí; nebo je předepsáno z mnoha osob určit jen několik, které jsou sympatické atd. Takový postup je výhodný, když počet posuzovaných jevů je příliš rozsáhlý nebo když předpokládáme, že posuzovatelé by je obtížně srovnávali a sestavovali by je do pořadí se značnou nepřesností. Neúplnou pořadovou škálu užíváme i tehdy, když očekáváme, že sestavování všech podnětů do pořadí by bylo pro posuzovatele značně nezajímavé či únavné. I za těchto podmínek je možno sestavit intervalovou škálu; je však třeba jinak postupovat. V extrémním případě - při dostatečně velkém počtu posuzovatelů - lze dokonce volit i takový postup, při kterém se určuje pouze ten podnět, který patří na první místo v pořadí.

Nejprve ilustrujeme postup při vytváření škály, máme-li stanoveno pořadí prvních k-podnětů z celkového počtu n-podnětů. Použijeme opět metodu srovnávání posudků. Předpokládejme, že se opět posuzuje 15 podnětů; k dispozici jsou data od 100 posuzovatelů. Každý z nich měl seřadit prvních 5 podnětů, které považuje za nejlepší /výchozí data jsou převzata z Guilfordovy učebnice/. Výsledky nejprve shrneme do tab. IV.3.

Jde nám nyní o zjištění, jak často je volen na prvních 5 místech určitý podnět S_i ve srovnání s ostatními podněty. Počet jeho voleb označíme C_i, počet všech voleb označíme T. Platí, že $T = k \cdot N$, kde k je počet vybíraných podnětů a N je počet posuzovatelů. V našem příkladě $T = 5 \cdot 100 = 500$. Poměr, ve kterém je určitý podnět hodnocen výše než všechny ostatní podněty souhrnně, je dán vztahem

$$p_{j>CS} = \frac{c_j}{c_j + \frac{T}{n}}$$

kde $c_j = \sum_{i=k}^{n} f_{ij}R_i - 0{,}5N_k$

$N_k = \sum f_{ij}$

n = počet řazených jevů /15/

Výpočet hodnoty / $p_{j>CS}$ / je zachycen v řádcích 6 - 9 tab. IV.3. Vypočtené hodnoty se opět převedou na hodnoty z a sestaví se konečná škála /řádek H/, jejíž nejnižší bod je opět zakotven v počátku /odečtením konstanty 0,838/.

Guilford /1954/ upozorňuje, že v některých případech může být určitý rozdíl mezi škálami, které se vytvářejí z úplných nebo neúplných pořadových dat. Tam, kde můžeme získat úplná pořadová data, je třeba tomuto postupu dát přednost.

IV.4. M e t o d a j e d i n é h o v ý b ě r u

Tato metoda byla užívána již Fechnerem v klasické psychofyzice a označována pojmem "Wahlmethode". Jedná se vlastně o extrémní případ předchozí techniky při $k = 1$. Posuzovatel uvádí pouze jediný podnět, a to ten, který řadí na první místo. Je použitelná jen v takových případech, kdy počet posuzovaných jevů je malý a počet posu-

Tab.IV.3. Frekvence výběru /f_{ij}/ 15 herců při určení pěti nejoblíbenějších

R_i	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
15	9	17	15	8	10	4	4	1	3	1	0	9	2	2	15	
14	7	11	15	8	8	3	6	6	4	2	1	8	1	8	12	
13	18	9	11	8	5	2	3	6	6	3	2	5	5	4	13	
12	11	11	16	4	8	7	2	5	5	1	4	6	8	7	5	
11	9	8	9	6	9	9	6	3	6	4	5	3	7	10	6	
$f_{ij}R_i$	698	746	869	450	522	311	273	270	305	138	143	417	282	388	688	6500
$0,5N_K$	27	28	33	17	20	12,5	10,5	10,5	12	5,5	6	15,5	11,5	15,5	25,5	250
C_j	671	718	836	433	502	298,5	262,5	259,5	293	132,5	137	401,5	270,5	372,5	662,5	6250
$P_{j\ CS}$	0,953	0,956	0,962	0,929	0,938	0,899	0,887	0,886	0,898	0,799	0,804	0,923	0,890	0,918	0,952	
z_j	1,675	1,706	1,774	1,469	1,579	1,281	1,210	1,205	1,270	0,838	0,843	1,421	1,222	1,392	1,665	
H	0,837	0,868	0,936	0,631	0,741	0,443	0,372	0,367	0,432	0,000	0,005	0,583	0,384	0,554	0,827	

zovatelů značně velký. Uvádí se, že poměr N/n by měl být větší než 50 /tedy při 10 posuzovaných podnětech alespoň 500 posuzovatelů/. Z toho plyne, že metoda jediného výběru je vhodná při výzkumech spíše sociologického charakteru či při výzkumech veřejného mínění apod., kdy se zjišťuje, kterému druhu zboží se dává přednost, jaká forma dovolené se považuje za nejlepší a j. Většinou se předkládá několik možných odpovědí a určuje se nejlepší z nich. Celkově se však jedná o metodu, která není příliš přesná a poskytuje poměrně málo informací.

IV.5. Metoda pořadí v podnětových triádách

Jakýmsi kompromisem mezi metodou neúplných pořadových škál a metodou jediného výběru je sestavování pořadí ze tří podnětů. Postupuje se tak, že jsou předloženy vždy tři podněty a posuzovatel je má pořadově seřadit. Postup se blíží metodě párového srovnávání, protože sestavené pořadí v trojici podnětů můžeme chápat jako srovnání tří podnětových párů. Většinou se používají všechny možné triády. Při n-podnětech je jejich počet dán vztahem $n(n-1)(n-2)/6$.

Získaná data se většinou zpracovávají analogicky jako při metodě párového srovnávání podnětů. Pokud se pracuje se všemi triádami, pak každá dvojice podnětů se vyskytne spolu v n-2 trojicích. Při slučování dat je třeba mít na zřeteli, že se často směšují data získaná od několika posuzovatelů s daty od téhož posuzovatele. Je ovšem možno sestavit podnětové triády tak, aby se každá dvojice podnětů vyskytla spolu jen jedenkrát /případně u různých posuzovatelů v různých kombinacích s ostatními podněty/.

Při úplném počtu triád jejich počet roste při zvětšujícím se počtu podnětů ještě rychleji než počet párů při použití metody párového srovnávání. Právě tato potíž vede k tomu, že metoda podnětových triád se používá jen ojediněle.

IV.6. Shoda mezi posuzovateli

Shodu mezi posuzovateli můžeme vyhodnotit pomocí Kendallova koeficientu shody W. Popsali jsme jej při jiné příležitosti /Břicháček, Hampejsová, 1963/ a najde se v běžných učebnicích neparametrických statistických technik. V příkladě uvedeném v tab. IV.1. po dosazení do vzorce pro výpočet koeficientu shody W vychází hodnota

$$W = \frac{\sum d^2}{\frac{1}{12} k^2 N(N^2 - 1)} = \frac{130\ 182}{\frac{1}{12} . 25^2 . 15 . 224} = 0{,}744$$

Významnost testujeme přes χ^2 -rozložení dle vzorce

$$\chi^2 = k(N-1)\ W = 25 . 14 . 0{,}744 = 260$$

což při $N - 1 = 14$ stupních volnosti je hodnota statisticky vysoce významná.

Tento nález svědčí o tom, že skupina 25 posuzovatelů seřadila srovnávaná díla zhruba stejně a zjištěné rozdíly se pohybují v rámci náhodného kolísání.

IV.7. S o u h r n

Celkově je možno říci, že metoda pořadových stupnic má řadu výhod, a to hlavně z hlediska posuzovatelů. Úkol je pro ně snadnější a zajímavější, než tomu bývá u metody párového srovnávání. Do jisté míry se vyhneme i různým nepravidelnostem, které se objevují při párovém srovnávání /kruhové triády/; posuzovatel, který má před sebou současně všechny podněty, třídí je pravděpodobně stále podle stejného kritéria, zatímco při párovém srovnávání může dojít k nejednotnému postupu. Podstatnou výhodou je i to, že můžeme pomocí metody pořadových stupnic srovnávat a škálovat i poměrně velký počet podnětů.

Pro použití metody pořadových stupnic platí zhruba tato pravidla:

a/ při malém počtu posuzovatelů se spokojíme s celkovým pořadím určeným buď součtem pořadových hodnot, které jsou přiřazeny jednotlivým podnětům, anebo z nich vypočteným mediánem, a nepokoušíme se sestavit intervalovou škálu;

b/ tam, kde můžeme očekávat normální rozložení používaných podnětů, vytváříme intervalovou škálu pomocí metody normalizovaného pořadí. Počet posuzovatelů není předepsán; metodu proto lze užít i při malém počtu posuzovatelů. Naproti tomu počet řazených podnětů nesmí být malý /předpoklad normality rozložení bychom nemohli ověřit/. Normální rozložení můžeme očekávat tam, kde vybíráme náhodně prvky z populace, která je normálně rozložena a není ničím zkreslena;

c/ metoda, založená na srovnávacích posudcích, je relativně nejvýhodnější, ale též nejpracnější. Rozložení posuzovaných podnětů není rozhodující. Obecně platí, že čím větší je počet posuzovatelů, tím přesnější škálu můžeme získat.

Ve všech případech je třeba uvažovat i o tom, zda zvolení posuzovatelé tvoří reprezentativní výběr, a zda tedy můžeme získané škály zobecňovat. Často dochází k chybě, jestliže škála je sestavena na základě posudku odborníků; jejich výsledky se nemusí shodovat s posouzením, které by prováděli laici.

Kapitola V.

TECHNIKA ZDÁNLIVĚ STEJNÝCH INTERVALŮ

V.1. Ú v o d

Techniku zdánlivě stejných intervalů zavedl v r. 1929 Thurstone /ve spolupráci s Chavem/ a dodnes patří k metodám velmi často užívaným v sociální psychologii, a to hlavně při výzkumu postojů. Detailní rozpracování této techniky je spojeno s monografií Edwardse /1957/. Literatura, která se zabývá touto metodou, je ze všech škálovacích metod pravděpodobně nejrozsáhlejší. Příkladem využití této techniky v naší literatuře je studie Váchové a Šmolky /1975/.

Metoda zdánlivě stejných intervalů byla vyvinuta pro situace, kdy v důsledku velkého množství podnětů /často i více než 50/ nelze použít metodu párového srovnávání, a kdy i metoda pořadových stupnic je těžko realizovatelná, ať již pro malé rozdíly mezi posuzovanými jevy, anebo pro nebezpečí, že sledované jevy je možno posuzovat z mnoha různých hledisek. Technika zdánlivě stejných intervalů je velmi náročná a pracná z hlediska experimentátora. Z hlediska pokusné osoby je však poměrně jednoduchá a vyžaduje pouze jeden úkon. Ve své struktuře je tato metoda jakýmsi opakem metody pořadového řazení. Pořadí mezi různými podnety je již na základě náročných předpokusů sestaveno a úkolem pokusné osoby je pouze určit, které podněty, resp. výroky nejlépe vyjadřují její vlastní názor, postoj či zaměření.

V.2. Sestavení škály

Při výkladu techniky zdánlivě stejných intervalů se soustředím hlavně na rozbor jednotlivých kroků, potřebných pro vytvoření škály. Existuje několik variant postupu; v tomto výkladu se přidržím klasického způsobu, který byl vyvinut pro vytváření postojových škál.

1. krok: Shromáždění výroků

Nejprve je větší počet osob /často i několik set/ požádán, aby vyjádřil svůj názor o té skutečnosti, pro kterou sestavujeme škálu. Shromážděné výroky je možno doplnit dalšími výroky známými z literatury či z jiných zdrojů. V této fázi je z hlediska metodologického třeba dbát na to, aby výběr dotazovaných osob byl reprezentativní. Kdybychom získali a analyzovali výroky skupiny vysokoškolských studentů, mohlo by se stát, že škála, kterou sestavíme, nebude přiměřená pro osoby starší či pro osoby z jiné sociální skupiny. Výroky, vyjadřující mínění o sportu, by asi nebylo správné shromažďovat bezprostředně po olympijských hrách, kdy řada lidí denně sleduje dobře režírované televizní přenosy; tato aktuální zkušenost může dočasně ovlivnit a pozměnit jejich názory a postoje.

2. krok: Utřídění výroků skupinou expertů

V další fázi se pracuje se skupinou expertů, kteří posuzují jednotlivé výroky a třídí je do určitého počtu kategorií od výroků výrazně pozitivních až po výroky výrazně negativní. Počet expertů by měl být dostatečně velký; často se uvádí, že jich má být alespoň 50. Kategorií, do kterých se výroky třídí, by měl být lichý počet, aby bylo možno zatřídit i výroky neutrální. Počet kategorií bývá většinou 7 - 11. Od expertů se žádá, aby třídili výroky tak, aby "vzdálenosti" mezi jednotlivými kategoriemi byly stále stejné /odtud pochází název metody - zdánlivě stejné intervaly/. Dále se předpokládá, že užité kategorie jsou umístěny na jediném kontinuu, to jest, že vytvářejí jednodimenzionální stupnici. Při třídění výroků se očekává, že experti jsou schopni odhlédnout od svých vlastních postojů či názorů a vycházejí pouze z obsahu předložených výroků. V instrukci se od nich též žádá, aby se snažili využívat všechny kategorie /i když ve většině případů extrémní kategorie pozitivní i negativní nebývají užívány tak často jako ostatní/.

Předpokládejme, že jsme zvolili škálu s devíti body, která má tuto strukturu:

1	2	3	4	5	6	7	8	9
příznivé a pozitivní výroky			neutrální výroky			nepříznivé a negativní výroky		

3. krok: Výpočet škálových hodnot pro jednotlivé výroky

Údaje získané od expertů se shrnou do tabulek a vypočte se škálová hodnota pro jednotlivé výroky. Jako míra střední tendence se často používá průměr, ale vhodnější je medián, který není ovlivňován extrémními hodnotami či asymetrickým rozložením dat.

Uveďme si hypotetické příklady posouzení 2 výroků, které posuzovalo 200 expertů. Jejich výsledky jsou shrnuty v tab. V.1.

Tab. V.1.
Posouzení 2 výroků (X , Y) podle údajů 200 expertů

	Kategorie								
	1	2	3	4	5	6	7	8	9
Výrok X									
f	18	44	70	46	20	2	-	-	-
p	0,09	0,22	0,35	0,23	0,10	0,01			
cp	0,09	0,31	0,66	0,89	0,99	1,00			
Výrok Y									
f	3	17	32	47	39	23	18	13	8
p	0,015	0,085	0,160	0,235	0,195	0,115	0,090	0,065	0,040
cp	0,015	0,100	0,260	0,495	0,690	0,805	0,895	0,960	1,000

V této tabulce jsou v řádcích označených písmenem f uvedeny počty osob, které určitý výrok zatřídili do dané kategorie /výrok X umístilo do kategorie č. 1 celkem 18 expertů, výrok Y pouze 3 experti atd./. V řádcích označených p se uvádí relativní frekvence, s jakou je daný výrok zařazen do některé z 9 kategorií. Tyto hodnoty jsou v řádcích cp kumulativně sečteny.

Výpočet mediánu se běžně provádí podle vzorce:

$$Me = u + \frac{0{,}50 - \sum p_u}{p_i}$$

kde u = spodní hranice intervalu, ve kterém leží medián,

$\sum p_u$ = relativní frekvence hodnot, které leží pod intervalem, ve kterém je medián,

p_i = relativní frekvence hodnot v intervalu, ve kterém leží medián.

Mediánovou hodnotu pro výrok X vypočteme takto: medián leží zřejmě ve 3. kategorii; proto $u = 2$. Do kategorií 1 a 2 bylo zařazeno 31% výroků = $\sum p_u$. Ve 3. kategorii leží celkem 35% výroků = = p_i. Tedy

$$Me_X = 2 + \frac{0{,}50 - 0{,}31}{0{,}35} = 2{,}543$$

Mediánovou hodnotu výroku Y vypočteme analogicky:

$$Me_Y = 4 + \frac{0{,}50 - 0{,}495}{0{,}195} = 4{,}026$$

Takto se vypočítají hodnoty mediánu pro všechny výroky. V případě potřeby je možno srovnávat podskupiny expertů navzájem /např. skupinu mužů a žen/, případně je náhodně rozdělit a ze srovnání jejich výsledků soudit na reliabilitu získaných výsledků.

4. krok: Výběr vhodných položek pro definitivní škálu

Nejprve se vylučují ty výroky, o kterých se mínění expertů rozcházejí; projevuje se to velkou variabilitou jejich posudků. Jedná se zřejmě o výroky, které jsou víceznačné, a různí posuzovatelé je různě interpretují. Z uvedených příkladů sem zřejmě patří výrok Y, který je řadou expertů chápán jako pozitivní a jinými jako negativní. Variabilita se většinou vyjadřuje kvartilovým rozpětím Q. Numericky postupujeme stejně jako při zjišťování mediánu; rozdíl je pouze v tom, že nehledáme hodnotu případu, který dělí všechny data na polovinu, ale stanovíme hodnotu, která odděluje první čtvrtinu dat od druhé čtvrtiny /resp. třetí čtvrtinu od čtvrté/. Zjištěné hodnoty značíme Q_1 a Q_3 a odečteme je od sebe; čím vyšší číslo získáme, tím větší je kvartilové rozpětí.

V našem případě pro výrok X platí:

$$Q_1 = 1 + \frac{0,25 - 0,09}{0,22} = 1,727$$

$$Q_3 = 3 + \frac{0,75 - 0,66}{0,23} = 3,391$$

$$Q = 3,391 - 1,727 = 1,664$$

Pro výrok Y platí:

$$Q_1 = 2 + \frac{0,25 - 0,10}{0,16} = 2,937$$

$$Q_3 = 5 + \frac{0,75 - 0,69}{0,115} = 5,522$$

$$Q = 5,522 - 2,937 = 2,585$$

Po vyloučení výroků s vysokou hodnotou Q vyhledáme takové výroky, jejichž mediánové hodnoty leží rovnoměrně podél celé škály v přibližně stejných intervalech. Tím se opět vyloučí řada výroků, které do tohoto systému nezapadají. Pokud máme pro určitý bod škály k dispozici několik výroků, pak je buď necháme všechny a užíváme je jako příklady, anebo volíme ten z nich, u kterého jsme zjistili nejnižší hodnotu Q .

5. krok: Definitivní forma škály, která se předkládá pokusným osobám

Sestavenou škálu je možno předkládat v dvojí podobě:

a/ Sestaví se škála, která má 5 - 9 bodů se stejnými intervaly mezi sebou a každý z nich je popsán tím výrokem, který mu odpovídá. Instrukce pak zní: "Jestliže váš postoj nejlépe vyjadřuje výrok č. 1, pak zaškrtněte číslo 1; jestliže váš postoj nejlépe vyjadřuje výrok č. 2, pak zaškrtněte číslo 2 ..." atd. K jednotlivým bodům škály je možno uvést i několik dalších příkladů, které mají zhruba stejnou škálovou hodnotu. Tento způsob předkládání škály je velmi jednoduchý a z hlediska pokusné osoby zcela nenáročný a rychlý.

b/ Druhý způsob prezentace je poněkud složitější, ale je výhodnější při detailním studiu postojů. Zvolí se větší počet výroků /třeba 20 - 40 / a předkládají se v náhodném pořadí. Pokusná osoba určuje, s kterým tvrzením souhlasí a s kterým nikoliv. Protože známe škálovou hodnotu každého výroku, je možno vyjádřit /většinou opět mediánem/ celkový postoj každé osoby, případně též průměrný postoj skupiny pokusných osob. Získané hodnoty pak můžeme srovnávat jak mezi jednotlivými osobami, tak mezi skupinami, můžeme sledovat jejich změny v čase apod.

V.3. M e t o d o l o g i c k é p r o b l é m y v y t v á ř e - n í š k á l y

Při vytváření škály pomocí techniky zdánlivě stejných intervalů se objevují mnohá nebezpečí. Je nezbytné upozornit stručně alespoň na nejzávažnější z nich.

Velké úsilí bylo věnováno výzkumu, zda vlastní názor expertů ovlivňuje třídění jednotlivých výroků. Pokud by zde existoval nějaký vztah, pak by celá tato metoda byla velmi sporná. Ukázalo se však, že skupiny expertů s různými formami jednání, případně i s jinými, odlišnými charakteristikami, vytvořily škály, které byly téměř identické. Korelace mezi nimi byly vyšší než 0,90. Při detailnějších rozborech se však přece jen našly jisté vazby mezi postojem posuzovatele a jeho hodnocením řady výroků. Posuzovatelé, kteří mají výrazně extrémní osobní postoj k jevu, který je škálován,

a/ mají tendenci užít jen menší počet kategorií, které jsou k dispozici; vytvářejí tedy méně citlivou škálu;

b/ lépe diferencují ty výroky, které jsou ve směru jejich vlastního přesvědčení; výroky opačné diferencují mnohem méně;

c/ navzájem se méně shodují při posouzení výroků neutrálních než při posouzení výroků pozitivních či negativních.

Tyto poznatky vedou k závěru, že experti s extrémními postoji by neměli být žádáni o spolupráci.

Při použití metody zdánlivě stejných intervalů je třeba uvažovat i o interakci mezi experty a sledovaným souborem pokusných osob. Často se může stát, že experti tvoří relativně homogenní skupinu osob s určitým vzděláním a s určitým sociálním statusem, zatím co pokusné osoby mohou být z jiné sociální skupiny, s jinými hodnotovými systémy, a proto také s jinými postojovými strukturami. Sestavená škála pak není reprezentativní a dochází k chybám, které celý výzkum zkreslují či dokonce znehodnocují. Podobných chyb bychom v současné sociální psychologii nalezli hodně.

Je též zřejmé, že metodicky velmi pochybné je nekritické přebírání škál, sestavených v určitých společenských a kulturních podmínkách. Aplikace škály může v jiných společenských podmínkách selhat či vést ke zcela zkresleným výsledkům. Situace je o to obtížnější, že systematickému srovnávacímu výzkumu postojových škál v různých společenských a kulturních podmínkách nebylo zatím věnováno příliš mnoho pozornosti.

Další slabina metody zdánlivě stejných intervalů je v tom, že škála nebývá dostatečně citlivá. Posuzovatel, který musí rozdělit řadu výroků do určitého počtu kategorií, zařadí často do stejného intervalu výroky, které by dokázal diferencovat /třeba při párovém srovnávání/. Tuto nesnáz není možno odstranit zvětšováním počtu kategorií, neboť třídění by pak bylo obtížnější a méně přehledné, případně zatížené jinými závažnými chybami. Jisté zlepšení v tomto směru přináší metoda následných intervalů, která bude probrána v VI. kapitole.

V.4. Zjednodušená forma techniky zdánlivě stejných intervalů

Technika zdánlivě stejných intervalů se používá i v jednodušší podobě při výzkumech psychofyzického typu. Pokusné osobě se předloží určité kontinuum a úkolem je rozdělit je do několika stejně vzdálených částí. Kontinuum může být ohraničeno dvěma body a osoba vyhledává či sestavuje takový podnět, který by ležel uprostřed mezi oběma body. Takto vzniklé intervaly znovu půlí atd. V obecnější podobě se postupuje takto: jsou dány dva standardní podněty i, j. Úkolem je rozdělit interval mezi nimi do n subjektivně stejných dílů.

Úloha může být zadána i v této podobě: jsou opět dány podněty i, j. Jde nyní o to určit hodnotu variabilního podnětu k tak, aby interval mezi i a j /d_{ij}/ byl subjektivně stejný jako interval mezi j a k /d_{jk}/. Při větším počtu podnětů se žádá, aby byly umístěny na kontinuu tak, aby vzdálenost mezi sousedními podněty zůstávala subjektivně stejná; tedy $d_{ij} = d_{jk} = d_{kl} = d_{lm} = \ldots$

Technika zdánlivě stejných intervalů může být použita i v následující podobě. Pokusné osobě se předloží /ať již současně či v určité posloupnosti/ řada podnětů s instrukcí roztřídit je do řady kategorií, mezi kterými by měly být subjektivně stejné intervaly. Jako podněty je možno používat například umělecké reprodukce, výsledky pracovních činností, sportovní výkony /krasobruslení/ apod. Jedná se vlastně o jistou analogii kroku č. 2, který byl popisován v předchozí části.

Uvedené úpravy techniky zdánlivě stejných intervalů se užívají méně často; nebývají ani příliš spolehlivé. Přesto však mají své oprávnění, a to hlavně při orientačních šetřeních, kdy si chceme v předpokusech ověřit, zda určité výzkumné záměry jsou nosné. Užívají se s úspěchem při intenzívních výzkumných projektech, ve kterých se záměrně pracuje s jednou osobou a sledují se její výsledky za různých podmínek a v různém čase.

Kapitola VI.

TECHNIKA NÁSLEDNÝCH INTERVALŮ

VI.1. Ú v o d

Tato technika odstraňuje některé nesnáze techniky zdánlivě stejných intervalů. Do jedné kategorie se často zahrnují výroky, které nejsou zcela identické a bylo by je třeba detailněji diferencovat. Mnohdy dochází k chybám u extrémních výroků. Jestliže posuzovatel označí určitý výrok za zcela extrémní, a pak je mu předložen výrok ještě extrémnější, nezbytně oba začlení do stejné kategorie. Pokud začne škálu opravovat, pak se rozkolísá celý systém, který si vytvořil a vznikají další nepřesnosti. Technika následných intervalů se těmto chybám vyhýbá.

Základním předpokladem metody následných intervalů je očekávání, že posudky různých podnětů jsou různě distribuovány; přitom jednotlivé distribuce mohou být i značně asymetrické. Škálová hodnota podnětů je určena centrální tendencí jejich distribucí. Touto hodnotou je pak definováno umístění podnětů na základním škálovém kontinuu. Neopíráme se o předpoklad stejných intervalů na škále, ale naopak se snažíme určit jejich přesné hranice, které nemusí být ekvidistantní. Vychází se v podstatě pouze ze subjektívní rovnosti intervalů, kterou určuje skupina expertů.

Postup při výpočtu škály má opět několik kroků, které postupně probereme s využitím příkladů, uváděných Edwardsem /1957, str. 124 a dále/.

VI. 2. V y t v o ř e n í   š k á l y

První dva kroky jsou podobné jako u předchozí metody. Použije se opět určitý počet výroků a skupina expertů je třídí do kategorií. V tab. VI.1. jsou jako příklad uvedená posouzení dvou výroků, které hodnotilo 200 posuzovatelů. Výsledky se shrnou do frekvenčních tabulek /řádek f /, četnosti se kumulativně sečtou /řádek cf /, a potom se převedou na kumulativní součty relativních frekvencí /řádek cp / tak, že hodnoty v řádku cf se dělí počtem posuzovatelů /N/.

Tab. VI.1.

Posouzení 2 výroků 200 posuzovateli

Výrok	Posuzovací kategorie Nepříznivé						Příznivé		
č. 1	1	2	3	4	5	6	7	8	9
f	2	4	12	12	26	52	60	26	6
cf	2	6	18	30	56	108	168	194	200
cp	0,01	0,03	0,09	0,15	0,28	0,54	0,84	0,97	1,00
č.14									
f	0	0	0	2	8	36	64	80	10
cf	0	0	0	2	10	46	110	190	200
cp	0,00	0,00	0,00	0,01	0,05	0,23	0,55	0,95	1,00

Z tab. VI.1. je již na první pohled zřejmé, že výrok č. 14 je posuzován příznivěji, než výrok č. 1. Distribuce posudků je v obou případech značně nesymetrická: u výroku č. 1 je variance posudků zřejmě větší než u výroku č. 14; je tedy výrok č. 14 chápán relativně jednoznačněji, než výrok č. 1. Nesymetrické distribuce posudků naznačují, že šíře jednotlivých kategorií nebude stejná a jejich hranice nebudou ekvidistantní.

Celkem bylo posuzováno 14 výroků; kumulativní proporce všech 14 výroků jsou shrnuty v tab. VI.2. Zcela stejným způsobem můžeme škálovat i jiné jevy - např. posuzovat různé osoby z hlediska určitých komplexních vlastností/pracovní zásluhy, ochota ke spolupráci, postoje k určité skutečnosti atd./, posuzovat estetickou hodnotu uměleckých děl, sportovní výkony, užitečnost zboží, přiměřenost chování v určité situaci atd.

Při dalším kroku hodnoty cp_{ij} /proporce posudků, které i-tý výrok zařazují pod horní hranici j-té kategorie/ převedeme na hodnoty z_{ij}. Přitom se vypouštějí hodnoty $p \leq 0{,}01$ a $p \geq 0{,}99$. Proto je v tab. VI.3. vynechán sloupec 9 a ve sloupci 8 hodnoty u 10. a 11. výroku; mizí také řada hodnot v 1. sloupci /a též některé v dalších sloupcích/. Hodnoty z_{ij} vyhledáváme v tab. A v příloze; postupujeme tak, jak bylo popsáno již v kapitole III. na str. 49 a dále.

Každou hodnotu z_{ij} můžeme chápat jako míru vzdálenosti mezi hraničním bodem jednotlivých kategorií a průměrnou hodnotou jednotlivých výroků. Průměrné hodnoty jednotlivých výroků kolísají a stejně tak kolísá i distribuce posudků; získáváme vlastně tolik různých škál, kolik výroků bylo posuzováno.

Škálovací procedura nyní vyžaduje, abychom stanovili šířku

Tab. VI.2.

Kumulativní proporce /cp_{ij}/ , zjištěné při posouzení 14 výroků /200 posuzovatelů/.

Výrok	Přiřazená kategorie 1	2	3	4	5	6	7	8	9
1	0,01	0,03	0,09	0,15	0,28	0,54	0,84	0,97	1,00
2				0,01	0,03	0,13	0,57	0,94	1,00
3	0,01	0,07	0,23	0,37	0,55	0,81	0,93	0,98	1,00
4	0,04	0,13	0,29	0,46	0,71	0,85	0,95	0,97	1,00
5	0,02	0,14	0,32	0,51	0,58	0,74	0,86	0,96	1,00
6			0,01	0,03	0,10	0,30	0,67	0,85	1,00
7		0,03	0,09	0,20	0,35	0,55	0,83	0,97	1,00
8		0,02	0,07	0,15	0,36	0,61	0,85	0,97	1,00
9	0,01	0,06	0,18	0,38	0,63	0,86	0,95	0,98	1,00
10	0,01	0,07	0,27	0,49	0,66	0,85	0,95	0,99	1,00
11	0,01	0,02	0,10	0,19	0,39	0,55	0,86	0,99	1,00
12		0,01	0,04	0,10	0,22	0,40	0,69	0,95	1,00
13			0,01	0,02	0,07	0,32	0,67	0,94	1,00
14				0,01	0,05	0,23	0,55	0,95	1,00

Tab. VI.3.

Hodnoty z_{ij} , odpovídající datům z tab. VI.2.

Výrok	Přiřazená kategorie 1	2	3	4	5	6	7	8
1		-1,881	-1,341	-1,036	-0,583	0,100	0,994	1,881
2					-1,881	-1,126	0,176	1,555
3		-1,476	-0,739	-0,332	0,126	0,878	1,476	2,054
4	-1,751	-1,126	-0,553	-0,100	0,553	1,036	1,645	1,881
5	-2,054	-1,080	-0,468	0,025	0,202	0,643	1,080	1,751
6				-1,881	-1,283	-0,524	0,440	1,036
7		-1,881	-1,341	-0,842	-0,385	0,126	0,954	1,881
8		-2,054	-1,476	-1,036	-0,358	0,279	1,036	1,881
9		-1,555	-0,975	-0,305	0,332	1,080	1,645	2,054
10		-1,476	-0,613	-0,025	0,412	1,036	1,645	
11		-2,054	-1,282	-0,878	-0,279	0,126	1,080	
12			-1,751	-1,282	-0,772	-0,253	0,496	1,645
13				-2,054	-1,476	-0,468	0,440	1,555
14					-1,645	-0,739	0,126	1,645

každého intervalu pro jednotlivé výroky na škálovém kontinuu. Vycházíme přitom ze zjištěných hodnot z_{ij}, uvedených v tab. VI.3. Šíři určitého intervalu určujeme rozdílem mezi hodnotou z_{ij} a hodnotou $z_{i\ j-1}$, tj. rozdílem mezi hranicemi sousedních intervalů. Tento rozdíl se značí w_{ij}; platí o něm vztah

$$w_{ij} = z_{ij} - z_{i\ j-1}$$

kde w_{ij} je šířka j-tého intervalu pro i-tý výrok.

V našem případě pro 1. výrok dostaneme tyto hodnoty:

$$w_{1,3} = -1{,}341 - -1{,}881 = 0{,}540$$
$$w_{1,4} = -1{,}036 - -1{,}341 = 0{,}305$$
$$w_{1,5} = -0{,}583 - -1{,}036 = 0{,}453$$

atd.

Výsledky jsou shrnuty v tab. VI.4.

Určili jsme takto hranice všech intervalů kromě horní hranice 9. kategorie /je vesměs 1,00/ a spodní hranice 1. kategorie /0,00/. Jim odpovídající hodnoty z_{i9} či z_{i1} by byly nekonečně velké a tedy prakticky nevyužitelné.

Z dat, uvedených v tab. VI.4., získáme škálové kontinuum tak, že vypočteme průměrné hodnoty /$\bar{w}_j$/ ve všech sloupcích; tyto hodnoty jsou užívány pro stanovení hranice a tedy také šířky jednotlivých intervalů. Začátkem kontinua je horní hranice prvního intervalu, jeho další hodnoty jsou určeny kumulovanými součty průměrných hodnot /řádek $c\bar{w}_j$ /. Tyto hodnoty tvoří vždy horní hranici jednotlivých intervalů. Chceme-li jednotlivé intervaly charakterizovat reprezentativně jedním číslem, pak je vhodné určit pro každý z nich

Tab. VI.4.

Sestavení škálových intervalů pro data z tab.VI.3.

	Následné intervaly							Škálové hodnoty	
	2-1	3-2	4-3	5-4	6-5	7-6	8-7	Výchozí	Revidované
1		0,540	0,305	0,453	0,683	0,894	0,887	2,996	+0,818
2					0,755	1,302	1,379	3,760	+1,582
3		0,737	0,407	0,458	0,752	0,598	0,578	2,293	+0,115
4	0,625	0,573	0,433	0,653	0,483	0,609	0,236	2,001	-0,177
5	0,974	0,612	0,493	0,177	0,441	0,437	0,671	1,893	-0,285
6				0,599	0,758	0,964	0,596	3,523	+1,345
7		0,540	0,499	0,457	0,511	0,828	0,927	2,932	+0,754
8		0,578	0,440	0,678	0,637	0,757	0,845	2,807	+0,629
9		0,640	0,610	0,637	0,748	0,565	0,409	2,168	-0,010
10		0,863	0,588	0,437	0,624	0,609		1,949	-0,229
11		0,772	0,404	0,599	0,404	0,954		2,891	+0,713
12			0,469	0,510	0,519	0,749	1,149	3,369	+1,191
13				0,578	1,008	0,908	1,115	3,502	+1,324
14					0,906	0,865	1,519	3,762	+1,584
$\sum w_j$	1,599	5,855	4,668	6,236	9,230	11,039	10,311		
n	2	9	10	12	14	14	12		
$\bar{w}_j$	0,800	0,651	0,467	0,520	0,659	0,788	0,859		
$c\bar{w}_j$	0,800	1,451	1,918	2,438	3,097	3,885	4,744		
Me	0,400	1,126	1,685	2,178	2,768	3,491	4,315		
H	-1,778	-1,052	-0,493	0,000	+0,590	+1,313	+2,137		
Interval	2	3	4	5	6	7	8		

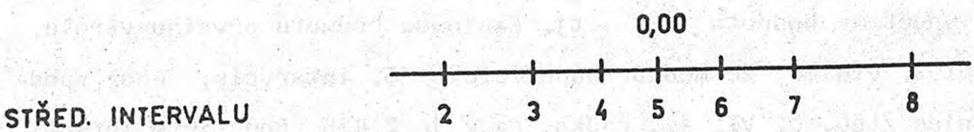
střední bod, tj. průměrnou hodnotu mezi hranicemi sousedních intervalů. U nejnižšího intervalu předpokládáme, že jeho spodní hranice se rovná nule. Takto vypočtené hodnoty jsou zachyceny v řádce Me.

Dále se stanoví nulový bod škály. Protože posuzovací kategorie jsou pozitivní, neutrální či negativní, získáváme zřejmě bipolární škálu. Je proto možno umístit nulový bod do středu prostředního intervalu /tj. intervalu číslo 5/ a ostatní body škály k němu vztáhnout tím, že odečteme od všech bodů hodnotu 2,178 /řádka H/. Takto vytvořené kontinuum je graficky znázorněno na obr. VI.1.

Z výpočtu i z grafického znázornění škály je zřejmé, že šířka jednotlivých intervalů na škále je proměnlivá. U extrémnějších posuzovacích kategorií je širší než u neutrálních kategorií. Pozitivní posouzení je spojeno se širšími intervaly než je tomu u negativních posudků. Vyplývá z toho, že v uvedeném příkladě je pozitivní posuzování relativně diferencovanější než negativní.

OBR.VI.1.

ŠKÁLOVÉ HODNOTY JEDNOTLIVÝCH VÝROKŮ.



VI.3. U m í s t ě n í j e d n o t l i v ý c h p o d n ě t ů n a š k á l e

Když jsme sestavili základní škálové kontinuum, můžeme vypočítat škálové hodnoty pro jednotlivé podněty. Používá se většinou mediánová hodnota z kumulativních proporcí, které stanovila skupina posuzovatelů pro jednotlivé výroky. Postupuje se dle následujícího vzorce:

$$S_i = u + \left(\frac{0{,}05 - \sum p_u}{p_j} \right) \bar{w}_j$$

kde S_i = škálová hodnota i-tého výroku

u = spodní hranice intervalu na škálovém kontinuum, do kterého spadá hodnota mediánu

$\sum p_u$ = součet relativních frekvencí, které leží pod intervalem, ve kterém leží medián

p_j = relativní frekvence posudků, ležících v intervalu, ve kterém je medián

$\bar{w}_j$ = šířka intervalu, ve kterém leží medián

Při výpočtu mediánových hodnot se musíme vrátit k tab.VI.2. a VI.4. Vypočteme hodnotu S_1, tj. škálovou hodnotu prvního výroku. Z tab.VI.2. vidíme, že medián bude ležet v 6. intervalu, jehož spodní hranice /tab. č. VI.,4., řádka $c\bar{w}_j$/ je 2,438. Pod tímto intervalem /tab. VI.2./ leží 0,28 případů; v intervalu, ve kterém leží medián, je 0,54 - 0,28 = 0,26 případů a jeho šířka /tab.VI.4./ je 0,659 /řádek $\bar{w}_j$/.

Pak tedy:

$$S_1 = 2{,}438 + \frac{0{,}50 - 0{,}28}{0{,}26} \cdot 0{,}659 = 2{,}438 + \frac{0{,}22}{0{,}26} \cdot 0{,}659 = 2{,}996$$

Pro 2. výrok dostaneme hodnotu:

$$S_2 = 3{,}097 + \frac{0{,}50 - 0{,}13}{0{,}57 - 0{,}13} \cdot 0{,}788 = 3{,}097 + \frac{0{,}37}{0{,}44} \cdot 0{,}788 = 3{,}760$$

Podobně postupujeme při výpočtu škálové hodnoty každého výroku, až konečně pro 14. výrok dostaneme hodnotu:

$$S_{14} = 3{,}097 + \frac{0{,}50 - 0{,}23}{0{,}55 - 0{,}23} \cdot 0{,}788 = 3{,}097 + \frac{0{,}27}{0{,}32} \cdot 0{,}788 = 3{,}762$$

Tak získáme škálové hodnoty jednotlivých výroků /předposlední sloupec v tab. VI.4./ a můžeme nyní pracovat buď se všemi, anebo vybereme jen některé, a to obvykle ty, které leží na vytvořeném škálovém kontinuu od sebe přibližně ve stejných vzdálenostech. Kdybychom chtěli vytvořit pětibodóvou škálu, zvolili bychom pravděpodobně výroky 5, 3, 8, 12, 14 se škálovými hodnotami 1,893 - 2,293 - 2,807 - - 3,369 - 3,762. Jejich diference nejsou sice zcela stejné, ale v dané situaci se tomuto požadavku relativně nejlépe přibližují. Vytvořili bychom vlastně novou škálu na principu techniky zdánlivě stejných intervalů.

Jestliže máme k dispozici větší počet podnětů s podobnými škálovými hodnotami, pak varianci posudků vypočítáme u všech. Podobně jako u metody zdánlivě stejných intervalů se stanoví kvartilové rozpětí Q. Vybíráme pak ty podněty, u nichž je rozpětí nejmenší, protože jsou zřejmě jednoznačněji posuzovány.

Škálové hodnoty jednotlivých výroků můžeme ještě upravit tak, aby byly ve shodě s konečnou podobou škály /řádek H v tab. VI.4/; od všech hodnot reprezentujících jednotlivé výroky, odečteme konstantu 2,178. Upravené hodnoty jsou zaznamenány v posledním sloupci tab. VI.4. a mohli bychom je též graficky zanést do obr. VI.1.

Je třeba ještě upozornit na jisté omezení popsaného postupu vytváření škály. Nelze určit šířku krajních intervalů /tj. spodní hranici intervalu č.1. a horní hranici intervalu č.9/. Kdyby se stalo, že více než 50% posudků určitého výroku by bylo kumulováno v jednom z nich, nebylo by možno vypočítat škálovou hodnotu právě toho intervalu, ve kterém leží většina posudků. Pro tento případ existuje úprava jak stanovit škálu. Protože se s podobnou situací setkáváme jen zřídkakdy /obvykle se totiž takový výrok vypouští jako příliš extrémní a málo diferencující/, nebudeme ji zde probírat; zájemce odkazujeme na studie Edwardse /1957/ nebo Krasemannové /1966/.

Technika následných intervalů má řadu dalších úprav a zjemnění. Existují techniky pro zpracování matic, ve kterých některé údaje chybí. Je možno sledovat vnitřní konzistenci získaných dat. Byly vypracovány techniky pro výpočet škálových hodnot jednotlivých posuzovaných podnětů při menším počtu posuzovacích kategorií; použití mediánu se totiž nedoporučuje, je-li počet kategorií menší než 9. S těmito detailnějšími úpravami základního postupu je možno se seznámit v monografii Guilfordově /1954, kap.10/ či Torgersonově /1958/.

VI.4. P o u ž i t í m e t o d y n á s l e d n ý c h i n t e r v a l ů

Zásadní výhodou právě popsané metody je její jednoduchost jak z hlediska posuzovatelů, tak z hlediska experimentátora. Výpočet intervalové škály je zdánlivě složitý, ale dá se dobře algoritmovat a použít výpočetní techniku. Chceme-li se dopracovat k intervalové škále, pak je výhodné používat tuto metodu, kdykoliv je to možné. Je zde sice omezení, které je dáno předpokladem normality distribuce získaných dat, ale protože tento předpoklad můžeme ověřit, není příliš veliké riziko, že se dopustíme chyby. Další výhodou je i to, že můžeme dodatečně přidávat nové podněty pro posouzení a srovnávat je s podněty posuzovanými dříve. Stejně tak můžeme metodu následných intervalů poměrně snadno několikrát aplikovat a použít ji jako prostředek pro sledování dynamiky názorů či postojů v čase. Pro tyto výhody bývá tato metoda často doporučována, i když není tak známá jako metody předcházející.

Jistou nevýhodou této metody - podobně jako je tomu i u metody zdánlivě stejných intervalů - je to, že se opírá o názory skupiny expertů, kteří vytvářejí jakousi normu. Tam, kde podobnou skupinu nemůžeme vytvořit, anebo kde pokusné osoby jsou značně jiné než je skupina expertů /např. věkově nebo v patologických případech/, naráží praktické použití obou technik na zásadní potíže. Proto se vyvíjely i takové postupy, ve kterých nejsme odkázáni na skupinu posuzovatelů, ale opíráme se pouze o odpovědi skutečných pokusných osob. Tím se sice ochuzujeme o normativní aspekt škálování a zůstáváme jen u popisné formy, ale vyhýbáme se problémům reprezentativnosti skupiny expertů /a také technickým potížím, spojeným s vytvářením takové skupiny/. V dalších kapitolách popíšeme škálovací metody, které vycházejí přímo z dat, získaných od pokusných osob.

Kapitola VII.

METODA SUMOVANÝCH ODHADŮ

VII.1. Ú v o d

Ve třicátých letech navrhl Likert /1932/ jinou škálovací metodu, která nahrazuje metodu zdánlivě stejných intervalů. Rozdíl je v tom, že vypustil časově náročný krok vytváření škály na základě mínění expertů. Jednotlivé výroky se přímo předkládají pokusným osobám, které pouze uvádějí svůj souhlas či nesouhlas s nimi. Tato metoda je v mnohé blízká školnímu známkování, ale i některým testovým technikám, a bývá proto též někdy uváděna pod názvem bodová /známková/ škála.

Práce s metodou sumovaných odhadů má dvě fáze: vytvoření škály a stanovení polohy určité osoby na škále.

VII.2. P o s t u p p ř i v y t v o ř e n í š k á l y

Zvolí se řada výroků či tvrzení, která se vztahují ke škálovanému tématu /třeba škála postojů k televizi/. Je dán určitý počet kategorií, které vyjadřují souhlas s jednotlivými výroky či jejich odmítnutí. Většinou se užívají tyto kategorie:

	velmi souhlasím -	souhlasím -	nevím -	nesouhlasím -	silně nesouhlasím
body:	5	4	3	2	1

Může se třeba vyskytnout výrok: "Televize vede člověka k pasívnímu vnímání a nerozvíjí jeho tvůrčí schopnosti", nebo: "Televize patří k životnímu stylu dnešního člověka", či "Hlavní význam televize je v pohotovém zpravodajství" atd. Ke každému výroku pokusná osoba zaujme své vlastní stanovisko. Při přípravě bodovacího systému je třeba pečlivě dbát na to, jak jsou odpovědi věcně orientovány. V prvé položce znamená souhlas negativní postoj k televizi, v druhé položce znamená souhlas naopak pozitivní postoj. Znamená to, že u některých otázek je třeba systém bodování obrátit tak, aby pozitivní postoj byl vždy hodnocen více body než postoj negativní.

Při sestavování výroků, které se mají posuzovat, je třeba racionálně zvažovat, které výroky skutečně vyjadřují postoj ke sledovanému jevu a které nikoliv. V našem případě pro zjišťování postoje k televizi by nebylo vhodné tvrzení: "Meteorologické předpovědi, které vysílá televize, jsou spolehlivé". Zde určitá odpověď vyjadřuje názor na kvalitu meteorologických služeb, ale nikoliv na televizi. Vhodné by nebylo ani tvrzení: "Většina lidí se domnívá, že televize je pro děti užitečná". Z odpovědi neplyne, zda osoba posuzuje názor ostatních lidí /např. souhlasí s tím, že většina lidí považuje televizi za užitečnou/, či zda posuzuje svůj vlastní názor na užitečnost televize pro děti.

I když nejsou přesné předpisy, jak vybírat jednotlivá tvrzení, přece jen můžeme stanovit obecnější zásady: Je na místě zvolit větší počet položek, protože často teprve při shromažďování se ukáže, že některé nejsou dostatečně citlivé či neměří sledovaný jev. Výroky mají být jednoznačně a pozitivne formulované. Všechny ukazatele by měly mít stejnou závažnost /např. při školním známkování nemá většinou známka z obsahu úlohy stejnou váhu, jako známka z její úpravy/. Jestliže pracujeme s výroky, které jsou opačně pólovány, pak je vhodné, aby jich bylo v obou směrech zhruba stejně.

Je vhodné již předem uvažovat o validitě i reprezentativnosti jednotlivých položek. Máme-li řádně sestaveny a promyšleny jednotlivé položky, shromáždíme odpovědi od většího počtu pokusných osob. Mělo by jich být alespoň 200. Řekněme, že u určitého výroku získáme data, shrnutá v tab. VII.1.

Tab. VII.1.

Postojová škála skupiny osob / N = 200 / k určitému tvrzení

	silně odmítám	odmítám	nevím	souhlasím	silně souhlasím
f	36	86	44	24	10
p	0,18	0,43	0,22	0,12	0,05
c.p	0,18	0,61	0,83	0,95	1,00
Me	0,09	0,395	0,72	0,89	0,975
z	-1,340	-0,292	0,583	1,221	1,960
H=z+1,340	0,000	1,048	1,923	2,561	3,300
zaokrouhleno	0	1	2	3	3

V první řádce tab. VII.1. je uveden počet osob, které odpověděly určitým způsobem. Tyto hodnoty jsou v řádce p vyjádřeny relativní četností, a v řádce c.p kumulativně sečteny. Dále potřebujeme stanovit střední bod v jednotlivých kumulativních kategoriích /řádek Me /. Vypočtou se dle vzorce

$$Me_x = cp_{x-1} + \frac{p_x}{2}$$

kde cp_{x-1} = relativní frekvence případů, které leží pod sledovanou kategorií;

p_x = relativní frekvence případů, které leží v rámci sledované kategorie.

V našem příkladě pro kategorii "nevím" platí

$$0,61 + \frac{0,22}{2} = 0,72$$

Další postup je obvyklý. Pro jednotlivé hodnoty ve 4. řádku tabulky se stanoví hodnota z. Škála se zakotví tak, aby její negativní pól měl škálovou hodnotu 0, což se docílí tím, že k hodnotám z přičteme nejvyšší hodnotu, která se vyskytla se záporným znaménkem /v našem případě 1,34/. Výsledná škála bývá někdy zaokrouhlována na celá čísla; v uvedeném příkladě by odpovědi "souhlasím" a "silně souhlasím" spolu splývaly.

Takto postupujeme u všech položek a v konečné fázi získáme pětibodové škály pro jednotlivé výroky. Během doby se ukázalo, že předvedený výpočet je možno zjednodušit a pracovat pouze s počtem odpovědí v jednotlivých kategoriích; konečné výsledky obou postupů se od sebe příliš neliší /Scheuch, 1962/.

VII.3. Stanovení polohy určité osoby na škále

Pro stanovení polohy určité osoby na vytvořené škále sečteme škálové hodnoty, které odpovídají tomu, jak daná osoba jednotlivé výroky posoudila. Odtud vznikl název - metoda sumovaných odhadů. V příkladě, uvedeném v tab. VII.1., by osoba, která odpoví "odmítám", získala 1,048 bodu; osoba reagující volbou kategorie "silně souhlasím" by získala 3,30 bodu atd. Výsledky jednotlivých osob se tedy ve všech položkách sečtou a jejich průměr definuje polohu dané osoby na škále. I zde se však v poslední době připouští, že je možno vycházet přímo ze zvolené odpovědi, aniž by se vážila zjištěnou škálovou hodnotou.

VII.4. Rozlišovací síla položek

Kromě polohy jednotlivých osob na škále můžeme na základě získaných dat uvažovat o diskriminační síle jednotlivých položek. Orientačně lze konstatovat, že málo diferencují ty položky, u kterých se většina dat nakupí na jednom extrémním pólu, resp. položky, na které se neodpovídá ve všech pěti kategoriích. Přesněji postupujeme tak, že z celkového počtu osob vybereme 25% s nejvyšším celkovým skóre a 25% s nejnižším skóre. Pro tyto skupiny vypočteme průměrné hodnoty všech položek a významnost rozdílu mezi nimi ověříme t-testem či některou jeho neparametrickou analogií. Výsledné hodnoty t seřadíme do pořadí a v další práci používáme jen ty položky, u kterých byla hodnota t nejvyšší.

Můžeme dále zjišťovat reprezentativnost jednotlivých položek, a to v podstatě analogicky, jako je tomu při položkové analýze v psychometrice. Srovnáváme rozložení dat v určité položce s rozložením dat pro všechny ostatní položky. Tam, kde je shoda malá, položku obvykle vyloučíme. Můžeme touto cestou ověřovat, zda soubor položek tvoří homogenní celek, či zda se nerozpadá do dvou či více podskupin.

Jestliže některé z položek vyloučíme /pro malou diskriminační schopnost či pro malou reprezentativnost či z jiných důvodů/, pak je třeba vypočítaná skóre - ať již pro celkovou škálu či pro polohu jednotlivých osob - opravit. V další práci se počítá s takto upravenými údaji.

VII.5. Použití

Likertova škála sumovaných odhadů je vcelku jednoduchým nástrojem, který se však v poslední době užívá poněkud méně než dříve. Pro svou jednoduchost se dá poměrně dobře použít v předběžných orientačních výzkumech. Sestavená škála nemá v postatě intervalový charakter; jedná se jen o zdokonalenou škálu pořadovou. Hodí se spíše pro zjišťování rozdílů v postojích mezi skupinami osob, než pro spolehlivé měření postojů určitého jedince.

Stejný numerický výsledek, zjištěný u několika jedinců, může vyplývat ze zcela jiné struktury odpovědí na jednotlivé položky, a tedy možná i z odlišných postojů. Toto tvrzení lze demonstrovat na následujícím příkladě výsledků 3 osob /A, B, C/ ve 3 položkách /I., II., III./.

	Kategorie				
Položka	1	2	3	4	5
I.	B	A		C	
II.	C	A,B			
III.	C	A	B		

Všechny 3 osoby získaly 6 bodů. ale osoba A odpověděla na všechny položky kategorií 2, osoba B odpovídala kategoriemi 1, 2, 3 a osoba C kategoriemi 4, 1, 1.

Škála sumovaných odhadů má značnou výhodu v tom, že je poměrně snadno sestavitelná, dobře se předkládá pokusným osobám a umožňuje nepřímé měření sledovaného jevu. Neptáme se přímo na postoj k určitému jevu, ale oklikou zjišťujeme stupeň souhlasu s jistými tvrzeními, která sledovaný postoj reprezentují. Ve své původní podobě se tato škála užívá dnes poměrně zřídka, ale některé její formy najdeme často v kombinaci s jinými škálovacími technikami.

Kapitola VIII.

ŠKÁLOGRAMOVÁ ANALÝZA

VIII.1. Ú v o d

Škálogramovou analýzu /někdy též zkráceně škálogram/ vypracoval N. Guttman začátkem čtyřicátých let. Cílem bylo vytvořit jednodimenzionální a homogenní škály. Všechny užité položky se mají vztahovat k jedinému faktoru, ale přitom by se od sebe měly lišit svou závažností. Můžeme zde vidět jakousi analogii s testy pro měření inteligence, ve kterých se obvykle určitý subtest skládá z řady položek. Všechny měří stejný faktor /např. prostorovou představivost/ a jsou seřazeny od nejlehčí k nejtěžší. Zaznamenává se, kolik z nich zkoumaná osoba správně zodpověděla.

Guttman viděl hlavní slabinu tradičních škálovacích metod v tom, že posuzované podněty mohou být hodnoceny z různých hledisek. Vznikají proto škály, které jsou vícedimenzionální, a proto těžko interpretovatelné. Tato skutečnost uniká pozornosti uživatelů a ve škálovací praxi není většinou brána v úvahu. Stává se pak, že některé položky korelují s celkovou škálou jen velmi nízko, a tím škálu znehodnocují. Jindy se objevují škály, jejichž položky se rozpadají do dvou či více trsů a vytvoření jediné škály je značně problematické.

Základní Guttmanovou snahou bylo sestavit řadu podnětových položek tak, aby bylo možno konstatovat, že osoba, která odpoví kladně na x-tou položku, získala vyšší skóre než osoba, která na ni neodpoví. Znamená to dále, že ten, kdo odpoví pozitivně na položku s určitou hodnotou, musí odpovědět pozitivně i na všechny položky, které mají nižší hodnotu. Známe-li počet bodů, který pokusná osoba dosáhla, pak můžeme jednoznačně stanovit strukturu jejích odpovědí, tj. určit, u kterých položek byla odpověď pozitivní a u kterých položek byla odpověď negativní.

Principy Guttmanovy techniky byly rozvedeny v řadě studií, z nichž mnohé jsou těžko dostupné; pro orientaci uvádím alespoň základní prameny: Guttman /1944, 1947, 1954/, Suchman /1950/, Krasemannová /1966/.

VIII.2. I l u s t r a t i v n í p ř í k l a d y G u t t m a n o v ý c h š k á l

Prvním příkladem Guttmanovy škály mohou být výsledky vzájemného utkání 6 mužstev odbíjené či košíkové na sportovním turnaji /tab. VIII.1./. Mužstvo A porazilo všechna ostatní, mužstvo B porazilo ostatní s výjimkou mužstva A atd. Mužstvo F pak prohrálo se všemi. Přitom se nemohou vyskytovat nerozhodné výsledky. Víme-li, kolik bodů určité mužstvo dosáhlo, pak můžeme jednoznačně určit jeho místo v pořadí a stanovit s kým vyhrálo a s kým prohrálo.

Podobná situace by nastala při měření postojů, kdybychom měli k dispozici 6 odstupňovaných kladných výroků, charakterizujících určitý postoj. Pak pokusná osoba, která odpoví pozitivně na i-tý výrok, musí mít kladnější postoj, než osoby, které na něj odpoví záporně. Osoby, které získají stejný počet bodů, musí mít i stejné postoje. Tím je překonán nedostatek většiny škálovacích technik, ve kterých se může snadno stát, že stejný počet bodů vyplývá z odlišných struktur odpovědí, jak jsme viděli na příkladu z konce předcházející kapitoly.

Tab. VIII.1.

Výsledky sportovního turnaje šesti mužstev /A - F/

	A	B	C	D	E	F	Počet vítězství
A		X	X	X	X	X	5
B	O		X	X	X	X	4
C	O	O		X	X	X	3
D	O	O	O		X	X	2
E	O	O	O	O		X	1
F	O	O	O	O	O		0

Pro ilustraci uvedeme často citovanou škálu, která má charakter škálogramu. Při komplexním výzkumu amerického vojáka, který byl prováděn v letech druhé světové války, se kromě jiného sledovaly též fyziologické projevy strachu před bojem. Zjistilo se toto jednoznačné pořadí /od projevů nejintenzivnějších po nejslabší/: 1/ Pomočení - 2/Pokálení - 3/Zvracení - 4/Pocit mdloby - 5/ Ztuhlost - 6/Žaludeční nevolnost - 7/Třes - 8/Zvedání žaludku - - 9/Tlukot srdce. Znamená to, že kdo pociťoval příznak č.5, pociťoval i příznaky 6 - 9. Ukázalo se též, že některé další fyziologické reakce /např. studený pot/ do této stupnice nezapadaly, a proto je nebylo možno začlenit do tohoto systému.

Jiným příkladem Guttmanových škál jsou jednotlivé dimenze v dotazníku interpersonálních vztahů FIRO b /viz Kožený, 1973/. Je konstruován tak, že každá škála má 12 položek, které postupně zachycují stále intenzívnější úroveň sledovaných potřeb.

VIII.3. P o s t u p p ř i v y t v á ř e n í š k á l o g r a m u

Sestavení jednodimenzionálních škál Guttmanova typu naráží v praxi na značné potíže a je i po technické stránce poměrně složité. Vlastní vývoj škály má dvě oddělené fáze :

a/ Vytvoření žebříčku položek, které by tvořily monotónní funkci /v optimálním případě lineární/.

b/ Sestavení intenzitní složky škálogramu, jejíž pomocí zjišťujeme, do jaké míry jsou si pokusné osoby jisty svými úsudky.

VIII.3.a. Postup při výběru položek

Nejprve se sestaví jednotlivé položky - ať již ve formě výroků či otázek. Postupuje se většinou intuitivně a zprvu se raději volí větší počet položek /alespoň 12/. Čím více položek je k dispozici, tím spíše se podaří z nich vybrat část, která splňuje požadavky škálogramu; na druhé straně je však celá procedura technicky náročnější. Obvykle se již v prvních orientačních pokusech ukáže, že se některé položky nehodí do očekávané struktury, a proto se vyloučí.

Tab. VIII.2.

Příklad výchozích dat pro sestavení škálogramu /ve sloupcích 1 jsou značeny pozitivní odpovědi, ve sloupcích 0 negativní odpovědi/

p.o.	A		B		C		D		E		Kladné body	Chyby
	1	0	1	0	1	0	1	0	1	0		
1.	X		X		X		X		X		5	0
2.	X		X		X		X		X		5	0
3.	X		X		X		X		X		5	0
4.	X		X		X		X			X	4	0
5.	X		X		X		X			X	4	0
6.	X		X		X			X	X		4	2
7.	X		X		X			X	X		4	2
8.	X		X		X			X		X	3	0
9.	X		X		X			X		X	3	0
10.	X		X		X			X		X	3	0
11.	X		X			X	X			X	3	2
12.	X		X			X	X			X	3	2
13.	X		X			X		X		X	2	0
14.	X		X			X		X		X	2	0
15.	X		X			X		X		X	2	0
16.	X		X			X		X		X	2	0
17.	X			X		X	X			X	2	2
18.	X			X		X		X		X	1	0
19.	X			X		X		X		X	1	0
20.	X			X		X		X		X	1	0
21.		X	X			X		X		X	1	2
22.		X		X		X		X		X	0	0
23.		X		X		X		X		X	0	0
24.		X		X		X		X		X	0	0
25.		X		X		X		X		X	0	0
f	20	5	17	8	10	15	8	17	5	20		12
chyby	0		1		0		3		2			

Zvolené položky se předloží většímu počtu pokusných osob a jejich odpovědi se zachytí do škálogramové tabulky. Jako příklad /tab. VIII. 2./ zde zvolíme situaci, kdy máme 5 položek, o kterých předpokládáme, že jsou citlivě odstupňovány z určitého hlediska /třeba co do obtížnosti nebo co do vyjádření intenzity postoje apod./ Skupina osob na ně odpovídá buď "ano" nebo "ne". Existují i situace složitější, např. na škále s větším počtem položek žádáme jako odpověď pouze jedinou volbu /na kterou z položek byste mohl odpovědět "ano", a od které byste tedy již odpovídal "ne"/, anebo několik voleb /určete 3 položky, které by nejlépe vyjadřovaly váš názor/ atd. Tyto možné postupy jsou popsány detailněji jinde /Torgerson, 1958, str. 307-317/.

Na první pohled je z tab. VIII.2. zřejmé, že obtížnost položek vzrůstá. Chceme určit, zda získaná data skutečně odpovídají požadavkům škálogramu. Existuje několik metod postupu, z nichž nejprve popíšeme Cornellskou techniku:

V každé položce je třeba stanovit bod, ve kterém se výsledky dichotomizují, tj. najít místo, kde se mění odpovědi a přecházejí z kategorie 1 do kategorie 0. Pro stanovení dělících bodů platí tato pravidla:

a/ uspořádat data tak, aby vzniklo co nejméně chyb. Za chybu sa považují odpovědi, které se liší od většiny odpovědí ležících nad či pod dělícím bodem;

b/ aby v žádné položce nebylo více chyb než správných odpovědí.

Dělící body jsou u každého indikátoru označeny vodorovnou čarou. V tab.VIII.2. zjišťujeme, že v položce A není žádná chyba, v položce B je jedna chyba atd. Celkem jsme zjistili šest chyb a tento počet již nelze dalšími úpravami tabulky VIII.2. dále snížit. Pořadí pokusných osob, které je uvedeno v tabulce, je sestaveno záměrně tak, abychom minimalizovali počet chyb a aproximovali ideální škálu. Nemusí nikterak odpovídat pořadí osob, v jakém bylo s nimi prováděno šetření či jak jsme je původně zaznamenali.

V uvedeném příkladě byl postup poměrně snadný, protože počet položek i počet osob byl poměrně malý. Pořadí položek z hlediska jejich obtížnosti je zřejmé. Také optimální seřazení pokusných osob /pořadí řádek/ bylo možno najít lehce - do jisté míry metodou pokusu a omylu. Při složitějších podmínkách, kdy pracujeme s větším počtem položek a s rozsáhlými soubory pokusných osob, je obvykle třeba použít samočinný počítač pro nalezení optimální sestavy.

Pro vyhodnocení, zda získaná data odpovídají ideální škále, navrhl Guttman koeficient reprodukovatelnosti / rk / . Určuje, zda zjištěný počet chyb je zanedbatelný či nikoliv. Výpočet se provede podle vzorce:

$$rk = 1 - \frac{F}{i \cdot N}$$

kde F = celkový počet chyb /v našem příkladě 6/

i = počet podnětů

N = počet pokusných osob

tedy: $rk = 1 - \frac{6}{5 \cdot 25} = 1 - 0{,}048 = 0{,}952$

Ideální případ, kdy rk = 1,00 se prakticky nevyskytne.

Jako dostatečné přiblížení k perfektní škále se obvykle uvádí hodnota rk = 0,85 až 0,90 . Guttman sám původně uváděl hodnotu 0,85, ale později zpřísnil požadavky na 0,90.

I když koeficient reprodukovatelnosti je základním kritériem existence jednodimenzionálních škál, potřebuje ještě určité doplnění. Čím více položek máme k dispozici, tím je věrohodnost koeficientu rk větší; při přísném postupu bychom měli u dichotomických dat mít alespoň 10 položek /z tohoto hlediska je výše uvedený příklad poněkud zjednodušující/. Každá jednotlivá položka má mít vlastní koeficient reprodukovatelnosti alespoň 0,85. Vypočte se analogicky podle vzorce pro celkové rk s tím rozdílem, že i /= počet podnětů/ = 1. V uvedeném příkladě jsou dílčí koeficienty reprodukovatelnosti tyto:

$rk_A = 1,00$, $rk_B = 0,96$, $rk_C = 1,00$, $rk_D = 0,88$, $rk_E = 0,92$.

Hodnota koeficientu reprodukovatelnosti se uměle zvýší i tím, když více než 80% osob odpovídá v mnoha položkách stejným způsobem; takové položky je lépe vyloučit.

Je užitečné sledovat podrobně ty osoby, u kterých se objevili chyby proti očekávaným výsledkům. Jestliže určitý počet osob se dopouští stejné chyby, pak je to signálem, že působí další dimenze. Proto není předpoklad jednodimenzionální škály pro získaná data přiměřený.

Získané hodnoty rk je možno interpretovat několikerým způsobem: Malý počet chyb znamená, že škála je vyhovující a blíží se uspokojivě k ideální škále. Velký počet chyb /koeficient reprodukovatelnosti je nižší než hranice 0,85/ může znamenat několik sku-

tečností:
a/ Většina chyb se vyskytuje pouze u jediné položky; pak je třeba tuto položku vyloučit a znovu vypočítat rk ; jeho hodnota se nepochybně zvýší. b/Soubor položek lze rozdělit do dvou podskupin - pak je třeba zvolit jednu z nich /případně ji rozšířit o nové položky/ a znovu vypočítat rk . c/Soubor položek nelze rozumně rozdělit, ale přece jen je zřejmá společná dimenze / rk ≙ 0,80 / ; vytvoří se tak zvaná quasi-škála. Může vzniknout z různých příčin - buď jsou chyby rozloženy náhodně ve všech položkách nebo se položky dělí do několika podskupin, anebo není homogenní skupina pokusných osob /někteří reagují konzistentně, jiní se odchylují od ideálního modelu v důsledku povrchního postupu či zavedením jakéhosi nového kontinua, či z řady jiných důvodů/.Quasi-škály mohou být použity jako základ pro analýzu latentních struktur /viz kap. X./. d/Soubor položek neodpovídá požadavkům škálogramu.

Jinou metodu pro stanovení stupně reprodukovatelnosti navrhl Goodenough /1944/; Edwards /1948/ ji po určitých úpravách považuje za přesnější. Při výpočtu vyjdeme z dat v tab. VIII.2. a stanovíme proporce, v jakých pokusné osoby odpovídaly na jednotlivé položky buď pozitivně či negativně. Výsledky jsou shrnuty v tab. VIII.3.

V řádce p_2 jsou hodnoty p zaokrouhleny. Kdyby se proporce pozitivních a negativních reakcí prosadily zcela pravidelně tak, jak by tomu bylo v ideální Guttmanově škále, dostali bychom takové kombinace výsledků, které jsou shrnuty v tab. VIII.4.

Tab. VIII.3.

Proporce pozitivních a negativních odpovědí /data z tab. VIII.2/

	Položky									
	A		B		C		D		E	
	1	0	1	0	1	0	1	0	1	0
f	20	5	17	8	10	15	8	17	5	20
p	0,8	0,2	0,68	0,32	0,4	0,6	0,32	0,68	0,2	0,8
p_2	0,8	0,2	0,7	0,3	0,4	0,6	0,3	0,7	0,2	0,8

Tab. VIII.4.

Kombinace odpovědí u 5 položek, odpovídající proporcím uvedeným v tab. VIII.3.

	Položky					Počet bodů
	A	B	C	D	E	
1.	1	1	1	1	1 p=0,2	5
2.	1	1	1	1 p=0,3	1	
3.	1	1	1 p=0,4	1	0	4
4.	1	1 p=0,7	1	0	0	3
5.	1 p=0,8	1	0	0	0	2
6.	1	1	0	0	0 p=0,8	
7.	1	1	0	0 p=0,7	0	
8.	1	0	0 p=0,6	0	0	1
9.	0	0 p=0,3	0	0	0	0
10.	0 p=0,2	0	0	0	0	

Z tab. VIII.4. vyčteme, že při zisku 5 bodů se musely objevit odpovědi 11111. Při 4 bodech 11110, při 3 bodech 11100 atd. Sledujeme nyní výsledky z tab. VIII.2. a zachycujeme, kolik odchylek od těchto ideálních pořadí se vyskytlo u jednotlivých osob; tato data jsou v posledním sloupci tab. VIII.2. Součet chyb je 12; dosadíme-li tuto hodnotu do vzorce pro výpočet rk, dostaneme

$$rk = 1 - \frac{12}{5 \cdot 25} = 1 - \frac{12}{125} = 1 - 0{,}096 = 0{,}904$$

Při použití Goodenoughovy techniky dostáváme pro rk hodnoty nižší. V našem příkladě je však koeficient reprodukovatelnosti, vypočtený podle obou metod, vyšší než 0,90, a proto můžeme uvedené výsledky považovat za data, odpovídající požadavkům jednodimenzionální škály.

Některé další metody pro zjištění, zda výsledky odpovídají požadavkům pro škálogramovou analýzu, uvádí Torgerson /1958, kap. 12./. V této souvislosti stojí za zmínku, že Guttmanův škálogramový postup se v podstatě shoduje s teorií homogenních testů, kterou vyvinula Loevingerová /1948/.

Škálogramovou analýzu je možno provádět i tehdy, když výchozí položky mají více možných odpovědí, které jsou pořadově uspořádány /3 možnosti: souhlasím - nevím - nesouhlasím; případně 5 možností/. V praxi se však jen málokdy zdaří vytvořit za těchto podmínek položky, které by odpovídaly škálogramovým požadavkům. Guttman doporučuje slučovat v rámci jednotlivých položek odpovědi v jednotlivých kategoriích tak, aby se počet kategorií snižoval a zároveň se snižoval i počet chyb. V extrémním případě redukce postupuje tak dlouho, až se odpovědi dichotomizují a postupujeme podle výše uvedených pravidel. Toto řešení však má řadu nesnází; často se uvádí, že jde o postup čistě formální, při kterém se mohou slučovat i kategorie, které spolu obsahově nesouvisejí. Rozbor těchto postupů, které se častěji vyskytují v sociologických výzkumech, a zároveň jejich zhodnocení lze nalézt u Krasemannové /1966/ či Torgersona /1958/.

VIII.3.b. Intenzitní složka škálogramu

Při sestavení škálogramu se kromě vlastní škály též sleduje, kde na něj leží nulový bod, tj. bod, ve kterém se negativní postoj mění v pozitivní /případně lehčí položky v obtížné apod./. Tam, kde škála má jen 3 nebo 5 bodů, můžeme zcela racionálně předpokládat, že nulový bod bude ležet v intervalu, který odpovídá prostřednímu bodu škály. Pokud je však škála složitější a nemá nulový bod, který by bylo možno vyčíst z formulace jednotlivých položek škály, pak musí být pro jeho nalezení užita speciální technika. U škály, která je uvedená na začátku kapitoly a týká se projevů strachu před bojem, je zřejmě rozdíl v závažnosti jednotlivých položek. První položky /pomočení, pokálení/ jsou příznaky těžkého strachu, zatímco poslední položka /tlukot srdce/, je příznakem jen velmi mírného strachu. Kdesi mezi nimi leží bod, který celou škálu dělí na dvě části: těžký a mírný strach. Přitom tento bod nemusí ležet přesně uprostřed škály, a ani nemusí být pro všechny osoby stejný. Nalezení dělícího bodu škálogramu se řeší následujícím postupem.

Když byl sestaven škálogram a získány odpovědi /typu "ano" - - "ne"/ od jednotlivých osob, zjišťuje se intenzita postoje či názorů dodatečnou otázkou, která zní zhruba asi takto: "Jak obtížně se vám na toto tvrzení /či na tuto otázku/ odpovídalo?". Pokusné osoby označí svůj dojem číselně v rozmezí 1 - 5:, číslo 5 znamená "odpovídal jsem snadno a bez váhání" a naopak číslo 1 znamená "rozhodoval jsem se velmi těžko". /Můžeme se ovšem ptát i jinými způsoby: "Jak silně jste o svém tvrzení přesvědčen?"; pak hodnota 1 znamená "velmi slabě", hodnota 5 "velmi silně" apod./ Postupně probereme všechny položky. Srovnáme-li nyní všechna získaná data /tj. obsahovou stránku odpovědí s intenzitou postoje/, dostáváme dvě řady údajů, které můžeme korelovat či jinak srovnávat. Ideální model je znázorněn v tab. VIII.5. Jednotlivé odpovědi, značené X, by měly tvořit jakýsi trojúhelník.

Tab. VIII.5.

Příklad jednoznačné vazby mezi intenzitou postoje a obsahovou stránkou odpovědí

Intenzita postoje	Obsah položek								
	negativní						pozitivní		
	1	2	3	4	5	6	7	8	9
5	X								X
4		X						X	
3			X				X		
2				X		X			
1					X				

Z idealizovaných dat plyne, že posuzovatel je zcela jednoznačně přesvědčen o svém postoji k extrémně formulovaným položkám – ať negativním či pozitivním /proto uvede na subjektivní škále intenzity postoje hodnotu 5/. Naopak položka č.5, která je vcelku neutrální, vzbuzuje také největší rozpaky; naznačuje shodu mezi neutrálním obsahovým bodem a nevýraznou intenzitou postoje. Takto sestavená data umožňují určit takový bod na škálogramu, který dichotomizuje postoje na kladné a záporné. U té položky, u které je určitou osobou /či skupinou osob/ uvedena nejnižší hodnota intenzity postoje, předpokládáme indiferentní či neutrální postoj.

Zpravidla nezískáváme tak ideální výsledky, jaké jsou označeny v tab. VIII.5.; obvykle je obsazena většina políček. Obecnější tendenci musíme teprve vypočítat. Postupujeme takto: Nejprve sestavíme data do tabulky, v jejíž jednotlivých polích jsou zachyceny počty osob, které na určitou položku /p_i/ reagovaly určitým stupněm intenzity postoje /i_j/. Stanovíme mediánové hodnoty jednotlivých sloupců. Vyjadřujeme je obvykle v relativních hodnotách. Rozsah intenzitní stupnice považujeme za 100% a mediánové hodnoty transformujeme tak, aby odpovídaly tomuto rozsahu.

Postup ilustrujeme na následujícím myšleném příkladě: byla vytvořena škála o 9 položkách a zároveň byla sledována intenzita postoje o 7 bodech. Získaná data jsou shrnuta v tab. VIII.6.

Z tab. VIII.6. plyne, že celkem 105 osob vyjádřilo výrazně pozitivní postoj /je vyjádřen bodem 1/ ke sledovanému jevu; z nich pak 53 uvádí při tom stupeň intenzity 7, tj. "zcela jednoznačně souhlasím", 21 vyjádřilo postoj intenzitou č. 2 atd. Z dat v každém sloupci určíme medián. Jako hrubou aproximaci uvádíme řádky, ve kterých leží; v tab. VIII.6. jsou označeny čtverečkem. Chceme-li však získat přesnější hodnoty mediánu, při jejichž výpočtu se respektuje jak nelineárnost vztahu mezi obsahy položek, tak mezi údaji o intenzitě, postupujeme takto:

$$Me_i = k\%\,(Me-1) + \left(\frac{\frac{\sum f_i}{2} - \sum f_i\,(Me-1)}{f_{Me}}\right)\left(k\%_{Me} - k\%\,(Me-1)\right)$$

kde $k\%\,(Me-1)$ = kumulativní součet procent případů, které se vyskytly v nižších kategoriích, než ve které je medián /hledej v posledním sloupci tab. VIII.6./;

$\sum f_i$ = celkový počet případů v i-té položce;

$\sum f_i\,(Me-1)$ = součet případů, které se vyskytly v nižších kategoriích, než ve které je medián;

f_{Me} = počet případů v kategorii, ve které je medián;

$k\%\,(Me)$ = kumulativní součet procent případů, které se vyskytly od nejnižší až po mediánovou kategorii.

Výpočet demonstrujeme numericky na postoji s obsahem č. 9. Medián leží v intenzitní kategorii č. 6. Po dosazení do vzorce dostaneme

$$Me_9 = 65,7 + \left(\frac{\frac{39}{2} - 14}{7}\right)(82,8 - 65,7) = 65,7 + \frac{5,5}{7}\,17,1 = 79,14$$

Pro položku č. 1, kde medián leží v intenzitní kategorii č. 7, dostaneme

$$Me_1 = 82,8 + \left(\frac{\frac{105}{2} - 52}{53}\right)(100 - 82,8) = 82,8 + \frac{0,5}{53}\,17,2 = 82,96$$

Tab. VIII.6.

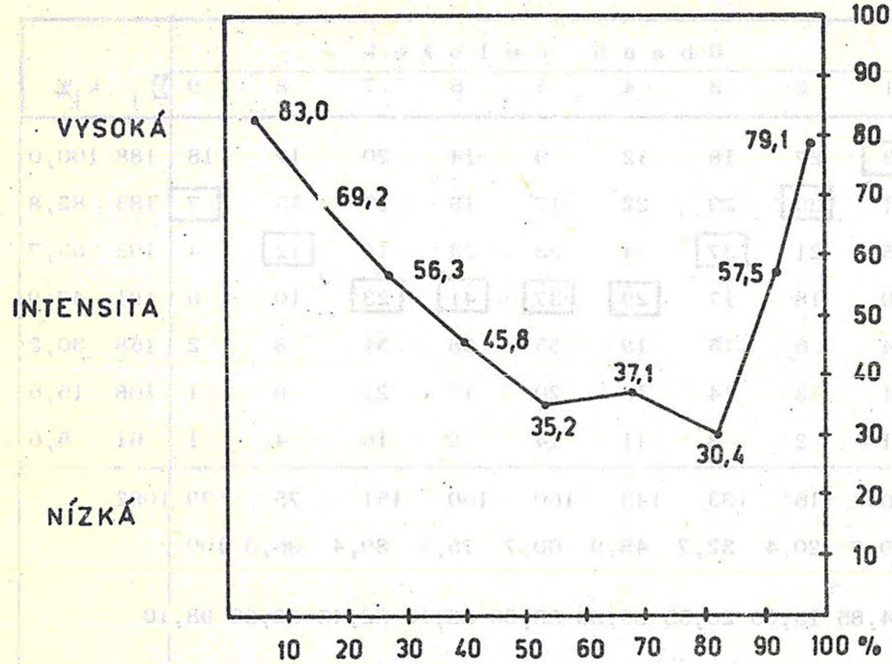
Stanovení nulového bodu při škálogramové analýze

Intenzita postoje	Obsah položek 1	2	3	4	5	6	7	8	9	$\sum f_j$	$k_j\%$
7	$\boxed{53}$	27	18	12	9	14	20	17	18	188	100,0
6	21	$\boxed{39}$	29	22	17	18	15	15	$\boxed{7}$	183	82,8
5	15	21	$\boxed{37}$	34	28	23	19	$\boxed{12}$	4	193	65,7
4	10	18	17	$\boxed{29}$	$\boxed{37}$	$\boxed{41}$	$\boxed{23}$	10	6	191	47,9
3	4	6	15	19	35	38	31	8	2	158	30,2
2	1	3	14	16	20	17	27	9	1	108	15,6
1	1	2	3	11	14	9	16	4	1	61	5,6
$\sum f_i$	105	116	133	143	160	160	151	75	39	1082	
$k_i\%$	9,7	20,4	32,7	45,9	60,7	75,5	89,4	96,3	100		
střed intervalu	4,85	15,05	26,55	39,30	53,30	68,10	82,45	92,85	98,10		
Me+/	83,0	69,2	56,3	45,8	35,2	37,1	30,4	57,5	79,1		

+/ vyjádřen v hodnotách stupnice $k_j\%$.

Výsledky pro všechny položky jsou shrnuty v poslení řádce tab. VIII.6. Tuto tabulku můžeme vyjádřit též graficky takto:

OBR. VIII. 1. INTENZITNÍ SLOŽKA ŠKÁLOGRAMU.

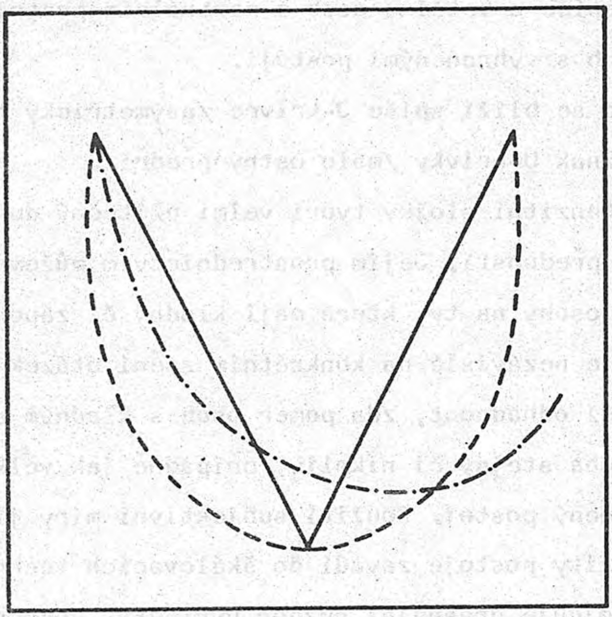


Z výsledků je zřejmé, že v tomto případě nejnižší bod intenzity postoje neleží uprostřed škály, ale je nápadně posunut k jejímu negativnímu pólu. Znamená to, že nejobtížněji se vyjadřoval postoj u položky č. 7 škálogramu; mohli bychom ji považovat za neutrální. Jednoznačně to však nemůžeme tvrdit, protože i u 5. a 6. položky dostáváme přibližně stejné hodnoty, což svědčí o tom, že hranice přechodu mezi kladným a záporným postojem není v tomto příkladě příliš ostrá.

Z pouhého náhledu na možné tvary křivek intezitní složky škalogramu, které jsou vyjádřeny v obr. VIII.2, můžeme očekávat minimálně 3 různé typy vztahů. Mezi jednotlivými druhy křivek, které značíme písmeny V, U, J, je ovšem plynulý přechod. Můžeme je interpretovat zhruba takto:

OBR. VIII. 2.

TYPY INTENZITNÍ SLOŽKY ŠKÁLOGRAMU.



Křivka tvaru V naznačuje, že sledovaný soubor osob se ostře dělí na 2 skupiny, které jsou zhruba stejně početné. Ostrý dělící bod umožňuje poměrně snadno vést dělítko mezi osobami s pozitivním a osobami s negativním vztahem k danému jevu. Neutrální bod leží zhruba v polovině škály.

Křivka tvaru J naznačuje, že dělítko se posouvá buď k pozitivnímu, či k negativnímu pólu; počet osob s jedním postojem je značně vyšší než počet osob s opačným postojem. Poměrně ostrý předěl umožní opět obě skupiny dobře diferencovat.

Křivka tvaru U naznačuje, že ve sledované populaci se vyskytuje značné procento osob s názory, které nejsou jednoznačně vyhraněny; proto se asi těžko rozhodují. Nelze mluvit o dvou vyhraněných skupinách, ale spíše o většině osob s neutrálním postojem a o dvou menších skupinách s vyhraněnými postoji.

Náš příklad se blíží spíše J-křivce /asymetrický tvar/, ale obsahuje též náznak U-křivky /málo ostrý předěl/.

Analýza intenzitní složky tvoří velmi užitečný doplněk škálogramu a má řadu předností. Jejím prostřednictvím můžeme poměrně snadno rozdělit osoby na ty, které mají kladný či záporný postoj; toto rozdělení je nezávislé na konkrétním znění otázek či položek škály. Můžeme též odhadnout, zda poměr osob s kladným a záporným postojem je zhruba stejný či nikoliv, případně jak velké procento osob má nevyhraněný postoj. Použití subjektivní míry jistoty rozhodnutí a intenzity postoje zavádí do škálovacích technik další dimenzi, která umožňuje přesnější rozbor výsledků. Podobný parametr je možno užívat i při jiných škálovacích metodách, avšak nejčastěji se užívá právě při škálogramové analýze.

VIII.4. Zhodnocení škálogramové analýzy

Vlastní metoda škálogramu byla řadou autorů kritizována. V praxi se totiž ukazuje, že vytvoření škály podle Guttmanových kritérií je velmi obtížné a v mnoha případech takřka neproveditelné. Původní požadavek, aby škálogram měl 10 - 12 otázek, nebývá často dodržen a autoři se po vyhodnocení řady položek spokojí s menším počtem položek. Vzniká pak otázka, zda konečná škála zachycuje skutečně celé kontinuum a zda nedochází k nežádoucímu zjednodušení či obsahovému posunu.

Někteří kritici dokonce tvrdili, že škálogram se skládá z položek, které jsou jen jistou variancí jediné základní položky. Toto tvrzení je do jisté míry oprávněné, je však třeba si uvědomit, že škálogram je určen výhradně a explicitně pro kvantifikování jednodimenzionálních jevů. V tom je jeho přednost i slabina. Jen málokdy sledujeme jev, který by byl jednoznačně jednodimenzionální. Škálovaná skutečnost bývá obvykle složitější a škálogram může někdy být nežádoucím gnoseologickým zjednodušením. Naproti tomu tam, kde sledovaný jev je přesně definován a pohybuje se dle našich zkušeností podél jediné dimenze, je využití škálogramu vhodné a umožňuje detailnější a přesnější analýzu než mnohé jiné škálovací metody. Právě proto bývá i dnes poměrně často užíván v mnoha oblastech psychologického výzkumu i praxe.

Existují různé úpravy základního postupu, hlavně po stránce výpočetní techniky; informace o nich lze najít ve studiích Torgersona /1958/ či Šafáře /1969/.

Kapitola IX.
KOMBINOVANÁ ROZLIŠOVACÍ ŠKÁLA

IX. 1. Ú v o d

Edwards a Kilpatrick /1948a, 1948b/ vyvinuli dvě formy jiné techniky, ve které se pokusili sloučit některé výhody škálogramu s výhodami techniky stejných intervalů a sumovaných odhadů. Základní otázkou, kterou se oba autoři zabývali, je problematika výběru položek pro určitou škálu /většinou pro měření postojů/. Škálogramová analýza nepřináší žádné pevné a objektivní kritérium; prvotní výběr položek je v podstatě intuitivní a vlastní výběr je veden pouze snahou vytvořit obsahově homogenní škálu. Diskriminační technika se snaží tuto slabinu překonat tím, že využívá přednosti jiných škálovacích technik. V podstatě jde o racionálnější postup, jak z velkého množství položek vybrat jen několik, které by splňovaly předpoklady Guttmanových škál, a přitom byly dostatečně reprezentativní pro sledovanou problematiku.

IX. 2. V y t v o ř e n í š k á l y

Vlastní postup vytváření škály má opět několik fází: Začátek postupu je shodný jako při použití techniky stejných intervalů. Shromážděné výroky /v původní práci Edwardse a Kilpatricka jich bylo 168/ se předloží skupině expertů a ti je třídí do určitého počtu kategorií. Poté se pro každý výrok vypočte kvartilová odchylka Q , která je v podstatě polovinou kvartilového rozpětí; jeho výpočet byl popsán již dříve.

Vyloučí se položky s vysokou hodnotou Q , tj. ty, u nichž se zjistí velká variabilita a které nejsou zřejmě jednoznačně posuzovány. Není stanoveno jednotné kvantitativní kritérium, podle kterého by se položky vypouštěly. V citované práci byly vyloučeny položky, u nichž se zjistilo, že Q > 1,29 ; šlo zhruba o polovinu výroků. Většinou se uvádí analogické empirické kritérium, podle kterého je na místě vyloučit alespoň polovinu výroků s vyšší variabilitou.

V další fázi se zbylé položky zpracovávají analogicky, jako je tomu při metodě sumovaných odhadů. Sestaví se několik kategorií, které vyjadřují stupeň souhlasu či nesouhlasu s jednotlivými položkami. Obvykle se užívá 5 - 6 kategorií. Edwards s Kilpatrickem použili 6 kategorií: silně souhlasí - souhlasí - mírně souhlasí - mírně nesouhlasí - nesouhlasí - silně nesouhlasí. Opět je třeba dbát na směr jednotlivých položek po stránce obsahové; u pozitivního tvrzení je třeba kategorie "naprosto souhlasím" hodnocena 5 body, zatímco stejná odpověď u negativního tvrzení nutně dostane 0 bodů /máme-li 6 kategorií odpovědí/.

Nyní je možno předložit takto připravený materiál řadě pokusných osob; i zde platí, že čím více osob pracuje, tím lépe. U každé osoby se zjistí celkový počet bodů, který je dán součtem bodů, uvedených v jednotlivých položkách. Pak se vybere čtvrtina osob, u kterých byl zjištěn celkově nejpozitivnější postoj, a čtvrtina osob s výrazně negativním postojem.

V další fázi se analyzují data v jednotlivých položkách. Postupuje se tak, že u každé položky se zjistí empirické rozložení posudků osob s pozitivním postojem /nejvyšší čtvrtina osob/ a u osob s negativním postojem /nejnižší čtvrtina/. Příklad dat tohoto typu je uveden v tab. IX.1., ve které jsou zachyceny výsledky 200 osob /100 osob s kladným postojem a 100 s negativním postojem/ ve 2 rozličných položkách /X,Y/; bylo použito stupnice s 5 kategoriemi.

Aby bylo nyní možno využít techniku škálogramu, je třeba pětistupňovou kategorizaci redukovat na 2 kategorie. Provede se to srovnáním výsledků skupiny "kladné" se skupinou "zápornou". Empiricky to znamená, že se hledá bod, podle kterého by bylo možno data dichotomizovat. Stanoví se dělítko, při kterém je minimalizován počet odpovědí, které jsou u skupiny "záporné" v kategoriích ležících nad dělítkem a u skupiny "kladné" pod dělítkem. Postup demonstrujeme na datech z tab. IX.1. u položky X. Kdybychom vedli dělítko mezi kategoriemi, které jsou bodovány 4 a 5, pak bychom měli 3 osoby ze "záporné" skupiny nad dělítkem a 35 + 13 + 7 + 7 osob z "kladné" skupiny pod ním /tedy celkem 65 osob/. Analogicky získáme:

při dělení mezi bodem 4 a 3: 3 + 5 + 13 + 7 + 7 = 35

- " - - " - 3 a 2: 3 + 5 + 28 + 7 + 7 = 50

- " - - " - 2 a 1: 3 + 5 + 28 + 30 + 7 = 73

Minimální hodnota se tedy objeví, jestliže dichotomizujeme kategorie mezi bodem 4 a 3.

U položky Y dostáváme tyto hodnoty:

dělení mezi bodem 5 a 4: 3 + 31 + 15 + 9 + 2 = 60

- " - 4 a 3: 3 + 7 + 15 + 9 + 2 = 36

- " - 3 a 2: 3 + 7 + 10 + 9 + 2 = 31

dělení mezi bodem 2 a 1: 3 + 7 + 10 + 34 + 2 = 56

V tomto případě je vhodné dělení mezi bodem 3 a 2.

Tab. IX.1.

Rozložení odpovědí osob s kladným a s negativním postojem u 2 vybraných položek sledované škály

Kategorie	Body	POLOŽKA X		POLOŽKA Y	
		"Kladná" skupina	"Záporná" skupina	"Kladná" skupina	"Záporná" skupina
silně souhlasí	5	38	3	43	3
souhlasí	4	35	5	31	7
není si jist	3	13	28	15	10
nesouhlasí	2	7	30	9	34
silně nesouhlasí	1	7	34	2	46

Na základě provedené dichotomizace získávají všechny kategorie, které leží nad dělícím bodem, váhu 1 a všechny pod ním váhu 0. Tento postup se provede postupně u všech položek a pak se určuje jejich rozlišovací síla. Edwards a Kilpatrick užívají čtyřpolní koeficient korelace – r_φ , který se užívá běžně pro výpočet vztahů mezi dvěma alternativními znaky /viz např. Siegel, 1956, nebo Břicháček, Hampejsová, 1963/. Pro jeho výpočet platí vztah:

$$r_\varphi = \frac{ad - bc}{\sqrt{(a+b)\ (c+d)\ (a+c)\ (b+d)}}$$

přičemž a, b, c, d jsou četnosti ve 4 polích dle schématu, uvedeného v tab. IX.2.

Tab. IX.2.

Schematické zachycení dichotomizace kategorií

Kategorie	"Záporná" skupina	"Kladná" skupina	$\sum$
1	a	b	a+b
0	c	d	c+d
$\sum$	a+c	b+d	N

U položky X, kde dělítko bylo vedeno mezi kategoriemi 4 a 3, platí a = 8, b = 73, c = 92, d = 27, a tedy

$$r_{\varphi} = \frac{73 \cdot 92 - 8 \cdot 27}{\sqrt{81 \cdot 119 \cdot 100 \cdot 100}} = \frac{6500}{9818} = 0{,}662$$

Pro položku Y dostaneme hodnotu r_{φ} = 0,693

Nyní je možno definitivně vybrat položky pro konečnou škálu. Data se většinou vyjádří graficky v osovém systému; na ose x se nanášejí hodnoty jednotlivých položek, získané metodou sumovaných odhadů, a na ose y jejich diskriminační síla, definovaná koeficientem r_{φ}. Vyberou se ty položky, které leží na sledovaném kontinuu zhruba ve stále stejných intervalech a které přitom mají poměrně nejvyšší diskriminační sílu. Často je možno sestavit i dvě paralelní formy škály.

Takto sestavená škála se znovu předloží dalším osobám a ověřuje se, zda její položky skutečně tvoří jednodimenzionální škálu; postupuje se tak, jak bylo popsáno v kapitole o škálogramové analýze.

IX.3. Z h o d n o c e n í

Popsaný postup je technicky i z hlediska vynaloženého času značně náročný a pracný, ale má řadu předností. V prvé řadě odstraňuje subjektivnost při volbě položek škálogramové analýzy. Použití kvartilové odchylky Q vylučuje položky víceznačné či položky s velkou variancí a použití koeficientu r_{φ} vyloučí položky, které nediskriminují dostatečně citlivě. Kromě toho využití výsledků několika dostatečně velkých skupin pokusných osob v různých fázích přípravy škály zajišťuje i celkově její větší objektivitu /ovšem za předpokladu, že skupiny osob jsou vybrány podle běžných pravidel pro výběr z populace/. Naproti tomu značná pracnost postupu a zároveň i jistá skepse vůči reálnému významu jednodimenzionálních škál vede k tomu, že se tato technika užívá v praxi poměrně zřídka.

Edwards a Kenney /1946/ ve své kritice postojových škál považovali obvykle používané postupy za příliš zjednodušující - ať se již jedná o Thurstonovu techniku stejných intervalů či o Likertovu techniku sumovaných odhadů. Z tohoto rozboru vyplynul požadavek kombinovat několik škálovacích technik, využít jejich předností a vyhnout se jejich slabinám, a dospět tak k dokonalejší škále. Odtud vyplynula snaha vytvořit kombinované škálovací techniky. Toto úsilí nebylo zatím doceněno - snad pro technickou obtížnost. Je však zřejmé, že se jedná o velmi plodný námět, jak o tom svědčí úspěšná studie Claussova /1962/.

Kapitola X.

ANALÝZA LATENTNÍ STRUKTURY

X.1. Ú v o d

P.F. Lazarsfeld vyvinul zajímavou škálovací techniku, která tvoří jakýsi protiklad k úvahám Guttmanovým. Jeho základní myšlenka je zhruba tato: Zjištěné vztahy mezi různými položkami škál /ale i dotazníků, testů atd./ můžeme vysvětlit působením faktorů, které nemůžeme přímo pozorovat /jsou latentní/. Lazarfeld usiluje o to, aby mohl ze získaných dat odvodit takovou latentní prostorovou strukturu, kterou by vysvětlil empirická data. Analýza latentní struktury tedy nevychází z předem odvozené či teoreticky očekávané obecné klasifikace sledovaných jevů. Hledaná klasifikace se teprve odvodí z empirického materiálu a ze statistických údajů o jednotlivých indikátorech.

Na první pohled by se mohlo zdát, že se zde jedná o základní myšlenku faktorové analýzy, která se zde aplikuje na kvalitativní data; ve skutečnosti jde o zcela jinou techniku, Guttman se zaměřil na analýzu přímých vztahů mezi různými škálovými položkami v jediné dimenzi; Lazarsfeld usiluje o odkrytí hypotetické struktury, umožňující vysvětlit zjištěná data, která jsou často nepřehledná a zatížená mnohými chybami či nepřesnostmi. Tím se liší od Guttmanova základního požadavku - vytvořit jednoduché kontinuum a minimalizovat různé chyby či nepřesnosti. Lazarsfeldův přístup umožňuje pracovat s větším počtem položek a získat jejich prostřednictvím přesnější poznatky /pro Guttmana je větší počet položek spíše nepřekonatelnou překážkou/. Z jistého hlediska je možno Guttmanův přístup chápat jako speciální případ Lazarsfeldovy koncepce, a to takový, kdy získaná data lze vysvětlit existencí jediné latentní dimenze.

Matematický model je dosti složitý a uvedeme ho zde ve značně zjednodušené podobě. Zájemce o bližší výklad odkazujeme na původní práce Lazarsfeldovy, z nichž pro psychologa jsou asi nejlépe přístupné studie z r. 1954 a 1959, či na výklad Greenův /1954, str. 359 a dále/. Z uvedených Lazarsfeldových studií má první spíše intuitívní charakter, druhá obsahuje důležitý výklad psychologické teorie, z které autor vycházel při vývoji celé metody, a detailní popis jednotlivých kroků numerického postupu. Složitost celé procedury vedla však k tomu, že se analýza latentních struktur užívala jen zřídka; lze však očekávat, že využití moderní výpočtové techniky tuto metodu oživí.

X.2. I n f o r m a t i v n í   v ý k l a d   a n a l ý z y   l a t e n t n í   s t r u k t u r y

Předpokládejme, že existuje jednodimenzionální kontinuum, které budeme dále sledovat. Jedná se třeba o známou škálu lži, užívanou v některých dotaznících /tento příklad není zcela korektní, ale užijeme ho pro jednodušší výklad/. Škála se skládá z řady dichotomických položek, o kterých předpokládáme, že měří skutečně tu vlastnost, kterou sledujeme. Dále předpokládáme, že jednotlivé osoby získávají určitý počet bodů, kterými je definováno jejich umístění na sledovaném kontinuu. Rozložení odpovědí jednotlivých osob a rozložení bodů, které získávají, nám předem není známo a ani nemáme důvodu očekávat určitý typ rozložení. Po shromáždění dat můžeme jednak stanovit pozici každého jedince na škále lži, jednak můžeme určit, v jakém poměru se u jednotlivých položek vyskytují odpovědi pozitivní a negativní.

Řekněme, že užíváme pro škálování "lži" položky ze známého Eysenckova dotazníku:

1/ Přišel jste někdy pozdě na schůzku nebo do práce? ano - ne

2/ Míváte občas názory či myšlenky, o nichž byste si přál, aby je jiní lidé neznali? ano - ne

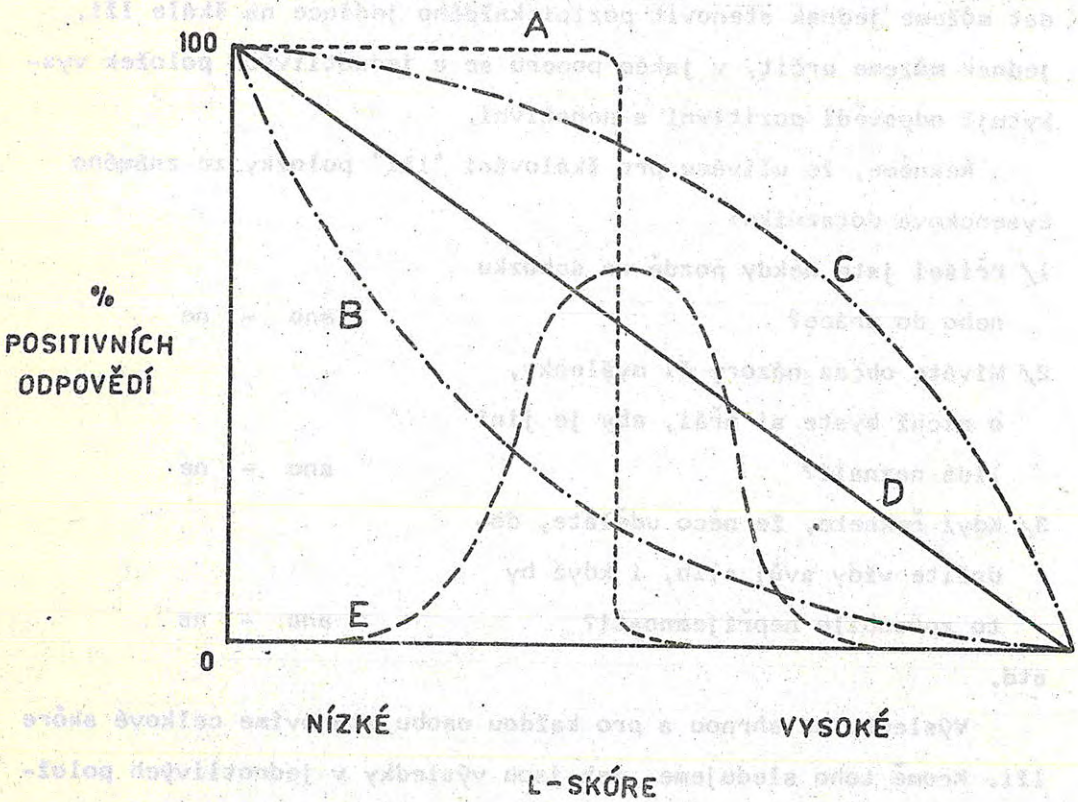
3/ Když řeknete, že něco uděláte, dodržíte vždy svůj slib, i když by to způsobilo nepříjemnosti? ano - ne

atd.

Výsledky se shrnou a pro každou osobu stanovíme celkové skóre lži. Kromě toho sledujeme, jak jsou výsledky v jednotlivých položkách rozloženy. Předpokládáme, že pravděpodobnost p_i - tj. pokusná osoba odpoví pozitivně na i-tou položku - je funkcí, f_i /X/, její pozice je kontinuu X , tj. na škále lži. Výsledky můžeme znázornit graficky a dostaneme tak stopové křivky /trace line/, zachycující změny pravděpodobnosti výskytu pozitivních či negativních reakcí v jednotlivých položkách v závislosti na latentním kontinuu, tj. na celkovém skóre lži. Hypotetické příklady jsou znázorněny na obr. X.1.

OBR. X.1.

PŘÍKLADY STOPOVÝCH KŘIVEK.



Z tohoto grafu můžeme vyčíst řadu poznatků o vlastnostech jednotlivých položek. Položka A plně odpovídá požadavkům Guttmanovým. Osoby s vysokým skóre lži na ni vesměs odpovídají negativně, osoby s nízkým skóre lži /L/ na ni odpovídají pozitivně. Přitom dělítko je velmi ostré a můžeme je přesně určit. Ostatní křivky mají v různých bodech různou pravděpodobnost výskytu pozitivní odpovědi. Křivky B, C a D vykazují kladné odpovědi u osob s nízkým L; jejich pravděpodobnost však postupně klesá. Tento pokles má pokaždé poněkud jiný charakter. U křivky B je pokles zprvu velmi rychlý; většina osob tedy rychle přejde ke "lživým" odpovědím. U křivky C je průběh opačný a u křivky D je pokles lineární. Křivka E má jiný charakter; pozitivní odpovědi se objevují nejčastěji u osob s průměrným L-skóre a klesají, když L-skóre přechází k extrémním hodnotám /podobný typ vztahu by se v našem příkladě asi neobjevil, ale v řadě jiných situací bývá častý/. Stopové křivky se sestrojují postupně pro jednotlivé položky. Základním problémem je sledovat pravděpodobnost současného výskytu určitých odpovědí /třeba kladných/ v několika položkách a přitom na různých místech latentního kontinua.

K dispozici máme vždy data, vyplývající z odpovědí jednotlivých osob na určitý počet položek. Odpovědi jsou kvalitativní /ano - ne/. U každé osoby známe strukturu jejích odpovědí; tak např. na 5 položek odpověděla osoba A systémem $\{ - + + + - \}$, tj. na první a pátou otázku odpověděla negativně a na druhou, třetí a čtvrtou odpověděla pozitivně. Je zřejmé, že se může vyskytnout celkem 2^k různých kombinací pozitivních a negativních odpovědí /k = počet položek/; v našem případě při pěti položkách je celkem 32 možností. Vyšetříme-li velké množství osob, zjistíme, že frekvence výskytu některých kombinací je častější než jiných. Proporci osob, které odpověděly stejně jako osoba A, označuje Lazarsfeld $p_{234/15}$, tj. pravděpodobnost, že v daném souboru osob se vyskytnou pozitivní odpovědi na položky č.2, 3, 4 a negativní na položky 1, 5. Stejně tak je možno stanovit frekvenci jakékoliv dvojice /p_{ij}/ či trojice /p_{ijk}/, odpovědí na zvolené položky i, j, k ... Jednotlivé položky jsou do jisté míry na sobě závislé /měří jednu latentní vlastnost, tj. tendenci podávat ne zcela pravdive odpovědi/, a proto ten, kdo odpoví pozitivně na položku i , odpoví i na položku j spíše pozitivně než negativně. Proto platí pro kteroukoliv dvojici položek vztah $p_{ij} - p_i p_j > 0$. Tato nerovnost je jakousi analogií interakcí v čtyřpolních tabulkách.

Poněkud jinak vypadá situace, analyzujeme-li homogenní skupinu pokusných osob, tj. takovou skupinu, u které je L-skóre stejné /řekněme skupinu osob, které na 2 z 5 položek odpověděly pozitivně/. Pokud položky měří určitou vlastnost a přitom nejsou sestaveny dle principu škálogramu - tj. od položky, která je nejčastěji pozitivně zodpovídána až po položku nejméně často pozitivně zodpovídanou - pak jsou na sobě nezávislé. Lazarsfeld používá pojmu latentní nezávislost. Pozitivní odpověď na jednu položku ještě nic neříká o tom, jaká odpověď se vyskytne u položky jiné. Tento fakt vyjadřuje Lazarsfeld vztahem

$$p^x_{234/15} = p^x_2 \; p^x_3 \; p^x_4 \; q^x_1 \; q^x_5$$

kde x znamená skupinu osob, patřících do určité homogenní třídy,

p_i = proporce osob, které na i-tou otázku odpověděly pozitivně

$$q_i = 1 - p_i$$

Uvedená rovnice zahrnuje dva další předpoklady: a/ v rámci homogenní latentní třídy x se frekvence pozitivních odpovědí na jednotlivé položky blíží latentní pravděpodobnosti p_i ; b/ v rámci homogenní skupiny jsou odpovědi v jednotlivých položkách na sobě nezávislé.

Lokální nezávislost platí pouze pro jednotlivé latentní třídy; není splněna, sloučíme-li data z jednotlivých tříd. Toto tvrzení je zdánlivě paradoxní, můžeme je však objasnit na následujícím příkladě: Uvažujme o prvních dvou položkách pro měření L-skóre, které jsme uvedli výše. Řekněme, že z hlediska sledovaného parametru existují tři latentní třídy osob: 1. s vysokým L-skóre, 2. s průměrným skóre, 3. s nízkým skóre. Při vytvoření těchto tříd se opíráme o údaje, které byly získány z rozboru odpovědí na 10 položek. Osoby s vysokým L-skóre na většinu otázek odpovídají způsobem, který není obvyklý. Tyto odpovědi budeme značit znaménkem + ; druhý typ odpovědí značíme - . Hypotetické výsledky jsou shrnuty v tab. X.1.

U první položky se objevilo ve skupině osob s vysokým L-skóre celkem 79 pozitivních reakcí ze 103 možných /tj. 77%/; ve druhé skupině to bylo 56% a ve třetí skupině 31%. U druhé položky byly počty zjištěných pozitivních reakcí tyto: 81%, 49% a 16%. Orientačně můžeme z těchto dat vyčíst, že druhá položka je poněkud citlivějším ukazatelem než prvá. Určíme ještě, jaké procento osob v jednotlivých skupinách odpovědělo pozitivně na obě položky /údaje v levém horním rohu tabulek/: 65%, 28% a 3%.

Tab. X.1.

Vztahy 2 položek u 3 skupin osob, klasifikovaných podle celkového L-skóre

	Třída 1. Vysoké L 1. položka			Třída 2. Střední L 1. položka			Třída 3. Nízké L 1. položka		
	+	-	Σ	+	-	Σ	+	-	Σ
2. položka +	67	17	84	35	26	61	3	14	17
2. položka -	12	7	19	34	29	63	30	59	89
Σ	79	24	103	69	55	124	33	73	106

Sloučíme-li výsledky všech tří skupin, dostaneme data, shrnutá v tabulce X.2.

Tab. X.2.

Vztahy dvou položek u celého souboru osob /data z tab.X.1./

		Položky 1 +	Položky 1 -	Σ
Položka 2	+	105	57	162
	-	76	95	171
	Σ	181	152	333

Při sloučení všech dat je zřejmá závislost mezi oběma položkami. Osoby, které odpověděly ve směru L-skóre na první položku, odpoví v druhé položce spíše pozitivně, než by tomu bylo u osob, které v první položce odpověděly proti směru L-skóre. Na první položku odpovědělo pozitivně 181 osob a z nich 105 /58%/ odpovědělo pozitivně i na druhou položku. Naproti tomu ze 152 osob, které v první položce reagovaly negativně, jen 57 /42%/ odpovědělo na druhou položku obráceně. Výše jsme uvedli, že při závislých výsledcích platí nerovnost

$$p_{ij} - p_i p_j > 0$$

Po dosazení našich dat:

$$\frac{105}{333} - \frac{181}{333} \cdot \frac{162}{333} = 0{,}3153 - 0{,}2644 = 0{,}0509 > 0$$

Pokud stejný výpočet provedeme u dat rozdělených do tří podskupin /tab. X.1./, pak dostaneme výsledky jiné; asociace v první skupině je mírně pozitivní, v druhé skupině nulová a ve třetí skupině mírně negativní.

Tím je prokázáno výchozí tvrzení; můžeme je interpolovat asi takto: odpovědi na 2 položky určité škály jsou v pozitivní vazbě, protože obě jsou indikátorem sledované vlastnosti. Jestliže však sledujeme skupinu osob, které si jsou z hlediska této vlastnosti velmi podobné, pak sledované indikátory jsou na sobě nezávislé, Obráceně je možno říci, že za homogenní skupinu osob můžeme považovat takovou skupinu, u které zjistíme nezávislost zvolených indikátorů. Při sloučení homogenních skupin osob do jednoho celku se objeví statistické asociace mezi jednotlivými indikátory.

Přejdeme k dalšímu myšlenkovému kroku, který se provádí při analýze latentní struktury. Latentní prostor není sice ještě znám, ale můžeme již říci, že půjde o takovou strukturu, která dokáže rozložit výchozí populaci osob do homogenních podskupin. V našem příkladě jsme si skupinu rozdělili do tří skupin podle výsledků na jediném latentním kontinuu /L-skóre/. Při obecné analýze však nemáme k dispozici předem stanovená pravidla třídění, podle kterých bychom klasifikovali zjištěná data /tj. frekvence určitého typu odpovědí v jednotlivých položkách/. Takové třídění je třeba odvodit ze vztahů mezi daty, která byla zjištěna v jednotlivých položkách.

Východiskem jsou opět stopové křivky, které vznikly na základě rozboru empirických údajů. O jejich konstrukci jsme mluvili na začátku kapitoly. Jde v podstatě o zachycení změn frekvence kladných odpovědí v dané položce a ve všech homogenních latentních třídách. Úvaha je dále vedena asi takto: Předpokládejme, že máme c homogenních tříd a n položek. Uvažujme o dvou položkách. Frekvence určité kategorie odpovědí / třeba kladných/ v těchto položkách v celé populaci /tj. při sloučení c-tříd/ je možno stanovit z následujících rovnic:

$$p_1 = \sum_{x=1}^{c} v^x \; p_1^x$$

$$p_2 = \sum_{x=1}^{c} v^x p_2^x$$

$$p_{12} = \sum_{x=1}^{c} v^x p_1^x p_2^x,$$

kde p_1^x znázorňuje, pro kterou podskupinu x se stanoví pravděpodobnost p ;

v^x znázorňuje tu část populace, která spadá do podskupiny x .

Součtové znaménko značí, že slučujeme výsledky přes všechny podskupiny.

Protože však sledujeme větší počet položek, dostáváme rovnice složitější, jako např.:

$$p_{12345} = \sum_{x=1}^{c} v^x p_1^x p_2^x p_3^x p_4^x p_5^x$$

Jestliže sestavíme dostatečně složité rovnice, které odpovídají experimentálním podmínkám, pak je celý problém vyřešen. Rovnice právě zmíněného typu nazývá Lazarsfeld rovnice vysvětlující /accounting equations; v německých textech se objevuje pojem Gleichungen der Rechnungslegung/.

Vysvětlující rovnice jsou základem pro celou analýzu. Jejich soubor dovoluje určit z empirických dat parametry latentních struktur. Formalizují algebraicky diagnostickou proceduru, jejímž prostřednictvím dospíváme od empirických dat k latentní pozici dané osoby. Vyjadřují vztah mezi latentním a aktuálním prostorem. Kdyby se měnila pozice osoby v očekávané struktuře, pak by se v závislosti na této změně měnila i pravděpodobnost odpovědí.

Ve skutečnosti jen výjimečně můžeme sledovat jednu osobu opakovaně v řadě situací; většinou sledujeme řadu osob pouze jedenkrát. Předpokládáme, že jsou umístěny na různých místech latentní struktury. Můžeme ovšem najít několik osob, které jsou umístěny na stejném místě; považujeme je z tohoto hlediska za identické.

Tento výklad základních myšlenek analýzy latentní struktury je velmi schematický. Vlastní postup je značně složitější a nebudeme jej v této souvislosti podrobněji rozebírat. Zájemce se s ním seznámí poměrně nejlépe na instruktivním příkladě, který uvádí Lazarsfeld /1954/ či z obecného rozboru Torgersonova /1958, str. 361 a dále/, ve kterém jsou popsány i některé další analogické škálovací postupy.

X.3. S o u h r n

Analýza latentní struktury je nesporně velmi podnětným prostředkem pro škálování složitých vztahů, které existují v realitě. Myšlenkové základy jsou nosné a na první pohled přiměřené pro řešení řady otázek experimentální i užité psychologie, a to tam, kde máme k dispozici řadu dichotomických údajů.

Metoda sama je ovšem značně náročná. V prvé řadě se dá rozumně aplikovat pouze tehdy, máme-li k dispozici rozsáhlá data od stovek osob; je tedy výhodnější pro sociology či demografy než přímo pro psychology. Také matematická stránka celé techniky je dosti složitá a často překračuje možnosti běžných psychologických pracovišť. Snad teprve využití samočinných počítačů usnadní postup a zpřístupní ji většímu počtu psychologů. V současné literatuře se s analýzou latentní struktury setkáváme poměrně málo. Zřejmě teprve budoucnost jasněji ukáže, zda tato na první pohled slibná metoda nevede do slepé uličky či zda ji bude možno nahradit relativně jednoduššími postupy.

I při tomto poněkud zdrženlivém postoji se nemohu ubránit dojmu, že Lazarsfeldova technika patří k nejdůležitějším teoriím škálování, které vznikly v posledních 30 letech. Jsem k tomu veden třemi důvody: a/ Jedná se o postup velmi obecný a proto aplikovatelný na velkou šíři problémů. b/ Jedná se o postup, který je přiměřený pro rozbor multidimenzionálních jevů, s nimiž se v psychologii velmi často setkáváme. c/Umožňuje tato data beze zbytku zpracovat, aniž by bylo třeba redukovat je na menší počet dimenzí či aspektů. Scheuch /1962/ se dokonce domnívá, že řadu jiných škálovacích technik můžeme považovat za speciální případy analýzy latentní struktury. Z uvedených důvodů považuji informaci o této technice za potřebnou. Bylo by jistě velmi žádoucí, kdybychom získali přímé zkušenosti, které u nás zatím chybí.

Kapitola XI.

COOMBSOVA TEORIE DAT

XI.1. T ř í d ě n í d a t

C.H. Coombs, přední představitel matematické psychologie, se zabýval v řadě studií /1950, 1953, 1960 aj./ obecnou charakteristikou dat, která se v psychologii vyskytují. V rozsáhlé monografii /1964/ pak tyto úvahy rozvinul a navrhl řadu nových škálovacích postupů. Coombsův systém je velmi obecný a řadu metod, které jsme probírali v předchozích kapitolách, je možno do něj zařadit. Při výkladu se soustředím na základní škálovací metodu, kterou Coombs navrhl a detailně propracoval. Dříve je však na místě uvést stručně jeho výchozí úvahy o teorii dat a měření.

Coombs se ve své teorii snaží abstrahovat od obsahové stránky jednotlivých metod analýzy psychologických dat. Vytváří modely pro charakterizování psychologických měření; klasifikuje a uvádí je do vzájemných vztahů. Z hlediska teorie měření člověk srovnává podněty, a to buď mezi sebou navzájem či vzhledem k absolutnímu standardu či k osobnímu vztahovému bodu /většinou subjektivní optimum/. Podněty mohou být buď samostatné nebo se vyskytují v párech. Mohou se srovnávat buď co do vzájemné shody /či podoby/ nebo z hlediska dominace /preference/. Toto vyjádřil Coombs takto /IV. axiom teorie dat/:

Všechna psychologická data existují ve třech dichotomiích:

I. Existuje vztah mezi párem bodů /dyáda/ nebo párem dyád.

II. Srovnávané prvky /body či jejich dyády/ lze vzíti buď ze dvou rozličných množin či z jedné množiny.

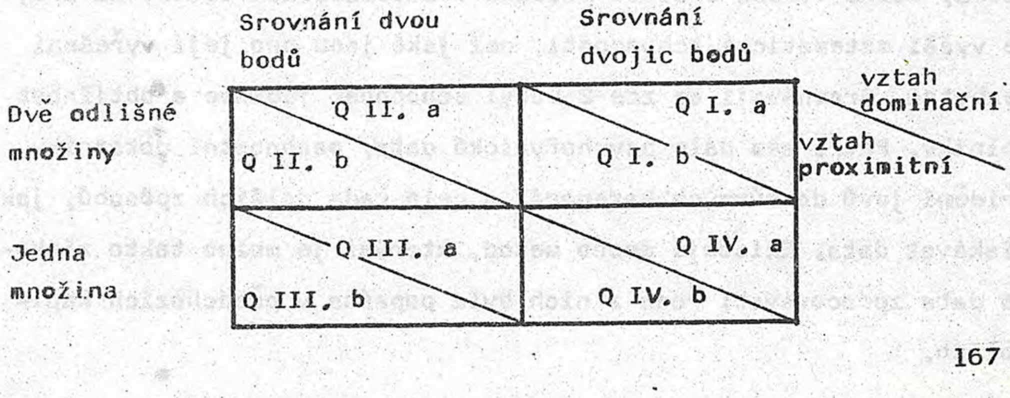
III. Vztah mezi body je buď preferenční a dominantní /A je úzkostnější než B/ nebo proximitní /jak mnoho jsou si podněty blízké či podobné - A je stejně úzkostný jako B/.

Základem každého psychologického modelu měření je asociovat s každým objektem zájmu - ať se jedná o podnět či o osobu - určitý bod v psychologickém prostoru. Účelem modelu pak je najít takový postup, který by umožnil rekonstruovat psychologický prostor, pakliže jsou nám známy výsledky pozorování a předběžné koncepce uvažovaného prostoru. Předběžných koncepcí psychologického prostoru může být mnoho a proto teorie měření musí být značně obecná.

Ze 3 dichotomií, které jsou formulovány ve IV. axiomu, vyplývá 8 tříd dat, které lze znázornit krychlí o rozměrech 2 x 2 x 2. Systém dat můžeme schematicky znázornit na obr. XI.1.

Obr. XI.1.

Coombsův systém psychologických dat



Stručně naznačím, jaká psychologická data odpovídají jednotlivým třídám. Výklad je velmi zjednodušený a pouze orientační; pro podrobné studium je třeba vyjít přímo z původních Coombsových prací.

Do kategorie Q I. a patří takové postupy, ve kterých jsou pokusné osoby tázány, kterou ze 2 alternativ /A, B/ by preferovaly. Přitom podněty A, B jsou vzaty z různých množin. Předpokládáme, že existují individuální "ideální" body pro obě množiny /I_A, I_B/. Člověk volí zřejmě tu alternativu, která je "blíže" k odpovídajícímu ideálnímu bodu. Tak vlastně srovnáváme 2 páry bodů /vzdálenost A, I_A se vzdáleností B, I_B/. Pro zpracování těchto dat se nejlépe hodí Coombsova technika rozvinutí, kterou popíši dále; v případě multidimenzionální analýzy lze užít i některé formy faktorové analýzy.

V oktantu Q I. b jsou data analogická; existují dvě alternativy a osoba udává, zda je či není schopna mezi nimi preferovat. Data v této podobě se prakticky vyskytují jen velmi řídce a proto zde nevznikly žádné specifické škálovací metody.

Data, která se zahrnují do oktantu Q II. a,, jsou v psychologii nejčastější. Patří sem všechny didaktické či osobnostní testy; osoba, která vyřeší určitou položku v matematickém testu, má zřejmě vyšší matematické schopnosti, než jaké jsou pro její vyřešení nezbytné. Srovnávají se zde 2 body: schopnost jedince a obtížnost položky. Patří sem dále psychofyzická data, osobnostní dotazníky, třídění jevů do různých kategorií a celá řada dalších způsobů, jak získávat data. Existuje mnoho metod, kterými je možno takto získaná data zpracovávat; řada z nich byla popsána v předchozích kapitolách.

Do kategorie Q II. b zařazuje Coombs data, která se získávají z pokusů, v nichž osoby uvádějí, zda s určitým tvrzením souhlasí či nikoliv. Srovnává se zde vlastně pár bodů; jeden je vlastní názor osoby a druhý je postoj, vyjádřený v posuzovaném výroku. Přitom oba body patří zřejmě k odlišným množinám. Klinická diagnóza, volba partnerů, příslušnost do určitých organizací, a do jisté míry i posuzovací škály - to vše jsou příklady dat tohoto typu. Pro jejich zpracování existuje na jedné straně řada poměrně jednoduchých metod /např. metoda průměrné chyby, známá z psychofyziky/, naproti tomu lze však zde užít i poměrně složité techniky, jako je Lazarsfeldova analýza latentní struktury, a to hlavně tam, kde nemáme žádné předběžné znalosti či hypotézy o vztazích mezi jednotlivými prvky uvnitř obou tříd.

Poměrně častá jsou i data, která můžeme zařadit do kategorie Q III. a . Srovnáváme dva podněty z jedné množiny a zjišťujeme, který z nich má větší hodnotu v určité vlastnosti či dimenzi /světlejší barva, pilnější žák atd./. Srovnáváním obou podnětů docházíme k závěru o jejich vztahu. Jako příklady podobných dat můžeme uvést psychofyzické výzkumy, škálování sportovních mužstev z jejich výsledků v turnaji atd. Protože podněty tvoří body z jedné množiny, je výsledkem jednodimenzionální škála, která je velmi často intervalového či dokonce poměrového typu. Intervalové škály se mohou vytvářet např. pomocí Thurstonových metodik; poměrovou škálu lze od- odvodit z Lucova modelu výběrového jednání /Luce, 1959/.

Analogicky jako v předchozích, tak i ve třetím kvadrantu se data typu Q III. a liší od dat Q III. b pouze tím, že se zjišťuje, zda podněty jsou stejné či nikoliv. Podobná data se vyskytovala též v psychofyzice, i když méně často. Úkolem pokusné osoby je najít mezi řadou podnětů takový, který by byl nejblíže k podnětu standardnímu, případně se sleduje, jak často je který podnět s jinými zaměňován. Postupně se vytvoří prostor vztahů mezi různými podněty. Mezi daty, která spadají do III. kvadrantu, jsou i sociometrické údaje. Zjišťujeme-li sociometricky, kdo je komu nadřízen, kdo koho ovládá apod., pak jde o data typu Q III. a ; chceme-li znát, kdo si koho zvolil za společníka, s kým by šel do kina apod., pak máme data typu Q III. b.

Do oktantu Q IV. a patří data toho typu, kdy člověk porovnává dva páry podnětů a určuje, který pár si je podobnější. Výsledné údaje mohou za určitých podmínek vést k jednodimenzionální metrické škále, ale mohou vést i k multidimenzionálnímu řešení. Data typu Q IV. b se vyskytují řídce a opírají se o situace, kdy zjišťujeme, zda člověk je s to zachytit rozdíly mezi dvěma páry podnětů či nikoliv.

Třetí používanou dichotomii /vztah dominance a blízkosti/ můžeme považovat za sekundární a redukovat proto základní psychologická data na čtyři základní typy: preferenční výběr /I. kvadrant/, posouzení jediného podnětu /II. kvadrant/, srovnávání podnětů /III. kvadrant/ a podobnostní data /IV. kvadrant/.

Mezi daty z uvedených kategorií je řada vztahů a v některých případech je jejich zařazení opřeno jen o velmi jemný rozdíl. Existuje též celá řada jiných typů dat, jako je např. latentní doba reakce, velikost /amplituda, síla.../ reakce, nepravidelnost odpovědí, chyby atd. I tato data je možno interpretovat jako vztahy mezi body či páry bodů a zatřídit je po určitých transformacích do uvedených modelů psychologického měření. Tak třeba čím horší je rozlišitelnost dvou podnětů, která vede často k jejich záměně, tím déle trvá i výběr mezi nimi, a tím menší je zřejmě též jejich vzdálenost v psychologickém prostoru /většinou lze uvedená data transformovat do podoby dat kategorie Q III. a/.

Pro studium škálovacích metod je však důležité to, že Coombs vyvinul řadu nových metod, popsaných hlavně v monografii z r. 1964. Není možné je všechny probírat; některé jsou poměrně abstraktní a je k dispozici jen málo výzkumů, které by umožnily posoudit jejich užitečnost a nosnost pro potřeby psychologie. Uvedu v další části základní Coombsovu metodu, která je dobře propracována, ve výzkumné praxi ověřena a s úspěchem užita.

XI. 2. Metoda rozvinutí dat v jedné dimenzi

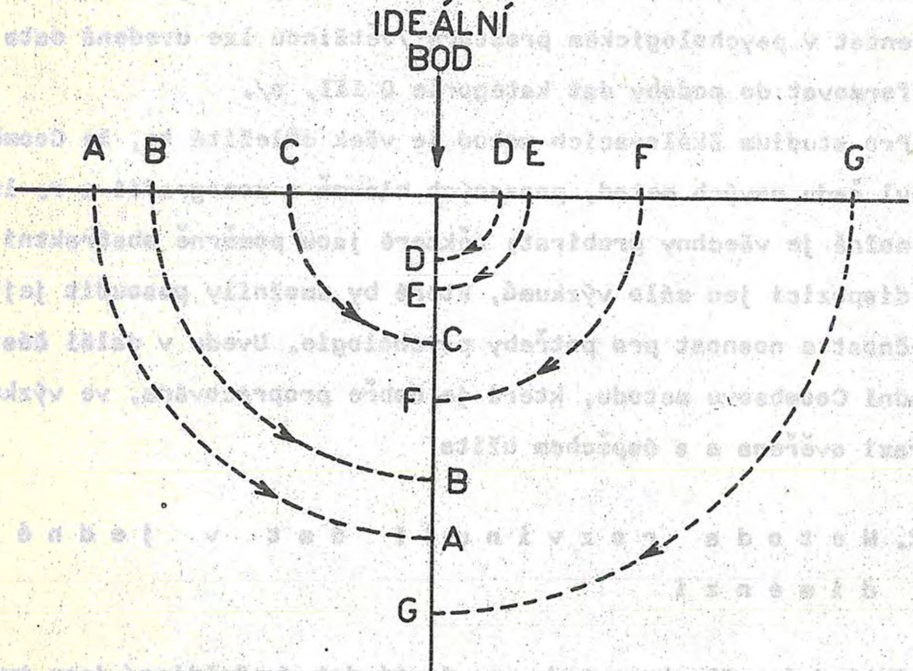
Základní myšlenky metody rozvinutí dat /unfolding/ jsou tyto: každý podnět je reprezentován bodem na jedné společné dimenzi, kterou Coombs nazývá J-škála. Při srovnávání jednotlivých podnětů získáme jejich pořadí. Projeví se v něm absolutní vzdálenost jednotlivých podnětů od ideálního /optimálního/ bodu.

Individuální systémy preferencí se vyjadřují I-škálami, které vzniknou jakýmsi přeložením J-škály kolem ideálního bodu jednotlivých osob /viz zjednodušené grafické vyjádření na obr. XI.2/.

OBR. XI. 2.

I-ŠKÁLA VZNIKÁ PŘELOŽENÍM J-ŠKÁLY

(ZJIŠTĚNÉ POŘADÍ: D,E,C,F,B,A,G)



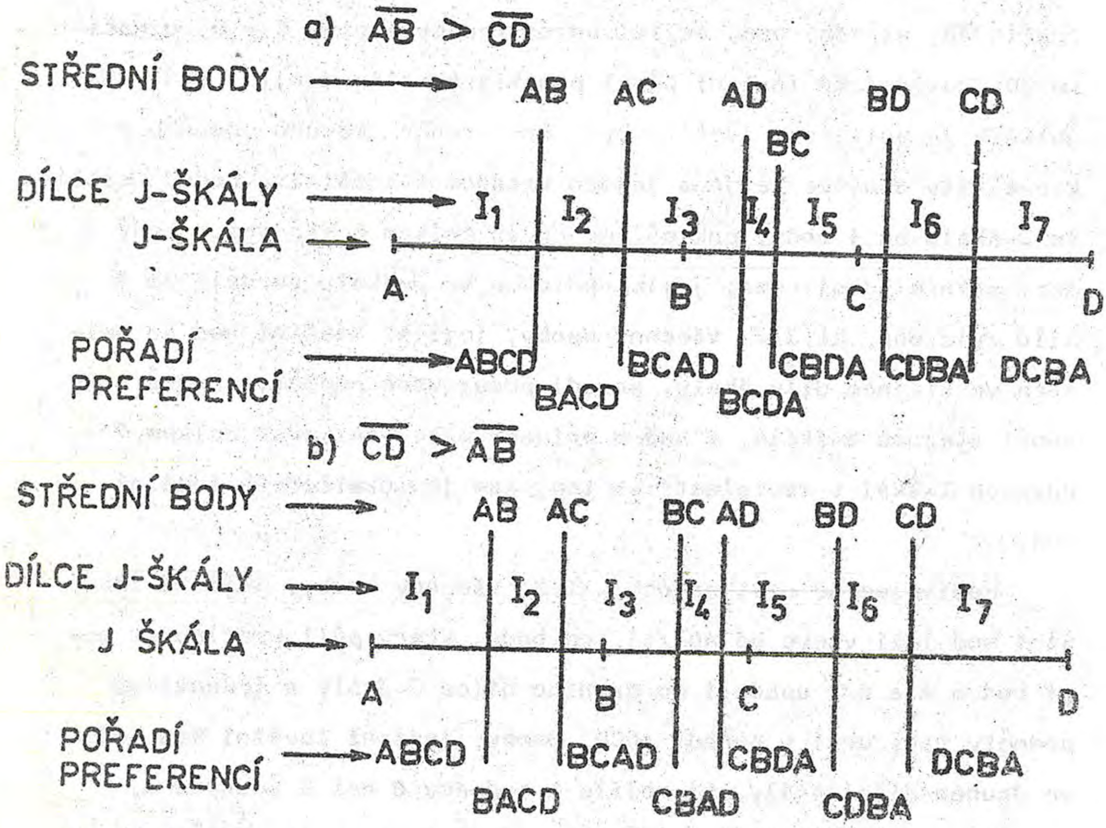
Při výzkumu máme zachycenou řadu individuálních I-škál a snažíme se zpětně rekonstruovat J-škálu. V principu tedy předpokládáme, že existuje jednodimenzionální kontinuum /čili jeden latentní parametr/, které určuje systém preferencí mezi danými podněty; v tomto smyslu jde o jistou analogii s Guttmanovým škálogramem. Coombs svůj model rozpracoval i pro vícedimenzionální škály; pro účel výkladu však budeme předpokládat pouze jedinou dimenzi.

Coombsovy základní úvahy, z kterých je odvozena metoda přeložení dat, si předvedeme na následujícím příkladě: Máme 4 podněty A, B, C, D, které leží na J-škále; přitom jejich rozdíl /vzdálenost jejich polohy na škále/ je dostatečně velký, takže se jedná o podněty dobře rozlišitelné. Vzdálenost mezi bodem A a B budeme značit $\overline{AB}$; střední bod, ležící uprostřed mezi body C a D, označíme CD; analogické značení platí pro kteroukoliv dvojici bodů. Na J-škále je umístěn i ideální bod. Preference jednoho podnětu z kterékoliv dvojice je dána jejich vztahem k ideálnímu bodu. Jestliže J-škála má 4 body, pak můžeme určit celkem 6 středních bodů mezi možnými dvojicemi; jejich polohou se J-škála rozdělí do 7 dílů /viz obr. XI.3./. Všechny osoby, jejichž ideální bod je umístěn ve stejném dílu škály, seřadí posuzované podněty shodně a vytvoří stejnou I-škálu. V našem případě může vzniknout celkem 7 různých I-škál v závislosti na tom, kam je lokalizován ideální bod.

Podívejme se opět na obr. XI.3. Všechny osoby, jejichž ideální bod leží vlevo od AB /tj. od bodu, který půlí vzdálenost mezi bodem A a B/, spadají do prvního dílce J-škály a jednotlivé podněty tedy určí v pořadí ABCD. Osoby, jejichž ideální bod leží ve druhém dílci škály, tj. blíže k podnětu B než k podnětu A, zatímco body C i D leží ještě dále, musí preferovat podněty v pořadí BACD. Ve třetím dílci jsou osoby, které mají ideální bod nejblíže k podnětu B, a dále je blíže k podnětu C než k podnětu A; jejich preferenční systém je tedy BCAD. Ve čtvrtém dílci se však mohou objevit dvě různá pořadí, a to buď BCDA, anebo CBAD.

OBR. XI. 3.

DVĚ MOŽNÉ J-ŠKÁLY PŘI 4 PODNĚTECH.



První pořadí vznikne tehdy, když $\overline{AB} > \overline{CD}$ /tj. vzdálenost mezi podněty A, B je větší než mezi podněty C, D/, zatímco druhá varianta je důsledkem vztahu opačného: $\overline{CD} > \overline{AB}$. Pořadí v dalších dílcích J-škály jsou opět jednoznačná.

Jestliže známe J-škálu, pak snadno stanovíme množinu I-škál, které jí odpovídají. Ve většině případů však stojíme před obráceným úkolem; známe I-škály, které odpovídají výsledkům jednotlivých osob, a chceme konstruovat J-škálu. Přitom většinou ani nevíme, zda vůbec pro daná data existuje jednodimenzionální řešení. Vytvoření J-škály probíhá v několika fázích.

Prvním krokem je určení pořadí podnětů na J-škále. Z obr. XI.3. je zřejmé, že I-škály mohou končit pouze dvěma podněty. Bude to buď první nebo poslední podnět v pořadí z J-škály, kterou chceme vytvořit. Žádný jiný podnět se na posledním místě objevit nemůže. Pro názornost se vraťme k obr. XI. 3.; snadno si představíme, že ať je ideální bod kdekoliv na J-škále, pak při "překlopení" škály kolem tohoto bodu musí I-škála končit buď prvním nebo posledním podnětem, tj. buď bodem A nebo bodem D. Dále snadno zjistíme, že existují pouze dvě škály, které jedním z těchto podnětů začínají a druhým končí; tyto škály jsou zrcadlově obráceny. Lze též dokázat, že žádné dvě další I-škály nemají tuto vlastnost; kdyby se přece vyskytly, znamenalo by to, že předpoklad jednodimenzionální škály není správný. Kterákoliv z obou zrcadlově položených škál zachycuje pořadí podnětů na J-škále; rozdíl je pouze v tom, že jedna škála je uspořádána vzestupně a druhá sestupně /v našem příkladě je situace snad zjednodušena tím, že podněty jsme označili prvními písmeny abecedy/. Takto sestavenou škálu nazývá Coombs kvalitativní J-škálou.

Druhý krok míří k vytvoření kvantitativní škály, která by byla škálou s metrickými vlastnostmi /tedy alespoň intervalovou/.

Zde bude postup již složitější.

Z obr. XI.3. je vidět, že první tři I-škály vyplývají jednoznačně z J-škály. Pouze prostřední škála I_4 z ní nevyplývá jednoznačně, ale má dvě možné podoby, a to buď BCDA, anebo CBAD. Z toho, která podoba škály I_4 se objeví, vyplývá nová informace metrického charakteru: buď platí $\overline{AB} > \overline{CD}$ /pokud I_4-škála byla BCDA/, anebo naopak $\overline{CD} > \overline{AB}$ /pokud bylo zjištěno pořadí CBAD/.

Coombs zavádí následující pravidlo: jestliže zjistíme, že střední bod mezi B a C předchází na J-škále před středním bodem mezi A a D, pak z toho plyne, že vzdálenost mezi C a D je větší než mezi A a B, což symbolicky vyjádříme takto: BC, AD $\rightarrow \overline{CD} > \overline{AB}$. Naopak platí: AD, BC $\rightarrow \overline{AB} > \overline{CD}$.

Více metrických informací při 4 podnětech nemůžeme získat. Mnohem složitější situace vzniká při větším počtu podnětů. Budeme uvažovat o 7 podnětech /A - G/, které jsou řazeny dle preferencí. Objeví se celkem $\binom{7}{2} + 1 = 22$ možných I-škál. Dále platí, že všechny I-škály končí buď podnětem A nebo G, že existuje jedna škála začínající podnětem A a končící podnětem G a zároveň druhá škála, která je jejím zrcadlovým obrazem, a konečně ostatní I-škály jsou sestaveny tak, že jeden ze sousedních párů podnětů se objeví v následující škále v opačném pořadí; na pořadí ostatních prvků se přitom nic nemění. Jedna z možných množin I-škál, která odpovídá těmto požadavkům, je zachycena v tab. XI.1. Bylo by možno sestavit celou řadu podobných množin. Jejich počet je možno přesně určit; o tom se zmíním později.

V prvním sloupci tab. XI.1. jsou čísla jednotlivých škál. Ta-

Tab. XI.1.

Množina I-škál při 7 podnětech

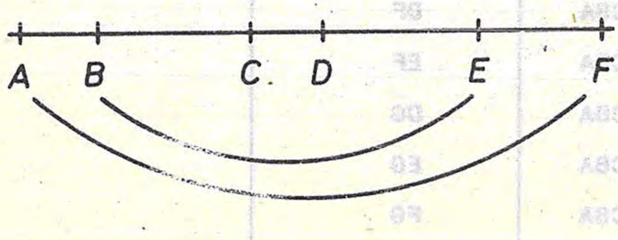
Číslo I-škály	I-škála	Spodní hranice daného dílu J-škály	Metrická informace
1	ABCDEFG		
2	BACDEFG	AB	
3	BCADEFG	AC	
4	CBADEFG	BC	BC, AD $\rightarrow \overline{CD} > \overline{AB}$
5	CBDAEFG	AD	
6	CBDEAFG	AE	AE, BD $\rightarrow \overline{AB} > \overline{DE}$
7	CDBEAFG	BD	BD, AF $\rightarrow \overline{DF} > \overline{AB}$
8	CDBEFAG	AF	
9	DCBEFAG	CD	CD, BE $\rightarrow \overline{DE} > \overline{BC}$
10	DCEBFAG	BE	CD, AG $\rightarrow \overline{DG} > \overline{AC}$
11	DCEFBAG	BF	BF, CE $\rightarrow \overline{BC} > \overline{EF}$
12	DECFBAG	CE	BF, AG $\rightarrow \overline{FG} > \overline{AB}$
13	DEFCBAG	CF	CF, DE $\rightarrow \overline{CD} > \overline{EF}$
14	DEFCBGA	AG	CF, BG $\rightarrow \overline{FG} > \overline{BC}$
15	EDFCBGA	DE	AG, DE $\rightarrow \overline{AD} > \overline{EG}$
16	EDFCGBA	BG	DE, CG $\rightarrow \overline{EG} > \overline{CD}$
17	EDFGCBA	CG	CG, DF $\rightarrow \overline{CD} > \overline{FG}$
18	EFDGCBA	DF	
19	FEDGCBA	EF	
20	FEGDCBA	DG	
21	FGEDCBA	EG	
22	GFEDCBA	FG	

to čísla zároveň pořadově určují jednotlivé díly J-škály tak, jak jsou postupně ohraničeny středními body, které dělí vzdálenosti mezi jednotlivými páry daných podnětů /viz též ve zjednodušené podobě obr. XI.3/. Každému dílu J-škály odpovídá jiné pořadí preference podnětů, které je zachyceno v druhém sloupci tabulky. Střední bod, který tvoří spodní hranici jednotlivých dílů J-škály, je uveden ve třetím sloupci. V posledním sloupci jsou zachyceny metrické informace, vyplývající z pořadí, v jakém jsou jednotlivé podněty mezi sebou preferovány, tj. z pořadí, které je zachyceno v druhém sloupci tabulky.

Jde nyní o to, jak získáme tyto metrické informace. Přinášejí je pouze takové vztahy mezi středními body, které nelze vyčíst z kvalitativní J-škály. Jestliže konečné body určité části J-škály je možno "obalit" konečnými body jiné části J-škály /tj. leží uvnitř jiné, širší části škály/, pak pořadí středních bodů těchto dvou částí škály přináší metrickou informaci. Podívejme se na obr. XI.4., na kterém bude postup i celá úvaha zřetelnější.

OBR. XI. 4.

OBALOVÁNÍ PÁRŮ PODNĚTŮ



Část škály ohraničená body C, D je "obalena" částí B, E a tato část je opět "obalena" částí A, F. Pořadí středních bodů AF, BE, CD obsahuje metrickou informaci. Naproti tomu střední body dílů BE, CF představují dvě části škály, které se vzájemně překrývají; jejich srovnání je jednoznačně dáno polohou jejich konečných bodů vzhledem k ideálnímu bodu a nepřináší žádnou novou metrickou informaci.

Není však nezbytné sledovat všechny dvojice středních bodů těch párů podnětů, které se "obalují". Zjistíme-li např., že BD /tj. střední bod mezi podnětem B a D/ předchází na J-škále před AE, pak zřejmě i BC musí předcházet před AE /protože BC leží na škále nutně před BD/.

Sledujme nyní postup, jak byla sestavena tab. XI.1. Předpokládejme, že se jedná o 7 podnětů, které mají být seřazeny dle pořadí preference. Osoby, které je seřadily do pořadí ABCDEFG, vytvořily I-škálu č.1. Spodní hranici prvního intervalu nelze určit; známe pouze jeho horní hranici, která je na J-škále tvořena středním bodem mezi podněty A a B. Tento bod je zároveň spodním bodem I-škály č.2 - s pořadím BACDEFG /viz též analogie na obr. XI.3./. I-škála č.3 má spodní bod uprostřed intervalu AC, pořadí podnětů je BCADEFG. Až sem je postup zcela jednoznačný a nemůžeme získat žádné metrické informace. První dvojznačná situace vzniká u I-škály č. 4, která může být buď CBADEFG /a pak spodní hranice je BC/, anebo BCDAEFG /se spodním bodem AD/. Protože v našem příkladě se objevila první varianta, plyne z toho, že vzdálenost $\overline{CD}$ je větší než $\overline{AB}$ /viz též znovu obr. XI.3/. Tato skutečnost je zaznamenána v rubrice metrické informace v tab. XI.1.

Takto postupujeme dále celou tabulkou a odvodzujeme všechny dostupné informace o metrických vztazích mezi podněty. Pro usnadnění postupu uvedu některé návody, jak postupovat, aniž bych je zdůvodňoval. V první řadě si všimněme, že v systému I-škál se každá následující škála liší od předchozí pouze přehozením prvků v jedné dvojici sousedních podnětů; tyto přehozené podněty tvoří zároveň spodní hranici daného dílu J-škály.

Poněkud obtížnější je odvození metrických informací. Coombs doporučuje následující postup: Sestavme si nejprve pořadí spodních hranic jednotlivých dílů škály /sloupec č.3 v tab. XI.1./. Postupně budeme srovnávat každý bod se všemi následujícími, pokud jejich vztah neplyne automaticky z celkového pořadí. Je třeba zřejmé, že musí předcházet AB před AC, AD apod., stejně tak AB před BC či EF atd., ale z toho nelze odvodit žádnou metrickou informaci. Pro racionální postup je účelné si označit na sledu spodních intervalů jednotlivých dílů ty body, ve kterých je přirozený sled podnětů přerušen. Tento bod srovnáme se všemi ostatními body, kterými opět jiné přirozené sledy začínají. Přepíšeme nejprve sled bodů, jak ohraničují díly na J-škále /tj. 3. sloupec z tab. XI.1./ a pomlčkou označíme přerušení přirozeného sledu;

AB, AC, BC - AD, AE - BD - AF - CD - BE, BF - CE, CF - AG - DE - - BG, CG - DF, EF - DG, EG, FG.

Konec prvního přirozeného sledu je BC; srovnáme jej se středním bodem AD; s ostatními body jej srovnávat nemusíme buď proto, že vyplývají přímo z polohy na škále, anebo je můžeme bez nesnází odvodit /jestliže $\overline{CD} > \overline{AB}$, pak též $\overline{CE} > \overline{AB}$, $\overline{CF} > \overline{AB}$ atd./.

Další pravidlo pro odvození metrických vztahů stanovil Coombs takto: při srovnání polohy dvou středních bodů zapíšeme nejprve jejich sled tak, jak vyplývá z jejich pořadí. V našem příkladě se poprvé vyskytnou střední body BC, AD. Jestliže nyní k sobě přiřadíme obě první písmena z každého páru /BA/ a obě druhá písmena /CD/, vzniknou dva nové páry písmen, z nichž jeden je v abecedním pořadí a druhý nikoliv; pár, který je v abecedním pořádku, reprezentuje vždy větší vzdálenost. Proto platí $\overline{CD} > \overline{AB}$. Kdybychom měli však obrácené pořadí výchozích středních bodů AD, BC, pak by z toho plynulo $\overline{AB} > \overline{CD}$ /viz znovu též obr. XI.3./.

Všechny dostupné metrické informace jsou znázorněny ve 4. sloupci tab. XI.1. Všimněme si ještě jednou postupu: První možná metrická informace plynula ze srovnání BC, AD. Další získáváme ze srovnání AE, BD /AE s jiným bodem srovnávat nemusíme/; pak srovnáme BD s AF /srovnávat BD s AG by už bylo nadbytečné/. AF není třeba srovnávat s ničím /srovnání s CD je zbytečné, protože již dříve bylo srovnání s BD/. Bod CD však srovnáme dvakrát /s BE a s AG; oběma je totiž na škále "obalen"/ atd.

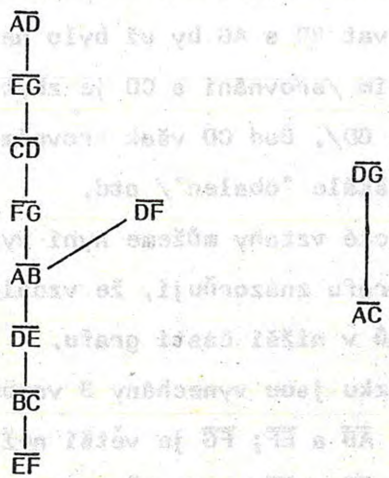
Získané metrické vztahy můžeme nyní vyjádřit graficky na obr. XI.5. Vyšší body grafu znázorňují, že vzdálenost uvedených bodů je větší než u bodů v nižší části grafu.

Na tomto obrázku jsou vynechány 3 vazby, které jsou nadbytečné /$\overline{CD}$ je větší než $\overline{AB}$ a $\overline{EF}$; $\overline{FG}$ je větší než $\overline{BC}$/. Získali jsme jeden řetězec vztahů /od $\overline{AD}$ k $\overline{EF}$/ a kromě toho ještě izolovaný vztah $\overline{DG} > \overline{AC}$. Můžeme však stanovit ještě další metrické vztahy. Zatím známe vztahy mezi všemi dvojicemi sousedních bodů. Můžeme srovnávat vzdálenost mezi body, které jsou od sebe vzdáleny ob jeden bod /tedy $\overline{AC}$, $\overline{BD}$, $\overline{CE}$, $\overline{DF}$, $\overline{EG}$/. Jejich vztahy odvodíme z již známých vazeb. Např. jestliže $\overline{CD} > \overline{AB}$, pak také musí být $\overline{BD} > \overline{AC}$, anebo ze vztahů $\overline{CD} > \overline{AB}$ a $\overline{DE} > \overline{BC}$ plynou nutně $\overline{CE} > \overline{AC}$ atd. Po zpracování všech údajů tohoto druhu odvodíme následující řetězec: $\overline{CE} > \overline{BD} > \overline{EG} > \overline{AC} > \overline{DF}$.

Analogicky odvodíme i další vztahy: $\overline{AD} > \overline{BE} > \overline{CF} > \overline{DG}$; dále $\overline{CG} > \overline{AE} > \overline{BF}$ a konečně též $\overline{BG} > \overline{AF}$ /což zřejmě plyne ze vztahu $\overline{FG} > \overline{AB}$ /. Vzdálenost $\overline{AG}$ je nutně největší. Po srovnání všech metrických údajů, které jsme získali, vytvoříme úplný metrický vztah. Znázorníme jej na obr. XI.6.

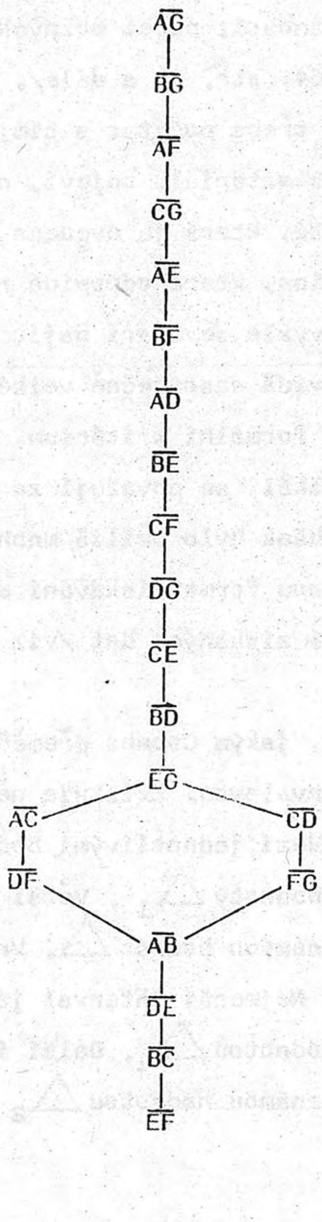
Obr. XI.5.

Metrické vztahy plynoucí z tab. XI.1.



Obr. XI.6.

Doplněné metrické vztahy z tab. XI.1.



Je však třeba upozornit, že uvedený příklad je značně idealizovaný. Soustava I-škál, z které jsme vycházeli, je pouze jedna z možných soustav. Teoreticky by jich bylo možno vytvořit při sedmi podnětech ohromné množství. Kvalitativních J-škál může být n!/2, při n = 7 to znamená 2.520 možností; počet možných škál je mnohonásobně větší /viz Coombs, 1964, str. 91 a dále/.

Při reálných pokusech je třeba počítat s tím, že všechny I-škály, které se v empirickém materiálu objeví, nelze zařadit do jediné množiny I-škál /třeba té, která je uvedena v tab. XI.1./. Vyhledává se vždy taková množina, které odpovídá relativně největší množství získaných dat. Obvykle se zdaří najít dominující kvantitativní J-škálu, která odpovídá dostatečně velkému počtu případů; Coombs však neuvádí žádné formální kritérium. Data, která neodpovídají zvolené množině I-škál, se považují za projev menší reliability výsledků. Pokud by bšak bylo příliš mnoho odchylek, doporučuje Coombs buď zvolit jinou formu získávání materiálu, anebo poněkud jiný způsob zpracování získaných dat /viz Coombs, Pruitt, 1960/.

Sledujme nyní další krok, jakým Coombs přeměňuje získanou pořadovou škálu na škálu intervalovou. Existuje několik postupů; popíši zde jen jeden z nich. Mezi jednotlivými body na škále existuje jakýsi interval neznámé hodnoty Δ_1 . Větší vzdálenosti vyjádříme zavedením dalších neznámých hodnot Δ. Vraťme se k obr. XI.6. a sledujme celou úvahu. Nejmenší interval je EF; označíme jej pozitivní, ale neznámou hodnotou Δ_1. Další interval v pořadí je BC, který je o jinou neznámou hodnotou Δ_2 větší než EF;

proto $\overline{BC} = \triangle_1 + \triangle_2$; analogicky $\overline{DE} = \triangle_1 + \triangle_2 + \triangle_3$ a $\overline{AB} = \triangle_1 + \triangle_2 + \triangle_3 + \triangle_4$.

Nyní však další pořadí není jednoznačné. Intervaly $\overline{DF}$ a $\overline{FG}$ jsou větší než $\overline{AB}$, ale jejich vztah nemůžeme přímo určit. Začněme intervalem $\overline{DF}$. Platí:

$\overline{DF} = \overline{DE} + \overline{EF} = /\triangle_1 + \triangle_2 + \triangle_3/ + \triangle_1 = 2\triangle_1 + \triangle_2 + \triangle_3$

Protože platí $\overline{DF} > \overline{AB}$, pak také

$$2\triangle_1 + \triangle_2 + \triangle_3 > \triangle_1 + \triangle_2 + \triangle_3 + \triangle_4$$

a proto $\triangle_1 > \triangle_4$

Budeme proto dále uvádět:

$$\triangle_1 = \triangle_4 + \triangle_5$$

a tento výraz zavedeme do všech vztahů, které jsme zatím odvodili. Bude tedy

$$\overline{DF} = \triangle_2 + \triangle_3 + 2\triangle_4 + 2\triangle_5$$

Sledujme nyní interval $\overline{AC}$:

$\overline{AC} = \overline{AB} + \overline{BC} = /\triangle_2 + \triangle_3 + 2\triangle_4 + \triangle_5/ + /\triangle_2 + \triangle_4 + \triangle_5/ = 2\triangle_2 + \triangle_3 + 3\triangle_4 + 2\triangle_5$

Víme též, že $\overline{EG} > \overline{AC}$, a proto k hodnotě $\overline{AC}$ přidáme další neznámou hodnotu $\triangle_6$; tedy

$$\overline{EG} = 2\triangle_2 + \triangle_3 + 3\triangle_4 + 2\triangle_5 + \triangle_6$$

Podívejme se nyní na druhou větev pořadových vztahů:

$\overline{FG} < \overline{CD} < \overline{EG}$. Protože $\overline{EG} = \overline{EF} + \overline{FG}$, můžeme interval $\overline{FG}$ stanovit jako

$\overline{FG} = \overline{EG} - \overline{EF} = /2\triangle_2 + \triangle_3 + 3\triangle_4 + 2\triangle_5 + \triangle_6/ - /\triangle_4 + \triangle_5/ = 2\triangle_2 + 2\triangle_3 + 2\triangle_4 + \triangle_5 + \triangle_6$

Dále víme, že $\overline{CD} > \overline{FG}$; zavedeme proto další hodnotu $\triangle_7$:

$$\overline{CD} = \overline{FG} + \triangle_7$$

a protože $\overline{EG} > \overline{CD}$, pak

$$2\triangle_2 + \triangle_3 + 3\triangle_4 + 2\triangle_5 + \triangle_6 > 2\triangle_2 + \triangle_3 + 2\triangle_4 + \triangle_5 + \triangle_6 + \triangle_7$$

a tedy

$$\triangle_4 + \triangle_5 > \triangle_7$$

Zavedeme

$$\triangle_4 + \triangle_5 = \triangle_7 + \triangle_8$$

a vyloučíme tak $\triangle_5$ ze všech zjištěných vztahů.

Nyní shrneme zatím zjištěné vztahy mezi sousedními podněty na škále.

$$\begin{array}{lcl}
\overline{AB} & = & \triangle_2 + \triangle_3 + \triangle_4 + \triangle_7 + \triangle_8 \\
\overline{BC} & = & \triangle_2 + \triangle_7 + \triangle_8 \\
\overline{CD} & = & 2\triangle_2 + \triangle_3 + \triangle_4 + \triangle_6 + 2\triangle_7 + \triangle_8 \\
\overline{DE} & = & \triangle_2 + \triangle_3 + \triangle_7 + \triangle_8 \\
\overline{EF} & = & \triangle_7 + \triangle_8 \\
\overline{FG} & = & 2\triangle_2 + \triangle_3 + \triangle_4 + \triangle_6 + \triangle_7 + \triangle_8
\end{array}$$

Další vztahy bychom mohli získat již mechanickým sčítáním, např.:

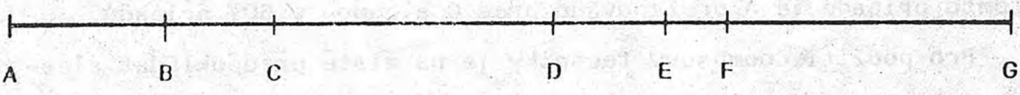
$$\overline{BD} = \overline{BC} + \overline{CD} = /\triangle_2 + \triangle_7 + \triangle_8/ + /2\triangle_2 + \triangle_3 + \triangle_4 + \triangle_6 + 2\triangle_7 + \triangle_8/ = 3\triangle_2 + \triangle_3 + \triangle_4 + \triangle_6 + 3\triangle_7 + 2\triangle_8$$

anebo

$$\overline{AG} = \overline{AB} + \overline{BC} + \overline{CD} + \overline{DE} + \overline{EF} + \overline{FG} = 7\triangle_2 + 4\triangle_3 + 3\triangle_4 + 2\triangle_6 + 7\triangle_7 + 6\triangle_8$$

Dané metrické vztahy je možno vyjádřit mnoha způsoby. Jedním z nich je předpoklad "stejných $\triangle$" /případně bychom mohli též říci průměrných $\triangle$/; každé hodnotě $\triangle$ přisoudíme hodnotu 1 , a pak $\overline{AB} = 5$, $\overline{BC} = 3$, $\overline{CD} = 8$, $\overline{DE} = 4$, $\overline{EF} = 2$, $\overline{FG} = 7$

Jestliže podnět A označíme na škále hodnotou 0 , pak hodnota podnětu B = $\overline{AB}$ = 5 atd. Výsledná škála tedy má tuto podobu:



Existuje řada dalších způsobů, jak vytvořit ze získaných dat intervalovou škálu. Mnohé z nich předpokládají využití počítačů. Coombs však srovnáváním zjistil, že vedou prakticky ke stejným výsledkům, a proto uvedený postup, který je relativně nejjednodušší, můžeme považovat za zcela přiměřený.

K informaci o škálovací technice rozvinutí je třeba připojit ještě Coombsovy úvahy o častých nepravidelnostech, ke kterým při jeho způsobu škálování dochází. Již ve III. kapitole jsme se zabývali nepravidelnými preferencemi, které při párovém srovnávání vedou k vytváření kruhových triád. Coombs rozeznává 3 druhy stochastické tranzitivity, které vedou k nepravidelnému kolísání výsledků:

a/ silná stochastická tranzitivita /SST/, o které platí:

$p/AB/ \geq 0,5$, $p/BC/ \geq 0,5 \longrightarrow p/AC/ \geq \text{maximum} \left[p/AB/ , p/BC/ \right]$

což můžeme ilustrovat asi takto: jestliže určitá osoba dává přednost A před B v 80% a B před C v 70%, pak při silné stochastické tranzitivitě preferuje A před C alespoň v 80% případů;

b/ mírná stochastická tranzitivita /MST/:

$p/AB/ \geq 0{,}5$, $p/BC/ \geq 0{,}5 \longrightarrow p/AC/ \geq$ minimum $\left[p/AB/, \; p/BC/ \right]$;

při platnosti výše uvedených hodnot by A muselo mít přednost před C alespoň v 70% případů;

c/ slabá stochastická tranzitivita /WST/;

$p/AB/ \geq 0{,}5$, $p/BC/ \geq 0{,}5 \longrightarrow p/AC/ \geq 0{,}5$;

v tomto případě je A preferováno před C alespoň v 50% případů.

Pro použití Coombsovy techniky je na místě předpokládat alespoň mírnou stochastickou tranzitivitu; při slabé stochastické tranzitivitě je technika rozvinutí nespolehlivým prostředkem škálování.

Základní myšlenkou pro výklad nepravidelnosti v posuzování je skutečnost, že posuzovací ideální bod není stabilní, ale má určitou variabilitu. I posouzení jednotlivých podnětů podléhá jisté, byť i menší variabilitě. Aktuální poloha ideálního posuzovacího bodu ovlivňuje posouzení a srovnání některých podnětů více, jiných méně, a některé prakticky neovlivní vůbec. Záleží to jednak na vzdálenosti jednotlivých podnětů od aktuálního bodu, jednak na tom, zda srovnávané body leží na stejné straně od ideálního bodu, či zda tento bod leží na škále kdesi mezi nimi. Situaci znázorníme schematicky na obr. XI.7.

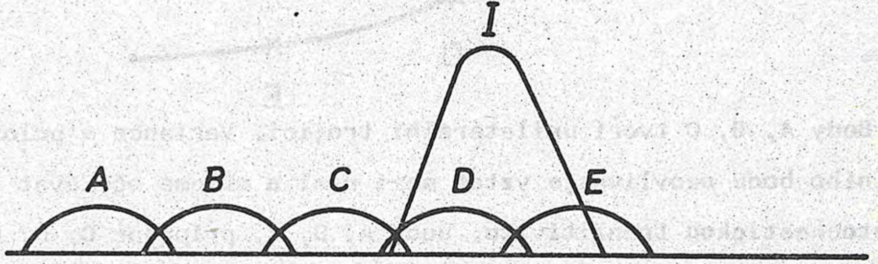
Ve většině případů bude na prvním místě posuzován podnět D a na druhém místě E. Jestliže se však ideální bod v rámci náhodného kolísání posune vlevo, pak sice na prvním místě zůstane podnět D, ale na druhém místě bude C. Posune-li se vpravo, pak na prvním místě bude podnět E a podnět D bude až na druhém místě. Naproti tomu vztah mezi podněty B a A není kolísáním polohy ideálního bodu nikterak ovlivněn.

Páry bodů, které leží na stejné straně J-škály od ideálního bodu, nazýváme unilaterální /např. AB či BC a pod./; páry bodů, ležících na různých stranách J-škály vzhledem k ideálnímu bodu, nazýváme bilaterální /např. CE apod./.

Uvažujeme-li o 3 bodech /a o jejich stochastické tranzitivitě/, dostaneme 3 možné případy, které jsou schematicky znázorněny na obr. XI.8.

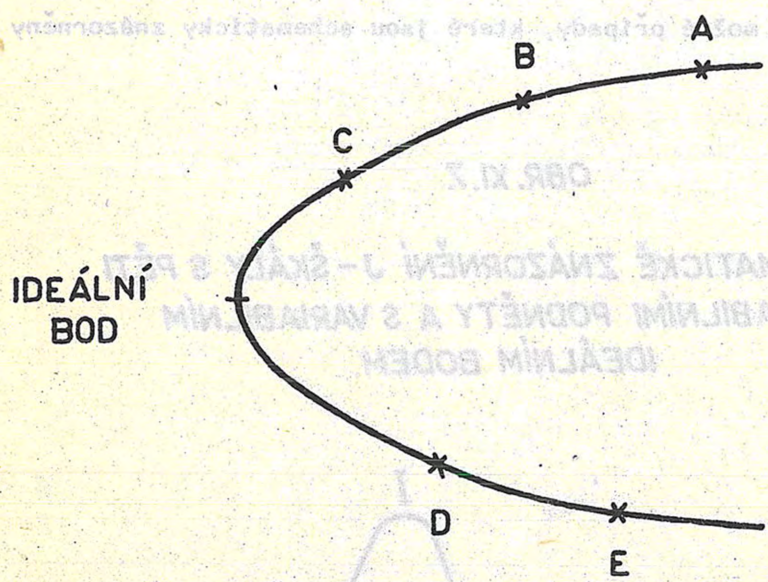
OBR. XI.7.

SCHEMATICKÉ ZNÁZORNĚNÍ J-ŠKÁLY S PĚTI VARIABILNÍMI PODNĚTY A S VARIABILNÍM IDEÁLNÍM BODEM.



OBR. XI.8.

RŮZNÉ TYPY TROJIC PODNĚTŮ.



Body A, B, C tvoří unilaterální trojici. Variance v poloze ideálního bodu neovlivňuje vztah mezi nimi a můžeme očekávat silnou stochastickou tranzitivitu. Body A, D, E, případně C, D, E tvoří přilehlou bilaterální trojici; dva body leží na stejné straně od ideálního bodu a třetí je na druhé straně, ale je buď blíže k ideálnímu bodu než oba druhé body, anebo je od něj dále. Můžeme zde očekávat mírnou stochastickou tranzitivitu, působenou variabilitou polohy ideálního posuzovacího bodu. Konečně body B, D, E tvoří rozštěpenou bilaterální trojici bodů; dva body jsou unilaterální /D, E/, ale třetí bod, který je k nim bilaterální, leží co do vzdálenosti od ideálního bodu mezi nimi /tedy pořadí by mělo být D, B, E/; stochastická tranzitivita bývá zde opět silná.

Coombs probírá uvedenou techniku mnohem podrobněji s řadou příkladů /Coombs, 1964, str. 122 a dále/. Pro informaci ještě uvádím některé empirické studie, které využily Coombsovu techniku rozvinutí; v nich je možno získat názor o její užitečnosti i slabinách. Patří sem např. Coombs /1958/, Coombs - Pruitt /1960/, Dember - Earl /1957/, Hall /1970/, Balcar - Kožený /1975/.

XI.3. Z á v ě r

Při rozboru Coombsovy teorie dat a měření se setkáváme s výraznou snahou zatřídit všechna psychologická data do jednoho systému a najít vhodné metody, jak tato data racionálně zpracovávat. Uvedli jsme zde ukázku jedné - snad nejvíce propracované - techniky. V základní Coombsově monografii /1964/ najde čtenář celou řadu dalších metod; stupeň jejich rozpracovanosti je však nestejný.

Nelze však nevidět ani druhou stranu mince. Příklady, které jsou uváděny, nejsou příliš reprezentativní pro potřeby současné psychologie. Ani vývoj v posledních 10 letech, která uplynula od vydání Coombsovy monografie, nenasvědčuje tomu, že by jeho náměty získávaly mimořádný ohlas a byly rozsáhle užívány. Zdá se, že se stále rozpracovávají /hlavně některými Coombsovými žáky/ a na své uplatnění, ale i kritické přezkoumání ještě čekají. Nepatří tedy zatím ke standardním technikám, které by se zcela běžně užívaly při škálování psychologických jevů, ale tvoří spíše zdroj námětů pro zpracování různých typů dat. Právě v upozornění na různé druhy psychologických dat a v netradičním pohledu na jejich zpracování vidím hlavní metodologický přínos Coombsův do problematiky psychologického škálování.

Kapitola XII.

SUBJEKTIVNÍ POSUZOVACÍ ŠKÁLY

XII.1. Ú v o d

Pojednáme nyní o posuzovacích škálách, které jsou v psychologii velmi populární a často užívané; vyžadují však značné opatrnosti jak při vlastní konstrukci a přípravě, tak i při interpretaci získaných dat. Jedná se o celý trs různých metod a postupů, jejichž společným rysem je to, že se opírají o subjektivní posouzení určitého jevu. Může se jednat o celkové hodnocení osob či jejich jednotlivých vlastností, reakcí, výsledků jejich činností, o posouzení určitých - různě složitých - podnětů, o posouzení vlastního výkonu či o odhad jistoty, zda určité jednání bylo správné, vhodné apod. Do této kategorie škál patří i řada speciálních technik, jako jsou sociometrie, sémantický diferenciál, subjektivní Q-třídění, řada psychofyzických technik atd.

Popularita subjektivních posuzovacích škál a jejich rozsáhlé využívání v aplikované psychologii je snadno pochopitelné. Mnohé jevy, které psycholog sleduje, jsou často velmi komplexní a běžné metody - ať již pozorování, objektivní registrace sledovaných dějů, zjednodušené experimentální modely aj. - jsou buď nepřiměřené, nebo příliš zjednodušující. Jindy se získávají objektivní, ale příliš rozsáhlá data /např. filmování určitých dějů, jako jsou sociální interakce dětí či pracovní činnosti během směny, nahrávání anamnestických rozhovorů či terapeutických sezení na magnetofonový záznam apod./, která nelze rozumně zpracovat v přiměřeném čase. Z těchto důvodů se vytvářely a neustále se tvoří nejrůznější posuzovací stupnice, kterými se snažíme zachytit a číselně vyjádřit sledované jevy na základě jejich komplexního pozorování.

Učinit si přehled o současném stavu problematiky po stránce věcné je prakticky nemožné; jen pro ilustraci uvádím, že Domas a Tiedeman již v roce 1950 zachytili více než 1000 titulů jen z jediné oblasti pedagogické psychologie /posuzování, které provádějí učitelé při vyučování ve škole/. Pojednám zde proto pouze o obecných metodologických pravidlech pro konstrukci subjektivních posuzovacích škál, upozorním na řadu úskalí, která se objevují, a v další kapitole se budu zabývat některými specifickými formami posuzovacích škál. Již v úvodu je třeba upozornit na to, že posuzovací škály mají celou řadu slabin a jejich gnoseologická hodnota při neobratném užití je velmi malá.

Posuzovací škály odpovídají zhruba metodě následných intervalů; většinou se odhaduje, do které z předem stanovených kategorií určitý podnět patří. Historicky jsou odvozeny z psychofyzických metod. V současné době byly detailně rozpracovány a mnohdy i matematicky formalizovány. Při výkladu vycházím z přehledných studií Clause /1968/, Guilforda /1954/ a Tenta /1970/, ve kterých lze najít i odkazy na další speciální literaturu.

XII.2. D r u h y   p o s u z o v a c í c h   š k á l

Guilford /1954/ rozděluje posuzovací škály do 5 skupin: a/ numerické škály, b/ grafické škály, c/ standardní škály, d/ kumulované škály, c/ škály s nucenou volbou. Jde vždy o posouzení určitých podnětů, a to buď jejich umístěním na nepřerušeném kontinuu anebo jejich zařazením do kategorií, do kterých je toto kontinuum rozděleno. Přitom předpokládáme, že se jedná o jevy, které lze posuzovat podle jedné dimenze anebo podle celkového dojmu, který může být sice ovlivněn různými posuzovacími dimenzemi, ale jejich váhy i vzájemné vztahy jsou relativně stálé. Pokud tomu tak není, mohou posuzovací škály vést k velmi nespolehlivým datům. Výsledkem posuzování bývá obvykle přiřazení čísel k jednotlivým podnětům. Toto přiřazení by mělo být citlivé, stálé a spolehlivé.

Jednotlivé druhy posuzovacích škál se od sebe liší postupem, kterým se škála vytváří, druhem i množstvím znaků, dle kterých se hodnocení provádí, a rozlišovací ostrostí, která se žádá od posuzovatele. Jednotlivé druhy posuzovacích škál popíši a v další části upozorním na nejčastější zdroje chyb, které mohou značně snížit hodnotu těchto škálovacích technik.

a/ Numerické posuzovací škály

Pokusným osobám se většinou předloží číselná stupnice, jejíž jednotlivé body jsou definovány, jasně popsány či alespoň vyjádřeny názornými příklady. Úkolem je posoudit jednotlivé podněty a zařadit je do této stupnice. Tak např. při rozboru určité profese je možno uvažovat o různých rysech osobnosti člověka a posuzovat, jak dalece jsou důležité pro pracovní úspěch v dané profesi. Významnost vlastností může být odhadována na čtyřbodové škále:

3 = velmi významná vlastnost pro úspěch v dané profesi

2 = méně významná vlastnost pro úspěch v dané profesi

1 = nevýznamná vlastnost pro úspěch v dané profesi

0 = vlastnost kontraindikující úspěch v dané profesi

Řada odborníků posuzuje nejrůznější vlastnosti z hlediska určité profese. Často se stanoví určitá empirická kritéria pro vyhodnocování dat i pro další postup /např. za podstatnou vlastnost se považuje ta, kterou alespoň 75% posuzovatelů označí číslem 3 nebo 2 atd./.

Numerická škála může mít i větší počet bodů. Při hodnocení afektivních hodnot barev použil Guilford 11-bodovou škálu, v níž jednotlivé body byly určeny takto:

10. Nejkrásnější barva, jakou si lze vůbec představit
9. Překrásná barva
8. Velmi pěkná barva
7. Středně pěkná barva
6. Mírně pěkná barva
5. Indiferentní barva
4. Mírně nehezká barva
3. Středně nehezká barva
2. Velmi nehezká barva
1. Zcela nehezká barva
0. Nejhorší barva, jakou si lze vůbec představit

Jinou škálu jsem použil /Břicháček, 1968/ při sledování subjektivní jistoty o tom, zda rozhodnutí a z něj plynoucí reakce pokusné osoby /rozlišování 2 bodů co do jejich polohy vzhledem k ideálnímu střednímu bodu na podnětovém panelu/ byla správná:

5. Jsem si zcela jist, že jsem reagoval správně

4. Domnívám se, že jsem reagoval správně

3. Asi jsem reagoval správně

2. Jen velmi těžko to mohu posoudit, ale snad jsem reagoval správně

1. Vůbec nedokáži posoudit, zda jsem reagoval správně

Podobných škál existuje celá řada; většinou jsou sestaveny apriorně a bez bližšího empirického ověření se předpokládá, že intervaly mezi sousedními čísly jsou zhruba stejné /alespoň subjektivně/. Tento předpoklad by bylo teoreticky možno ověřit přiměřeným výzkumem, ale většinou se to u posuzovacích stupnic neprovádí a jednotliví autoři se spokojují s více či méně oprávněným odhadem, že uvedený předpoklad platí.

Často diskutovaným problémem je zakotvení numerické škály, tj. stanovení extrémních bodů na škále. Ve výše uvedeném příkladě škály pro afektivní posuzování barev jsou oba póly škály formulovány velmi extrémně /nejkrásnější barva, jakou si lze vůbec představit - - nejhorší barva, jakou si lze vůbec představit/ a na první pohled až neúčelně, protože v praxi se asi nenajdou osoby, které by k nějaké barvě přiřadily podobně ostře formulované posouzení. Přesto se však zavedení podobných výroků doporučuje. Umožňují totiž jasně vyjádřit extrémní postoj, což pomůže pokusným osobám lépe si představit oba póly škály, a tím ji pevně zakotvit. Může se stát, že se najdou jedinci, kteří by byli ochotni daný podnět extrémně posoudit a přitom formulace označená číslem 9 nebo 1 /v případě s afektivním hodnocením barev/ by jim nevyhovovala. Pokud by nebyly k dispozici ještě extrémnější výroky, pak by museli zvolit bod 9 či 1, a tím by se škála zužovala a její citlivost by se snižovala. Navíc se v empirických výzkumech prokázalo, že i když extrémních formulací prakticky není využito, přece jen distribuční rozložení jednotlivých posudků je větší, než je tomu u škál, u kterých extrémní výroky nebyly zavedeny. Posuzovací škála s extrémními výroky je proto celkově citlivější. Obecně platí, že se objevuje tendence konečné body posuzovací škály používat jen vyjímečně, takže počet reálně užívaných bodů je na obou koncích škály snížen o bod. Doporučuje se proto předem stanovit, kolik bodů chceme mít na škále a jejich počet pak ještě zvýšit o dvě extrémní formulace, kterými se škála zakotví. V uvedeném příkladě s afektivním hodnocením barev byla sice škála 11-bodová, ale prakticky bylo využito jen 9 bodů; pokud by nebyly předloženy oba extrémní posudky, pak by se zbylá 9 bodová škála v praxi redukovala na 7-bodovou.

Jiným sporným bodem je umístění neutrálního bodu /jedná-li se ovšem o škálu, která je bipolární/. Jestliže neutrální bod škály označíme číslicí 0, pak nezbytně musíme použít negativní čísla, což nemusí být pro mnohé osoby zcela běžné. Může též vzniknout dojem, že na posuzovaném kontinuu se objevuje určitý zlom. Většinou se proto nedoporučuje pracovat s negativními čísly /viz příklad s afektivním hodnocením barev, kde neutrální bod je označen číslicí 5/.

Celkově lze říci, že numerické škály se velmi snadno vytvářejí, dobře se používají a snadno se vyhodnocují. Z hlediska pokusných osob se jedná o nástroj velmi jednoduchý; předpokládá se pouze schopnost přiřazovat čísla k jednotlivým podnětům ve shodě s jejich vlastnostmi. Záleží však též na škálovaném kontinuu a na jeho objektivním rozsahu. Tam, kde se škálují podněty, které jsou si objektivně velmi blízké, numerická škála s velkým počtem bodů je spíše matoucí.

b/ Grafické posuzovací škály

Grafické škály mají celou řadu forem a podob. Společným jmenovatelem je to, že se předloží úsečka dostatečně dlouhá a pokusná osoba má vyznačit, kam by na ni umístila posuzovaný podnět. Grafické škály mají mnoho variant. Předložená úsečka může být v poloze vodorovné či svislé, mohou na ní být jednotlivé vztažné body vyznačeny čísly či popisem, anebo mohou být pouze zakotveny koncové body škály apod.

Existuje celá řada návodů či pravidel, jak grafické škály sestavovat a prezentovat. Většina z nich se opírá spíše jen o racionální úvahu než o poznatky ze systematického empirického výzkumu. Uvedeme nejčastěji požadované zásady.

Grafické škály bývají většinou předkládány v poloze vodorovné, ale mnohdy jsou výhodnější svislé /"teploměr"/. Svislá škála se lépe popisuje; těsně k bodu, který je na úsečce graficky vyznačen, se připíše jeho charakteristika slovem či celou větou; u vodorovné škály se tento popis provádí hůře a celkový grafický dojem není příliš přehledný. Naproti tomu u svislé úsečky její horní bod sugeruje "lepší" pól škály, což může vést k systematické chybě; i u neutrálních dimenzí může být horní pól přeceňován. Ovšem i u vodorovných škál bývá někdy jejich levá část přeceňována.

Úsečky mají být dlouhé asi 12-15 cm. Příliš dlouhé úsečky nebývají citlivější. Dochází u nich často k tomu, že se jednotlivé posudky grupují do jakýchsi chomáčů, a škála proto není dostatečně využita.

Doporučuje se většinou, aby grafické škály byly nakresleny bez přerušení, čímž se podporuje představa kontinuity. Přerušovaná čára sugeruje dojem přerušované stupnice, jejíž jednotlivé body vznikají sloučením podobných posudků do jedné kategorie; v extrémních případech může vzniknout i dojem různých kvalit, což pak vede k celé řadě nežádoucích důsledků pro způsob posuzování. Nepřerušená čára má i tu výhodu, že si ji pokusná osoba může dělit na tolik dílců, kolik považuje za přiměřené k získaným datům; u přerušené čáry jsme odkázáni na předem stanovený počet dílů.

Pokusným osobám se mají předkládat jednotlivé posuzované podněty spolu s grafickou škálou vždy na nové stránce. Není vhodné nakreslit několik čar vedle sebe či pod sebou a na každé zaznamenat posouzení jiného podnětu. Osoba nutně srovnává své posudky navzájem a vytváří si vlastně pořadovou škálu /pokud ovšem srovnávání podnětů není záměrně sledováno/.

Jestliže se posuzuje několik podnětů z různých hledisek /např. různé osoby z hlediska různých rysů osobnosti/, pak je výhodnější nejprve jeden podnět /tj. určitou osobu/ posoudit ze všech hledisek a pak teprve přejít k druhému podnětu atd. Tento postup je lepší než začít s určitou vlastností a posuzovat v ní všechny osoby; není však vhodný, působí-li výrazně haló-efekt /viz str.213/.

Často se diskutovalo, zda je vhodné, aby "dobrý" či "vyšší" pól škály byl vždy stejně umístěn nahoře nebo vlevo, či zda se má náhodně měnit. I když se zdálo, že změna "pólování" je výhodnější, nejsou pro to empirické důkazy /viz Tent, 1970, str. 873/; naopak častá změna pólovaní posuzovatele spíše mate a vede k nekontrolovatelným chybám. Proto se většinou pozice "lepšího" bodu na škále nemění.

Jestliže pracujeme se škálou, na které chceme uvést několik bodů, a víme, že vzdálenost mezi nimi není stále stejná, pak i na grafické škále je vhodné rozložit body tak, aby prostorově odpovídaly očekávaným hodnotám. Může se stát, že neutrální bod pak není ve středu grafické škály; toto je zcela rozumné, jestliže očekáváme, že většina posudků se bude blížit k jednomu pólu škály.

U grafických škál se nedoporučuje užívat tak extrémní formulace na krajních bodech jako u numerické škály. Pokusné osoby se jim vyhýbají a posouvají svá řešení mnohem více ke středu, než tomu je při užití numerické škály. Většinou se proto doporučuje neumísťovat krajní formulace přímo na konec grafu, ale ponechat ještě malý kus čáry přesahovat přes poslední body. Tím je osobám naznačeno, že může existovat ještě vyhraněnější názor, než jaký je uveden v popisu škály.

Pokud je k dispozici samočinný počítač, je vhodné volit takovou techniku zápisu, aby záznam pokusné osoby mohl být ihned vyděrován a data byla připravena pro další zpracování bez mechanického proměřování grafických záznamů.

Grafické škály mají řadu výhod: snadno se předkládají pokusným osobám, jsou pro ně celkem zajímavé, snadno pochopitelné a mnohdy snáze představitelné než abstraktní čísla. Rychle se získávají rozsáhlá data. Z hlediska experimentátora jsou grafické škály mnohem jemnější než numerické. Ukázalo se /Taylor, Parker, 1964/, že grafické posuzovací škály jsou při jednodimenzionálním podnětovém kontinuu stejně validní a reliabilní jako Guttmanův škálogram.

Kombinaci numerické a grafické posuzovací škály vyvinul Taylor a jeho spolupracovníci /1968, 1970, 1972/ pro účely klinického posuzování. Používá se grafická škála se 100 body. Na této škále jsou různé body zakotveny příkladem či velmi stručným klinickým popisem určitého pacienta. Ukázala se velmi vysoká shoda mezi posuzovateli /trénovaní i netrénovaní; posudky za různých vnějších i vnitřních okolností apod./, a také mnohé experimentální chyby se výrazně snížily. Taylorův postup je sice poměrně pracný, ale pro posuzování komplexních podnětů je zřejmě velmi užitečný.

c/ Standardní posuzovací škály

Předpokladem pro použití tohoto typu posuzovací škály je to, že se v předpokuse pečlivě zvolí jakési příklady, které slouží jako standardní podněty. U jednotlivých standardů se zjistí jejich škálové hodnoty. Optimální situace nastává, jestliže škálová vzdálenost mezi nimi je stále stejná. Pokusné osoby jsou se všemi příklady řádně seznámeny, a teprve pak se jim předkládají jednotlivé empirické podněty, které mají posuzovat. Úkolem je přiřadit jednotlivé podněty ke stanoveným standardům či případně je umístit vhodnou interpolací mezi ně.

Jako příklad se často uvádí hodnocení písma či jiných výtvorů člověka. Pro posouzení kvality se uvede příklad rukopisu, který je možno hodnotit známkou jedna, příklad rukopisu známkovaný dvojkou atd. Rukopis určitého žáka se pak posuzuje srovnáním s předložený-mi typickými příklady. Podobných postupů se užívá často i při vy-hodnocování některých výkonových či projekčních testů, pro hodnoce-ní kvality výrobků atd.

Základním metodologickým problémem vytvoření posuzovacích škál tohoto typu je řádné sestavení standardních podnětů a stanovení je-jich škálových hodnot. Někdy se používalo zjednodušujících postupů, jejichž poznávací hodnota však je minimální. Pro standardy se uvá-děly jako příklady jednotlivé osoby, které byly pokusným osobám známy. Tak např. za člověka velmi iniciativního a aktivního, kte-rý dovede dobře řídit ostatní, považujeme pana X, za člověka jen mírně iniciativního a aktivního, který nedovede příliš dobře řídit ostatní, považujeme pana Y atd. K nim se pak přirovnávaly posuzova-né osoby. Tento postup se zdál zprvu velmi výhodný, protože abstrakt-ní čísla byla nahrazena konkrétními příklady. Ve skutečnosti se však objevily mnohé nesnáze; jednotliví posuzovatelé neznají stej-ně dobře všechny osoby, užívané jako standardy; nelze abstrahovat od afektivních postojů k nim atd. Tím se snižuje hodnota celého postupu a roste nekontrolovatelná experimentální chyba.

Jindy se místo jmen konkrétních osob používaly tak zvané slov-ní portréty, tj. popisy fiktivních osob, které by odpovídaly jed-notlivým bodům na škále. Úkolem pokusné osoby je stanovit, koho z posuzovaných konkrétních osob jednotlivé předložené popisy připomínají. Má-li být tento postup metodicky únosný, pak je třeba velkou pozornost věnovat škálování a přísnému kalibrování slovních portrétů, které budou předkládány jako standardy.

Obecně je možno konstatovat, že standardní posuzovací škály jsou rozumně použitelné pouze tehdy, máme-li dobře sestavenou množinu standardních podnětů. Na tom závisí hodnota celé metody. Sestavení standardů však nebývá jednoduchou záležitostí a vyžaduje většinou značné úsilí.

d/ Kumulativní posuzovací škály

Při kumulativních posuzovacích škálách se pro posuzovaný podnět zachycuje součet bodů /případně jejich průměr/, který byl získán z několika položek. Přitom je možné přisuzovat jednotlivým položkám různou váhu. Nejčastější formou těchto škál jsou různé zaškrtávací listiny, ve kterých je seznam vlastností, názorů nebo různých činností, a úkolem pokusné osoby je označit ty z nich, které odpovídají posuzovanému podnětu. Je možno sestavit seznam charakterových vlastností a zaškrtávat ty, které dobře popisují či vystihují posuzovanou osobu. Pozitivní vlastnosti se většinou skórují +1, negativní -1, a odpověď neutrální /"nevím"/ 0. Pokud se pracuje se seznamem, který má jen menší počet položek, pak často bývají jednotlivé odpovědi váženy složitými indexy. Váhy jednotlivých položek však musí být stanoveny na základě předpokusů, které bývají náročné po stránce časové i technické.

Při předkládání zaškrtávací listiny se většinou připouští neutrální odpověď /nejsem si jist - nevím/, ale doporučuje se žádat, aby tato odpověď byla udávána jen výjimečně. Je též lépe, musí-li posuzovatel reálně odpovídat na každou otázku /to znamená zaškrtnout buď odpověď "ano" či "ne"/, než když se žádá, aby se zaškrtávaly jen ty položky, které se pojí s pozitivní odpovědí.

Velmi rozsáhlou zaškrtávací listinu, obsahující 724 různých výroků charakterizujících pracovníky v průmyslu, sestavil Uhrbrock /citováno dle Guilforda, 1954, str. 272/. Každý výrok posuzovalo 20 mistrů; odhadovali, jak dalece jsou jednotlivé charakteristiky potřebné pro úspěch v práci mistra. Byla použita technika zdánlivě stejných intervalů; škála měla 11 bodů. Pak byly stanoveny průměrné škálové hodnoty jednotlivých výroků a zároveň i jejich variance. Tak např. výrok "Je velmi energický" měl škálovou hodnotu 8,5 /rozptyl 1,95/; výroku "Je dobrý rutinní pracovník" byla posouzena průměrná hodnota 6,0 /rozptyl 1,25/; výrok "Délává nesnáze" je škálován 2,0 /rozptyl 1,35/ atd. Takto získané hodnoty se staly základem pro posuzování jednotlivých osob. Zaznamenávají se škálové hodnoty těch položek, které byly určité osobě přiřčeny, a stanoví se medián jako míra střední tendence. Citovaná studie má však značně problematický základ. Poměrně malý počet posuzovatelů a operacionální definice posuzovacího kritéria /"Jak moc by se asi člověk, o kterém toto tvrzení platí, hodil za mistra v našem podniku?"/ značně zpochybňují externí validitu nálezů. Ani interní validita dat není asi příliš velká. Nelze rozumně předpokládat, že by tak velké množství položek odpovídalo jedné posuzovací dimenzi. Zřejmě se zde shrnují do jednoho indexu údaje z mnoha dimenzí, které jsou spolu jen volně vázány. Vyjádření dat jediným indexem není přiměřené a velmi snadno mohou vzniknout artefakta a nálezy, jejichž hodnota je sporná; případné závěry pro praxi pak vedou na scestí.

Zaškrtávací listina může být použita i v podobě popisu různých činností či případně relativně samostatných prvků činností. Zaškrtávají se ty činnosti, které posuzovaná osoba /či skupina osob/ v určitém čase provede. Tato technika se může užít v mnoha situacích, jako např. při posuzování pracovních činností, při sledování sociálních vztahů během diskuse, při výzkumu společného řešení obtížné úlohy apod. Při tom se škáluje buď aktivita jednotlivých osob nebo výskyt různých prvků či forem činností.

Objevuje se i jistá analogie slovních portrétů, o kterých byla zmínka při výkladu standardních posuzovacích škál. Používá se většinou několika poměrně stručných tvrzení /např. "Vždy se snaží udělat druhým radost nějakou maličkostí."/ a odhaduje se, pro které osoby z daného souboru jsou výstižná. Celkové skóre se získává sloučením všech položek /kladných i záporných/. I zde se ovšem vnucuje otázka, zda tvrzení tvoří homogenní skupinu, kterou bychom mohli vyjádřit jednodimenzionální škálou.

Kumulativní posuzovací škály mají nesporně řadu předností. Snadno se předkládají, pro pokusnou osobu je úloha poměrně lehká a snadno pochopitelná. Není třeba srovnávat posuzované jevy či osoby navzájem, ale stačí pouze odhadovat, zda určitou vlastnost mají či nemají, či zda určité činnosti se objevily či nikoliv atd. Hodí se dobře pro komplexní situace, jako je posouzení stavu pacienta v průběhu terapie, posouzení vhodnosti uchazeče o zaměstnání apod. V jiné formě - tam, kde se posuzuje jediná proměnná u řady osob - se vlastně přibližujeme k osobnostnímu dotazníku, který pokusná osoba vyplňuje nikoliv o sobě samé, ale o druhých lidech. Další výhodou těchto škál je to, že se dobře a rychle vyhodnocují.

Hlavní problematickou otázkou při použití kumulativních posuzovacích škál bývá zjištění, zda zvolené položky tvoří jedinou dimenzi. Pokud tomu tak není, může snadno dojít ke sloučení dat, která slučitelná nejsou. To pak může značně zkreslit či znesnadnit interpretaci výsledků.

e/ Posuzovací škály s vynucenou volbou mezi variantami

Posuzovací škály s vynucenou volbou jsou poněkud jinou formou škál, než byly předchozí typy. Ptáme se, zda posuzovaný podnět má určitou vlastnost ve větší míře než jinou vlastnost. Poměrně známou analogií této škály je Kuderova preferenční škála pro volbu povolání; ze 3 prvků se určuje ten, kterému je dána přednost, a dále ten, který je hodnocen nejníže /čili se vlastně sestaví pořadí/.

Při vytváření škál tohoto typu se často uvažuje, mezi kolika možnými odpověďmi má pokusná osoba volit. Highland a Berkshire /citováno dle Guilforda, 1954/ srovnávali různé počty i kategorie prvků, a to:

a/ ze dvou tvrzení /buď obě kladná či obě záporná/ označit to, které posuzovaný podnět lépe vystihuje,

b/ ze tří tvrzení /buď všechna kladná či všechna záporná/ vybrat nejlépe a nejhůře popisující, c/ ze čtyř tvrzení /všechna kladná/ určit dvě nejlépe odpovídající,

d/ ze čtyř tvrzení /všechna kladná/ určit jedno nejlépe a jedno nejhůře popisující,

e/ ze čtyř tvrzení /dvě kladná a dvě záporná/ určit jedno nejlépe a jedno nejhůře popisující,

f/ z pěti tvrzení /dvě kladná, jedno neutrální, dvě záporná/ označit jedno nejlépe a jedno nejhůře popisující.

Z rozsáhlé studie, ve které se uvažovalo i o reliabilitě a o validitě výsledků /srovnáním s úsudkem znalců/ vyplynulo, že nejvýhodnější je forma uvedená sub c/.

Jako příklad posuzovací škály s vynucenou volbou, která je blízká variantě c můžeme uvést studii Cosgrova /1959/, který zkoumal chování učitelů ve vyučování. Faktorováním dat, získaných v předpokuse, zjistil Cosgrove obecný faktor celkové působnosti učitele a čtyři specifické faktory: znalost a utřídění látky, přiměřený poměr k žákům, přiměřený plán a průběh vyučování, radost z práce s dětmi. Pro tyto čtyři faktory sestavil posuzovací škálu o 10 položkách. Ve všech položkách byla čtyři tvrzení, z nichž každé odpovídalo jinému specifickému faktoru. Úkolem žáků bylo sestavit je do pořadí podle toho, jak dobře popisují určitého učitele. Uvedeme příklady dvou položek:

č.3 - vyučování vede vždy dobře

- vyučování ho těší
- je spontánně přátelský
- myslí a postupuje logicky

č. 6 - má smysl pro humor
- úkoly zadává pravidelně a rovnoměrně
- žáci nemají strach se ptát
- učí látku ve shodě s novými poznatky

Zjišťuje se přitom, kolik bodů připadlo pro jednotlivé faktory. Po sloučení výsledků od všech žáků se vytvořil jakýsi čtyřrozměrný profil každého učitele. Srovnání jednotlivých učitelů je však velmi obtížné.

Solidní sestavení posuzovací škály s vynucenou volbou není jednoduchou záležitostí. Pokud není k dispozici nějaký empirický podklad /jako v předchozím příkladě/, pak má celý postup několik fází, které by neměly být opomenuty. Nejprve se shromáždí co nejvíce výroků, případně jmen vlastností, pokrývajících sledovanou problematiku. Uvažuje se, které z nich jsou nevýrazné a málo diferencující, a které budou pravděpodobně umístěny na pólech škály /jak na kladném, tak na záporném/. Experimentálně se pak určují dva indexy. Jsou to:

a/ Diskriminační index, kterým se vyjadřuje, zda daný výrok skutečně diferencuje mezi jevy či osobami, umístěnými na obou pólech škály.

b/ Preferenční index, kterým se zjišťuje, jak jednotlivé výroky jsou samy o sobě hodnoceny skupinou posuzovatelů /třeba z hlediska atraktivnosti, afektivní hodnoty, sociální žádoucnosti apod./. Tento druhý moment může totiž rušivě zasahovat do vytváření škály tím, že vlastní názory posuzovatelů mohou být v interakci s posuzováním. Experimentální chyba se tím zvětšuje.

Při stanovení indexů preference musíme mít k dispozici skupinu posuzovatelů, která je zhruba analogická s těmi, kteří budou později se škálou běžně pracovat. Při výběru položek se bere v úvahu jak diskriminační, tak preferenční hodnota jednotlivého výroku. Tak třeba při předkládání čtveřice výroků se obvykle užijí dva výroky s vysokou hodnotou preferenční, z nichž jeden má velkou hodnotu diskriminační a druhý nikoliv, a dva výroky s nízkou hodnotou preferenční /opět jeden s vysokou diskriminací a druhý nikoliv/. Péče je věnována i instrukci pro posuzovatele. Orientační přezkoušení sestavené škály a případná revize některých položek jsou obvykle nezbytné

Kritický přehled poznatků o posuzovacích škálách s vynucenou volbou uveřejnil Zavala /1965/. Ukazuje se, že tato forma posuzovací škály bývá výhodnější než předcházející formy, ale pouze za předpokladu, že je řádně předem připravena. Svědčí o tom několik studií, které byly zaměřeny na problematiku hodnocení lidí v zaměstnání /přehled viz Guilford, 1954, str. 276/. Není ovšem jisté, že by tomu tak muselo nutně být i v jiných oblastech, pro které se posuzovací škály užívají.

Výhodou posuzovacích škál s vynucenou volbou je značně vysoká reliabilita dat v čase. Škála nemusí být příliš rozsáhlá; 10 - 12 dobře vybraných položek vede zhruba ke stejně validním výsledkům jako škála s 28 - 32 položkami /viz též Tent, 1970, str. 898/. Z těchto poznatků vyplývá, že se jedná o metodu, která je nesporně velmi užitečná a bývá neprávem - snad pro další proceduru, spojenou s jejím sestavením - opomíjena.

Ze Zavalova přehledu /1965/ je však zřejmé, že vhodná kombinace různých posuzovacích škál vede ke kvalitnějším výsledkům, než izolované použití kterékoliv z nich. Tento poznatek je pochopitelný a platí jistě i pro jiné škálovací metody než jen pro posuzovací škály.

XII.3. Metodické problémy při práci se subjektivními posuzovacími škálami

Předpokladem pro přiměřené používání posuzovacích škál je v prvé řadě očekávání, že pokusná osoba dokáže posuzovat zkoumané jevy s určitým stupněm přesnosti a objektivity, a že toto posouzení alespoň v hrubých rysech odráží objektivní skutečnost. Naše úsudky jsou však ovlivňovány řadou chyb a často si ani subjektivně nebýváme příliš jisti, zda naše posouzení je přiměřené a správné. Z dlouholetých metodických zkušeností vyplynulo, že při posuzování se objevují jak chyby náhodné a těžko analyzovatelné, tak chyby systematické, kterých se za určitých okolností dopouští většina posuzovatelů. Zdroje systematických chyb je třeba znát a již při plánování výzkumu uvažovat o cestách, jak je odstranit, minimalizovat či kontrolovat.

V prvé řadě záleží na vlastním výzkumném tématu. Jiné chyby mohou nastat při posuzování žáků učitelem, jiné při posuzování pracovních zásluh, jiné při posuzování průběhu terapie, jiné při posuzování uměleckých děl, jiné při posuzování sportovních výkonů atd. Těmito speciálními aspekty práce s posuzovacími škálami v různých oblastech psychologie se nemůžeme zabývat. Výklad se zaměří na obecnější poznatky a pravidla. Opírám se přitom hlavně o studie Brogdena a Taylora /1950/ a Guilforda /1954/. Jako příklad detailní analýzy této problematiky pro specifické otázky pedagogicko- psychologického výzkumu efektů školního vyučování lze uvést studii Tenta /1970/.

XII.3.a. K o n s t a n t n í c h y b y a j e j i c h k o n t r o l a

U subjektivních posuzovacích škál se objevují některé typy experimentálních chyb, které mají obecný charakter a působí zcela pravidelně. Jsou latentním nebezpečím při všech výzkumech, které se opírají pouze o posuzovací škály. Obvykle se o nich hovoří jako o konstantních chybách. O většině z nich je nashromážděno poměrně hodně poznatků ze systematických empirických výzkumů. Základní poznatky stručně vyložím; je třeba je znát a uvažovat o nich již ve fázi přípravy výzkumu. Dodatečná snaha o odstranění vlivu konstantních chyb z posuzovacích škál bývá velmi obtížná a často již prakticky neproveditelná.

Hlavní druhy konstantních chyb jsou tyto:

1/ "Haló efekt"

Tato globální chyba v posuzování lidí je známa nejdéle a byla popsána již na začátku dvacátých let; její označení pochází od Thorndika. Posuzujeme jednotlivé rysy lidí podle celkového dojmu, který na nás učinili. Jestliže celkový dojem je kladný, pak i jednotlivé rysy osobnosti, ale často i výsledky jejich činnosti, bývají posuzovány kladně. Důsledkem je jednak zvýšení pozitivních korelací mezi posuzovanými vlastnostmi, jednak pokles validity posouzení některých /možná dokonce i všech/ vlastností, které jsou sledovány. Dochází zřejmě k tomu, že do posuzování se promítnou dimenze, které nejsou přiměřené, a smísí se s dimenzemi přiměřenými. Haló efekt je velmi vtíravý a asi se mu nelze zcela vyhnout. Prosadí se hlavně při posuzování vlastností, které nejsou přesně definovány, u vlastností, které mají vysokou morální hodnotu, u rysů osobnosti, které se nesnadno pozorují či které se jen obtížně oddělují od jiných rysů, u jevů, které se vyskytují jen velmi vzácně, a konečně i u vlastností, které se projevují v interakci s jinými lidmi /zde se dokonce může přidružit i haló efekt dalších lidí/.

Snížení haló efektu lze docílit důkladným zácvikem posuzovatelů, využitím techniky posuzování s vynucenou volbou a konečně též takovým uspořádáním pokusu, kdy se posuzuje nejprve rys A u všech osob, pak rys B u všech osob atd. Tento postup je přiměřenější než posuzovat nejprve osobu X po všech stránkách, pak osobu Y atd. Technicky to znamená, že na jednu stránku záznamového archu pro pokusnou osobu uvedeme jeden rys a všechny posuzované osoby /nikoliv tedy na jedné stránce jednu osobu a všechny posuzované rysy/.

2/ Logická chyba v posuzování

Rysy osobnosti či znaky činností, které spolu těsněji souvisejí /často i jen dle subjektivního dojmu či předsudku posuzovatele/, bývají hodnoceny podobně. Zvyšuje se opět uměle korelace mezi posuzovanými podněty. U haló efektu k tomuto zvýšení docházelo na základě celkového dojmu o určité osobě; zde je to působeno zdánlivými vztahy mezi vlastnostmi lidí. Odstranění či alespoň snížení této chyby je možno docílit tím, že se vyhýbáme posuzování abstraktních či špatně definovaných vlastností, kdy může snadno dojít k jejich sémantickému překrývání. Pro posouzení používáme, pokud to je možné, pozorovatelné formy či prvky činnosti. Logická chyba se objevuje sice nejčastěji při posuzování lidí, ale působí i při posuzování jiných podnětů.

3/ Vliv známosti

Obvykle se objevuje tendence posuzovat osoby, které známe či které nás zaujmou, poněkud lépe než osoby cizí. Tato chyba se odstraňuje velmi obtížně. Pokusné osoby, které jsou s ní předem seznámeny, mají obvykle tendenci opačnou a posuzují známé osoby příliš ostře; dochází pak k obrácené chybě.

4/ Chyba kontrastu

I tato chyba bývá dosti známá a častá. Projevuje se tím, že pokusná osoba, která má sama určitou vlastnost /anebo se domnívá, že ji má/, posuzuje ostatní lidi spíše v opačném směru. Podobný mechanismus vede třeba k tomu, že lidé, kteří se o sobě domnívají, že umí riskovat, očekávají, že ostatní riskují méně. Lidé, pro jejichž životní styl je příznačný přísný řád a pravidelnost, považují ostatní za velmi nepořádné a nedůsledné.

Existuje však i opačný druh chyby, kdy posuzovatel posuzuje ostatní lidi analogicky jako sebe. Tato chyba je méně častá, ale vyskytuje se při posuzování některých sociálně psychologických vlastností. Člověk, který má značně vyvinutý smysl pro kooperaci, často považuje i ostatní lidi za kooperantní, protože je sám ke spolupráci přitahuje. Zde se prosazuje do posuzovacích kritérií osobnostní struktura posuzovatele; tento druh chyb se jen poměrně těžko odkrývá a nebývá většinou ani kontrolován.

Podobný druh chyby se objevuje často při posuzování postojů. Posuzovatelé, kteří mají sami vyhraněný postoj, posuzují jednotlivé výroky či jevy ostřeji a často umísťují mnohé z nich blíže k extrémním škálovým hodnotám. Naopak osoby s nevyhraněným postojem obvykle kumulují svá posouzení do středu škály. Tento druh chyb v posuzování se také těžko odstraňuje. Při plánování výzkumu je proto třeba zvažovat, jak provedeme výběr posuzovatelů a zda se budeme snažit, abychom měli skupinu homogenní /pozor však na zobecňování výsledků/, či zda záměrně vytvoříme skupinu, která je ve svých postojích heterogenní /tím se zvýší variabilita výsledků, výsledná škála se hůře vytváří, ale má obecnější platnost/.

5/ Chyba blízké asociace

Relativně nejpozději byla popsána chyba, která mohla ovlivnit či zkreslit mnohé minulé výzkumy. Ukázalo se totiž, že prostorová či časová blízkost posouzení dvou podnětů vede k tomu, že jsou po- posuzovány analogicky. V jedné studii /Stockford a Bissel, 1949 - citováno dle Guilforda, 1954/ se prokázalo, že průměrná korelace mezi posouzením vlastností, které byly předkládány v seznamu těsně za sebou, byla + 0,66. Průměrné korelace mezi vlastnostmi, zapsanými v seznamu ob jedno, dvě a více míst, postupně klesaly a ustálily se na hodnotě + 0,46 , která se objevila mezi vlastnostmi oddělenými pěti jinými položkami. Když bylo ve stejném seznamu pozměněno pořadí podnětů, objevily se opět vyšší korelace mezi posudky sousedních položek - i když se nyní jednalo o zcela jiné dvojice položek. I tato chyba vede zřejmě k tomu, že korelace mezi posudky jednotlivých podnětů se v průměru uměle zvyšuje.

Praktické důsledky, které vyplývají ze znalosti tohoto typu chyb při posuzování, jsou tyto: vlastnosti, o kterých předpokládáme, že by mohly být spolu vázány, by neměly být předkládány k posouzení těsně za sebou, a naopak takto by měly být seřazeny ty, které by spolu souviset neměly. Kromě toho by bylo účelné různým pokusným osobám předkládat podněty v různém pořadí. Snížit lze tuto chybu též tím, že necháme posuzovat všechny posuzované osoby nejprve z hlediska jedné vlastnosti, a pak - pokud je to možné - až po určitém časovém odstupu z hlediska další vlastnosti atd. Do výsledků se však může vloudit jiná forma této chyby: to, co platí o vlastnostech, může platit o posuzovaných osobách. Je proto na místě měnit i pořadí osob, které jsou posuzovány. Zdánlivě neutrální seznam osob, daný abecedou, nemusí být z tohoto hlediska nejlepším řešením.

6/ Chyba centrální tendence

O této chybě jsme se již zmínili při popisu některých posuzovacích technik. Řada pokusných osob totiž váhá využít extrémní hodnoty posuzovací škály; posudky se kumulují kolem středu a snižuje se jejich variance. Tato chyba bývá tím větší, čím méně posuzovatel jednotlivé posuzované osoby či podněty zná. Plyne z toho, že je výhodnější užívat raději škály s větším počtem bodů a případně vytvářet takové škály, u kterých nejsou vzdálenosti mezi jednotlivými body lineární. Vzdálenost bodů kolem středu by měla být větší než na pólech škály.

7/ Posuzovací styl

Je běžně známou skutečností, že někteří posuzovatelé jsou spíše kritičtí a přísní, jiní jsou mnohem mírnější. Posouzení téhož podnětu pak může být až diametrálně odlišné. Rorer /1965/ v přehledné studii upozorňuje na to, že existuje jakási tendence spíše souhlasit s předloženými tvrzeními. Její stupeň je individuálně příznačný a stálý. S touto chybou je třeba počítat, a to hlavně při kumulativních škálách, u kterých může snadno dojít ke značnému zkreslení.

8/ Časová chyba

Jestliže se posuzuje průběh komplexních jevů v čase, může často dojít k chybě tím, že posuzovatel podcení ty jevy, které se vyskytují jen řídce a naopak přecení ty, které se objevují často. Tento druh chyby vede k nevyváženým posudkům, které mohou značně zkreslit výslednou škálu. Časová chyba se překonává jen velmi obtížně; může být částečně odstraněna řádnou instruktáží posuzovatelů. Dále se doporučuje, aby se k posouzení předkládaly raději delší časové úseky, ve kterých se jednotlivé posuzované prvky mohou plastičtěji projevit.

Bylo by možno najít jistě i další zdroje konstantních chyb, které mohou plynout z nepřesné instrukce pro pokusné osoby, z jejich nesprávného výběru, ze způsobu, jak je posuzovaný materiál předkládán /např. posouzení terapeutického rozhovoru buď z přepisu stenografického záznamu nebo z poslechu magnetofonového záznamu anebo konečně na základě zvukového filmového záznamu, ve kterém lze hodnotit i neverbální projevy./ Kdo se chce detailněji zabývat metodologickými problémy subjektivních posuzovacích škál, musí se nezbytně opřít o speciálnější literaturu, která však bývá většinou publikována pouze časopisecky.

XII.3.b. Metody pro snížení konstantních chyb při posuzování

Pro snížení konstantních chyb v subjektivních posuzovacích škálách nelze stanovit jednoznačný návod. Konkrétní podmínky výzkumu či praxe si vždy vyžadují konkrétní a racionální rozbor celé situace. Přesto však je možno stanovit několik obecnějších pravidel. Snad nejdůležitější podmínkou pro snížení konstantní chyby je opatrný výběr pokusných osob /pozor však na externí validitu!/, které by měly být jednak dostatečně motivovány pro spolupráci, jednak musí být i osobnostně disponovány k tomu, aby jejich posuzování bylo nezaujaté a zároveň i dostatečně citlivé. Důkladný trénink v předpokusech je zcela nezbytný. Při něm se ukáže, zda pokusné osoby správně a jednoznačně pochopily instrukci a zda si do experimentu nepromítají chybné interpretace o jeho zaměření a smyslu či o předpokládaných hypotézách experimentátora. Seznámení s různými typy konstantních chyb by mělo být součástí přípravy posuzovatelů. Pokud je to možné, doporučuje se uspořádat s pokusnými osobami skupinovou diskusi, ve které se mohou jednak ujasnit jejich úkoly, jednak vyměnit zkušenosti jak z předpokusů, tak i ze samého výzkumu. Zároveň experimentátor může z diskuse poznat, zda výzkum proběhl podle jeho záměru, či zda nebude třeba dodatečně - ovšem na základě důkladných argumentů - některé osoby vyloučit a jejich výsledky vůbec nezpracovávat. V diskusi může experimentátor též odhadnout, zda skupina pokusných osob může být považována alespoň zhruba za reprezentativní /zda nejsou příliš přísní či zda se nejedná o skupinu odborníků, která je profesionálně determinována ve svých postojích či v očekávaných výsledcích atd./.

Pokud se jedná o počet pokusných osob - pak obecně řečeno - čím více, tím lépe. Technické možnosti mnohdy nutí k práci s malým počtem osob /v mnoha studiích se setkáváme s výpočtem indexu shody mezi dvěma posuzovateli/; v tom případě je nezbytná značná opatrnost při interpretaci výsledků a hlavně při zobecňování závěrů.

Je třeba též uvažovat o možných interakcích mezi posuzovateli a posuzovanými. Různé předsudky /často mohou být generační - ve výzkumech, kde starší osoby hodnotí a posuzují nastupující generaci/ mohou značně zkreslit výsledky. Tato problematika je blízká otázkám výzkumu sociální či mezilidské percepce. Není zatím ještě dořešena a jak po stránce metodologické, tak i po stránce pojmové validity je zde mnoho nezmapovaných oblastí. I na tyto otevřené otázky a na možné zdroje chyb, které z nich mohou vyplynout, je třeba pokusné osoby včas upozornit.

Konstantní chyby v posuzování je možno též snížit jasnou a jednoznačnou definicí posuzovaných podnětů. Některé posuzované vlastnosti jsou zřejmé a pokusné osoby o nich mají zhruba stejnou představu /výkonnost, rychlost, originalita, energičnost/, ale jiné mohou být mnohoznačné /takt, odvaha, spolupráce, popularita/ a je třeba vymezit je buď definičně nebo alespoň vhodnými příklady. Je třeba se vyhýbat příliš komplexním pojmům /dobrý charakter apod./. Při popisech by neměly být užívány výrazy sémanticky neurčité /často, mnohdy, velmi, někdy apod./. Musí být též jasno, zda se posuzuje aktuální stav nebo minulost anebo naopak budoucí perspektivy; aktuální stav se posuzuje spolehlivěji.

Celou řadu dalších doporučení spolu s návodem pro výpočet velikosti konstantních chyb pomocí analýzy rozptylu uvádí Guilford /1954/.

V souvislosti s problematikou konstantních chyb při posuzování je vhodné se zmínit i o požadavcích na osobnost posuzujících osob. Je nesporné, že různí lidé nejsou stejně schopni posuzovat určité jevy dostatečně citlivě. Rozdílná posuzovací schopnost souvisí s řadou osobnostních rysů i s celkovým zaměřením zvolených posuzovatelů.

Vhodná motivace je nesporně jedním ze základních předpokladů činnosti posuzujících osob. Lépe postupují ti, kteří svůj úkol řeší se skutečným zájmem. Proto je třeba znovu zdůraznit význam instrukce i vhodné informovanosti pokusných osob. Rozhodně je třeba varovat před takovým postupem, kdy posuzovatele tvoří skupina, která je závislá na experimentátorovi /jeho podřízení, studenti apod./.

Pokusné osoby by měly být voleny z lidí uvážlivých, kteří nepodléhají snadno různým vnějším vlivům či prvním dojmům. Přílišná uspěchanost a vyslovení kategorických soudů bez přemýšlení nejsou obvykle slučitelné se schopností dobře a spolehlivě posuzovat.

Introverti většinou posuzují jiné osoby lépe než extroverti. Určitá úroveň empatie je dalším předpokladem pro výběr dobrých posuzovatelů. Znalost sebe sama v určitých vlastnostech bývá spojena s lepší schopností posoudit jiné lidi v analogických vlastnostech. Naproti tomu životní či pracovní zkušenost nemusí být zárukou dobrého posuzování. Někdy dokonce profesionální zkušenost může vést k jednostranným až chybným posudkům. Učitel matematiky může posuzovat celkové schopnosti žáka na základě svých vlastních zkušeností značně jinak než třeba učitel češtiny.

XII.4. S o u h r n

Subjektivní posuzovací škály se v psychologii - a to hlavně v různých oblastech užité psychologie - používají velmi často. Mají řadu výhod, ale zároveň i celou řadu úskalí.

Co můžeme považovat za hlavní výhody subjektivních posuzovacích škál? Velká přednost je v tom, že je lze užít pro posouzení značného množství podnětů. Metoda pořadových škál se stává nepřehlednou a nespolehlivou, máme-li zařadit více než třicet či čtyřicet podnětů. Technika párového srovnávání je reálně možná do sedmi či osmi podnětů a při větším počtu je technicky obtížně proveditelná. Subjektivní posuzovací stupnice jsou omezeny prakticky jen ochotou pokusných osob spolupracovat a jejich schopností pracovat delší dobu soustředěně.

Další předností subjektivních posuzovacích škál je možnost jejich užití na velmi rozsáhlou oblast problémů jak teorie, tak praxe. Setkáváme se s nimi prakticky ve všech odvětvích psychologie. Tato šíře záběru nabízí možnost srovnávacích studií, které jsou neproveditelné, pokud nemáme k dispozici data získaná alespoň zhruba stejnými metodami.

Z hlediska pokusných osob jsou vcelku zajímavé a přitažlivější než většina jiných škálovacích postupů. Blíží se každodenní zkušenosti člověka a i tím jsou výhodné i z hlediska ekologické přiměřenosti. Zanedbatelný není ani ten fakt, že práce s dobře vyvinutými posuzovacími škálami vyžaduje poměrně méně času než je tomu při použití jiných technik.

U komplexních a multidimenzionálních podnětů - jako je např. posuzování uměleckých děl, dojmů z vrcholných sportovních výkonů či celkové posouzení osobnosti - je subjektivní posuzování jednotlivých podnětů relativně nejvhodnější cestou. Párové srovnávání či sestavování pořadí bývá v těchto podmínkách obtížnější a často vede k tomu, že se posuzuje zjednodušeně a mnohdy jen z hlediska jediné dimenze.

Naproti tomu je třeba jasně říci, že řada subjektivních posuzovacích škál má spornou hodnotu. Intuitivně sestavené škály, které nejsou prověřeny v předběžných šetřeních, vedou často k závěrům, které jsou gnoseologicky chybné. Před jejich neodborným využitím v praxi, a to hlavně tam, kde se rozhoduje o osudech lidí /v poradenství, při hodnocení lidí, při posuzování celkového stavu člověka apod./, je třeba velmi důrazně varovat.

Řada chyb, které se při práci se subjektivními posuzovacími škálami objevují, snižuje jejich hodnotu. Znalost a předběžná analýza možných chyb umožňuje alespoň do jisté míry jejich kontrolu, a tím i zpřesnení údajů, které získáváme. Znamená to ovšem mnohem důkladnější a kritický rozbor používaných škál a jejich důkladnou přípravu jak před výzkumným, tak před rutinním využitím.

Validitu subjektivních posuzovacích škál nelze jednoznačně určit. Většinou se určuje korelací s vnějším, objektivním kritériem. Výsledky nejsou jednoznačné. Při posuzování rysů osobnosti se ukazuje, že některé vlastnosti jsou posuzovány poměrně správně, u jiných se projevuje velká chyba. Údaje o validitě jsou dále ovlivňovány i výběrem a zácvikem pokusných osob, podmínkami, ve kterých výzkum probíhal, i jinými momenty, které často znemožňují zobecnění nálezů. Je proto třeba validitu subjektivních posuzovacích škál ověřovat empiricky v konkrétních podmínkách. V mnoha případech však vnější kritérium chybí /např. při posuzování uměleckých děl, módních jevů, chutnosti potravin apod./ a validizační studia tohoto typu nejsou proveditelné.

Do jisté míry obdobná situace je i z hlediska reliability. Většinou se uvádí, že bývá poněkud vyšší, než je tomu u techniky pořadových škál. Existuje však řada potíží při jejím stanovení. Opakované posouzení nebývá vhodné, protože korelace se může uměle zvyšovat jinými faktory - hlavně paměťovými. Proto se většinou srovnávají posudky jednotlivých pokusných osob; je zde sice určité nebezpečí, že posuzovatelé podlehnou podobným chybám, ale toto zkreslení je relativně menší, než je zkreslení, které vzniká při opakovaném posuzování. O řadě studií, které se zabývaly reliabilitou subjektivních škál referuje Clauss /1968/. Pro stanovení různých aspektů reliability jsou k dispozici různé metody /adaptace Spearman-Brownovy formule, průměrná pořadová korelace mezi posuzovateli aj.; viz Guilford, 1954, nebo Clauss, 1968/.

Jednou z nejčastěji diskutovaných otázek je vztah mezi počtem bodů na subjektivní posuzovací škále a reliabilitou výsledků. Obvykle se uvádí, že optimální je sedmibodová škála, ale objevily se i jiné názory. Z pečlivé srovnávací analýzy i z rozsáhlého vlastního výzkumu došli Lissitz a Green /1975/ k závěru, že již při pětibodové škále jsou výsledky značně citlivé a zvýšení počtu budů se se projeví jen nepatrným zvýšením reliability. Záleží ovšem v prvé řadě na tom, zda body na subjektivní posuzovací škále jsou pro pokusné osoby dostatečně srozumitelné a jasné; není-li tomu tak, pak se ztrácí reliabilita nezávisle na tom, kolik stupňů má škála.

Celkově však lze konstatovat, že v mnoha případech při posuzování a hodnocení lidí či při vyhodnocování složitých procesů vnějšího světa nemáme zatím k dispozici vhodnější metody, a proto jistě i v budoucnosti budou subjektivní posuzovací škály užívány stejně často, jako je tomu dnes. I ne zcela přesná metoda je výhodnější než pouhá spekulace. Pokrok můžeme čekat spíše v metodickém prohlubování a zlepšování jednotlivých škál než v radikální změně postupů. Je však věcí každého autora, který se subjektivními posuzovacími škálami pracuje, aby nakládal se svými nálezy s patřičnou obezřetností a kritičností.

Kapitola XIII.

VYBRANÉ SPECIÁLNÍ ŠKÁLOVACÍ TECHNIKY VYCHÁZEJÍCÍ ZE SUBJEKTIVNÍCH POSUZOVACÍCH ŠKÁL

XIII.1. Ú v o d

Postupem doby byly z tradičních subjektivních posuzovacích škál odvozeny některé další speciální metody. Řada z nich byla detailně propracována a užita v řadě výzkumů i při řešení mnoha praktických otázek. Jedná se mnohdy o různá teoretická východiska. Stručně uvedu ty z nich, které jsou častěji užívány a lépe propracovány, a naznačím pouze jejich základní principy.

Zmíním se o čtyřech speciálních technikách, a to o Q-třídění, o sémantickém diferenciálu, o posuzovací škále se subjektivním zakotvením a o sociometrických postupech.

XIII.2. P o s o u z e n í   p o m o c í   Q - t e c h n i k y

Před více než 20 lety Stephenson /1953/ navrhl Q-techniku faktorové analýzy, jejímž principem je srovnávání výsledků jednotlivých osob. Cílem tohoto postupu je zjistit společné rysy různých osob a dát tak objektivnější podklady pro psychologickou typologii. Písmenem Q označil intraindividuální i interindividuální korelace; písmenem R pak značil korelace mezi výsledky z různých testů či zkoušek, o které se opírala klasická faktorová analýza. Q-technika faktorové analýzy byla velmi slibná pro potřeby výzkumu osobnosti a byla z ní odvozena i její posuzovací forma, kterou můžeme považovat za specifickou formu subjektivních posuzovacích stupnic. Detailní bibliografii o Q-technice a její metodologii publikoval Brown /1968/; uvádí v ní téměř 600 studií.

Q-technika posuzování /také Q-třídění či Q-sort/ předpokládá, že se předem vytvoří 60 - 100 výroků, které popisují člověka. Mohou být odvozeny z určitých teorií, z osobnostních dotazníků, z jiných posuzovacích škál, z umělecké literatury či z běžných formulací. Jednotlivé výroky jsou na kartách předloženy posuzovatelům s žádostí, aby uvážili, zda a do jaké míry přiměřeně popisují určitou osobu, která má být hodnocena. Výroky mají být rozděleny do určitého počtu kategorií /obvykle 7 - 11/, které jsou umístěny na jednorozměrném kontinuu od kategorie "Zcela přesně jej to vystihuje" až po kategorii "Pro něj zcela nepřiměřené". Přitom je důležité, že je předem stanoveno /většinou za předpokladu normálního rozložení/, kolik karet musí být umístěno v jednotlivých kategoriích. Tím je posuzovatel nucen své třídění podřídit stanovenému systému.

Uvedeme některé příklady předepsané Q-distrubuce při 11 kategoriích; některé další příklady uvádí Kerlinger /1972/.

Z tabulky XIII.1. je zřejmé, že při 90 položkách musí pokusné osoby vždy 3 prvky zařadit do extrémních kategorií, po 4 prvcích do kategorií č. 9 či č. 1, 16 prvků do střední kategorie apod.

Q-technika posuzování může být užita zhruba ve dvou směrech:

První způsob užití je vhodný z hlediska potřeb výzkumu osobnosti a její dynamiky. Pokusná osoba uvažuje, jak dalece jednotlivé výroky odpovídají jeho vlastnímu obrazu o sobě. Při druhém třídění je třídí podle své představy o tom, jaký by chtěl být, po třetí třeba podle toho, jak ho asi vidí jeho partner, rodiče, nadřízený atd. Podobných posuzovacích aspektů je možno zvolit celou řadu. Q-třídění je možné opakovaně užít v průběhu terapie, výcvikového kursu aj. Shoda mezi různými aspekty /třeba mezi reálným já a ideálním já, nebo mezi tříděním v různých fázích terapie/ se opět zjišťuje korelací.

Tab. XIII.1.

Distribuce položek při Q-třídění do 11 kategorií

Počet	Maximální souhlas						Maximální nesouhlas				
položek	10	9	8	7	6	5	4	3	2	1	0
90	3	4	7	10	13	16	13	10	7	4	3
80	2	4	6	9	12	14	12	9	6	4	2
70	2	3	5	8	11	12	11	8	5	3	2
	2	3	4	8	11	14	11	8	4	3	2

Druhý způsob užití je zaměřen na vyhledávání osob, které jsou navzájem podobné a tvoří relativně homogenní skupiny. Postupuje se tak, že jedna osoba posuzuje větší počet lidí a ze srovnání výsledků se odhaduje, zda posuzovaná skupina se rozpadá do několika podskupin či nikoliv. Jsou-li k dispozici odhady většího počtu posuzovatelů, pak je možno buď sledovat průměrné posouzení různých vlastností u jednotlivých posuzovaných osob, anebo se zaměří pozornost na posuzovatele a sleduje se stupeň podobnosti jejich třídění.

Q-třídění je možno užít i pro posouzení řady jiných jevů. Můžeme třeba třídit lidi do kategorií dle toho, jak jsou oblíbeni, můžeme třídit umělecká díla, třídit různé hodnoty podle toho, zda je přijímáme či nikoliv, třídit výroky, vyjadřující postoje atd.

Kritický přehled využití metody Q-třídění je možno nalézt ve studii Meehla /1960/ a Wittenborna /1961/. Celkově můžeme říci, že využití Q-techniky je nadějné a umožňuje formulovat ve výzkumu osobnosti otázky, které zatím nebyly řešeny převážně z důvodů metodologických. Zatím nelze jednoznačně říci, zda či případně za jakých okolností je Q-třídění výhodnější než prosté posuzovací škály či dotazníky. Je však jisté, že dobře odpovídá nejnovějším proudům v teorii osobnosti, které předpokládají, že prosazení se určité osobnostní tendence /třeba extraverze/ v určité konkrétní situaci je spoluurčováno i souhrou dalších vlastností osobnosti.

Otevřená zatím zůstává problematika statistického zpracování dat získaných Q-tříděním /viz též Gaito, 1962/. Guilford /1963/ navrhl určitou transformaci hodnot, získaných z Q-třídění, která umožňuje užívat faktorovou analýzu. Neparametrickou faktorovou analýzu pro podobná data vyvinul Kelly /1963/ a užil ji na rozbor vztahů mezi teoriemi psychoterapie a teoriemi osobnosti. Techniku stanovení vnitřní konzistence při Q-třídění vypracovali Neff s Cohenem /1967/.

Zřejmou výhodou Q-techniky je to, že při jejím použití nemůže působit tendence kumulovat příliš mnoho výroků ve středu škály, či posunovat hodnocení ke kladnému či zápornému pólu. I některé jiné typy konstantních chyb, které ovlivňují výsledky v subjektivních posuzovacích škálách běžného typu, jsou odstraněny. Q-technika se dá snadno opakovat a můžeme mnohem důkladněji a strukturovaněji sledovat změny v čase, než tomu bývá při pouhém sledování změn v průměrech. Hodí se mnohem více pro intenzivní studium jedince /či několika málo osob/ než pro extenzivní výzkumy s velkými výběry osob.

XIII.3. S é m a n t i c k ý d i f e r e n c i á l

Tato posuzovací technika bývá často srovnávána s Q-tříděním a uvádí se, že má podobné výhody, ale vyhýbá se jeho úskalím; bývá jí proto dávána přednost /Tent, 1970/.

Sémantický diferenciál je zdánlivě velmi jednoduchá metoda, která využívá grafické škály. Zavedl jej Osgood /1952/; teoretické východiska, metodologické předpoklady i řadu aplikací obsahuje monografie Osgood, Suci a Tennenbaum /1957/. V německé terminologii byl zaveden pojem dojmový diferenciál /Eindrucksdifferential - Ertel, 1965a, b/.

Sémantický diferenciál je sestaven z určitého počtu sedmibodových škál /Kerlinger, str. 550/. Na koncích jsou bipolárně zakotvena adjektiva, která vyjadřují určité vlastnosti. Škály vytvářejí sémantický prostor s neznámým počtem dimenzí, ale s eukleidovskými vlastnostmi. Každá škála je lineární a prochází počátkem tohoto prostoru. Dále je stanoven předmět či osoba, která má být posouzena. Od pokusné osoby se žádá, aby určený podnět postupně posoudila z hlediska všech předložených škál. Na každé z nich má značkou označit, na který bod sedmibodové stupnice by podnět umístila. Získá se tak kvantitativní hodnota daného podnětu v každé z uvedených škál. Po shromáždění dat od všech osob se vypočte průměr zjištěný v jednotlivých škálách a získá se tak celkový profil posuzovaného podnětu. Můžeme též zachycovat průměrné hodnocení, zjištěné u různých podskupin či u jednotlivých osob. Bližší informace o postupu při sestavování sémantického diferenciálu je možno najít i v učebnicích Kerlingera /1972, kap. 32/ či Krecha a kol. /1968/.

Srovnáme-li výsledky různých podskupin posuzujících osob, používá se často korelační koeficient Q, který vyjadřuje stupeň podobnosti mezi dvěma profily. V případě maximální podobnosti /tj. oba profily jsou paralelení/ získáváme hodnotu +1 ; při naprosté nepodobnosti /zrcadlový obraz obou profilů/ vyjde hodnota -1 .
Koeficient Q se vypočte takto:

$$Q_{AB} = 1 - \frac{\sum d_{AB}^{2} - k\,(\bar{X}_A - \bar{X}_B)^2 - (\sigma_A - \sigma_B)^2}{2\,.\,\sigma_A\,.\,\sigma_B}$$

kde d_{AB} je rozdíl hodnot, zjištěných v jednotlivých škálách,
$\bar{X}_A$, $\bar{X}_B$ jsou průměrné hodnoty získané u jednotlivých skupin posuzovatelů /či v jedné podskupině při posouzení 2 pojmů/,
σ_A, σ_B jsou směrodatné odchylky obou průměrů,
k je počet škál.

Postup si demonstrujeme na příkladě, který uvádí Hofstätter /1959/. Srovnával pojmy "osobnost" a "masa" a použil při tom sémantický diferenciál o 24 škálách. Výsledky, které získal od skupiny studentů, byly tyto: Průměrná hodnota, získaná při posouzení pojmu "osobnost", byla 4,3 , směrodatná odchylka 7,1 . Při posouzení pojmu "masa" získal $\bar{x} = 4,4$, $\sigma = 5,3$. Hodnota $\sum d_{AB}^2$ byla 107,30. Míra podobnosti mezi oběma pojmy tedy byla:

$$Q = 1 - \frac{107,30 - 24\,(4,3 - 4,4)^2 - (7,1 - 5,3)^2}{2 \,.\, 7,1 \,.\, 5,3} =$$

$$= 1 - \frac{103,82}{75,26} = -0,38$$

Výsledek naznačuje, že sledované pojmy jsou posuzovány protikladně. Z věcného rozboru jednotlivých škál je pak možno blíže analyzovat, v kterých položkách se od sebe výrazně liší a v kterých nikoliv. V tomto konkrétním výzkumu se však nutně vtírá otázka, zda získaný výsledek je zobecnitelný; vzhledem k výběru posuzovatelů /vesměs studenti/ je na místě určitá skepse.

Uvedený postup vyhodnocování výsledků nebývá příliš častý; používá se pouze tam, kde nemáme předem analyzovány a standardizovány jednotlivé škály a adaptujeme sémantický diferenciál pro určité konkrétní účely. Ve standardizované podobě sémantického diferenciálu se faktorovaly výsledky získané v jednotlivých škálách a zjistily se 3 základní dimenze: valence, potence, aktivita. Pod pojmem valence se v podstatě zahrnuje hodnocení /příklady škál: dobrý-špatný, cenný-bezcenný, čistý-špinavý apod./; zahrnuje zhruba

50% rozptylu. Pojem potence vyjadřuje v podstatě intenzitu působení sledovaného objektu /příklady škál: přísný - mírný, silný - slabý, dlouhý - krátký apod./; škály zahrnuté do tohoto pojmu vyčerpávají přibližně 25% variance. Pojem aktivita zahrnuje asi 20% variance /jako příklad můžeme uvést škály: rychlý - pomalý, aktivní - pasívní atd./. Sémantický diferenciál byl v této podobě standardizován, ale může být uzpůsoben k různým účelům tím, že se na základě rozboru situace některé škály přidají, jiné vyloučí apod. Při podobných úpravách však je třeba zjistit, zda provedené změny neovlivnily výrazně strukturu základních dimenzí.

Principiálně můžeme sémantický diferenciál užít pro posouzení jakéhokoliv podnětu, který je pokusné osobě znám. Mohou to být konkrétní osoby /otec, spolužáci,.../, profese, sociální skupiny, předměty /tato učebnice, dnešní oběd,.../, abstraktní pojmy /kázeň, studium,.../ aj.

Podobnost se v těchto částečně standardizovaných škálách měří většinou mírou vzdálenosti mezi sledovanými pojmy /přesněji mezi jejich faktorovou reprezentací/; tato hodnota se značí D_{ab} a vypočítá se ze vztahu

$$D_{ab} = \sqrt{\sum_j d_{ab}^2}$$

kde D_{ab} je lineární vzdálenost mezi pojmy a, b ;

d_{ab} je algebraická diference mezi souřadnicemi pojmu a , b v jedné dimenzi /či v určitém faktoru/.

Řekněme, že osoba X a osoba Y posuzovaly otce a matku, a objevily se následující výsledky, odpovídající faktorovým hodnotám /tab.XIII.2./. Pro oba posuzovatele vypočteme podobnost /pro zjednodušení výpočtu vycházíme ze zaokrouhlených hodnot/:

Tab.XIII.2.

Výpočet vzdálenosti mezi dvěma pojmy u dvou osob

	O s o b a X			O s o b a Y		
	Valence	Potence	Aktivita	Valence	Potence	Aktivita
Otec	+5	+3	+1	+3	-2	-2
Matka	+4	+2	-1	-3	+3	-4
d	1	1	2	6	5	2
d^2	1	1	4	36	25	4
$\sum d^2$		6			65	
$\sqrt{\sum d^2}$		2,45			8,06	

Čím větší je hodnota D, tím větší je vzdálenost mezi oběma pojmy pro daného posuzovatele. Není zatím stanovena závazná hranice významnosti pro podobnost sledovaných pojmů. Uvádí se, že při individuálních datech /jediný posuzovatel/ hodnota $D > 1,5$ znamená při užití výchozích škál již značnou nepodobnost. Při srovnávání faktorových skórů se uvádí hodnota $D > 2$. Při skupinových datech již hodnota $D > 0,5$ značí značnou nepodobnost. Jsou však i určité výhrady vůči tomuto postupu při stanovení výpočtu hodnoty D a diskuse není zatím uzavřena /viz též Nunnaly, 1962/. Proto je asi na místě pracovat s takto získanými výsledky jen pro srovnávací účely.

Byly navrženy i některé další formy sémantického diferenciálu /Coyne, Lolafaye, Holzman - 1966/, ale zatím nebyly dostatečně standardizovány. Důkladné statistické studie o zpracování dat získaných sémantickým diferenciálem publikoval Orlik /1965, 1967/. Prokázal, že faktorování korelací mezi profily vede k pseudofaktorům, které nemají žádný význam, a navrhl proto jiné postupy zpracování. Velmi podstatným příspěvkem ke standardizaci celého postupu jsou studie Ertelovy /1965 a, b/, který se důkladně zamýšlel nad výběrem škál a dospěl k šesti škálám pro každou dimenzi. Škály volil osmibodové a zakotvil je vždy adjektivně.

Celkově se sémantický diferenciál jeví jako všestranně použitelný výzkumný nástroj. Jeho vývoj po stránce metodologické není ještě dokončen, ale i přesto může být již dnes užíván, byť i s nezbytnou kritičností. Jeho použití je velmi rozsáhlé - v oblasti výzkumu osobnosti, v klinické psychologii, v poradenství i v průmyslové psychologii. Často se užívá při výzkumu v sociální psychologii, ale s úspěchem byl aplikován i ve forenzní a v pedagogické psychologii. Má nesporně výhody i v tom, že určitý pojem posuzujeme z mnoha hledisek, přitom však celkový postoj je vysuzován ze souhry či rozporů v parciálních hodnoceních. Tím se dociluje globálnějšího pohledu, než tomu bývá ve většině posuzovacích stupnic. Naproti tomu je nesporně zatížen chybami, které se vyskytují u všech grafických posuzovacích stupnic. Při nevhodně zvolených škálách může dojít k značným artefektům. Tedy i v tomto případě - tak, jako ve většině ostatních - platí, že zdánlivě jednoduchá metoda si vyžaduje dlouhou, pečlivou a kritickou přípravu.

XIII.4. Posuzovací škála se subjektivním zakotvením

Zmíním se o jedné velmi specifické variantě posuzovacích stupnic, která je na rozdíl od ostatních určena pouze pro sebeposouzení. Jedná se o postup navržený Kilpatrickem a Cantrilem /1960/. Autorům šlo o vytvoření takové metody, která by zachytila reálný obraz o sobě a o svém stavu tak, jak jej člověk v určitém časovém bodě percipuje. Škála není slovně formulována, aby názor pokusné osoby nebyl nijak ovlivněn použitými výrazy. Vytváří se teprve v průběhu pokusu, a to na základě rozhovoru s pokusnou osobou a z obsahového rozboru jejích očekávání, plánů, vzpomínek, minulých či aktuálních činností atd.

Pokusná osoba je nejprve vyzvána, aby popsala, jaký by pro ni byl ideální způsob života. V této exploraci se postupuje tak dlouho, až osoba nemá již co dodat. Její výroky se pokud možno doslovně protokolují. Obdobně se zjišťuje, co by pro ni bylo nejhorším způsobem života; i zde se provede detailně explorace a výroky se zachytí.

Pak se pokusné osobě předloží grafická neverbální škála s body 0 - 10 /většinou v podobě žebříčku, ve kterém je bod 10 na nejvyšší příčce a bod 0 je pod nejnižší příčkou/ a vyloží se, že ideální způsob života, tak jak jej osoba popsala, odpovídá číslu 10 a nejhorší způsob odpovídá číslu 0. Tím je subjektivní škála zakotvena a připravena k užití. Můžeme se ptát, kde se na tomto žebříčku člověk podle svého dojmu právě teď nalézá. Tím se získá třetí bod na škále /aktuální stav/. Podle potřeby či cíle šetření se můžeme ptát dále; třeba: kde jste byl na tomto žebříčku před 2 lety, jakou pozici očekáváte za 5 let, kde jste byl na žebříčku před svatbou, kde budete po dokončení studia atd. Zjistí-li se rozdíly v uváděné poloze, je možno sledovat, v čem se způsob života změnil vzhledem k minulému stavu, či co se očekává od budoucnosti. Cantril /1963/ použil škálu se subjektivním zakotvením i pro studium aspirací.

Autoři této techniky uvádějí, že se objevily srovnatelné nálezy v různých společenských podmínkách /u Afričanů, Japonců atd./. Je ovšem otázka, jak dalece se jedná o spolehlivý a reliabilní prostředek; zde chybí uspokojivá data. Je však jistě možné tuto subjektivní posuzovací škálu používat, a to buď ve srovnání s jinými prostředky /se sémantickým diferenciálem/, anebo pro srovnání subjektivního posuzování v různých časových úsecích /před a po terapii apod./. Můžeme přitom zjišťovat rozdíly nejen v udávané poloze na škále, ale i v jejím subjektivním zakotvení.

Celkově se jedná o metodu, která má význam při intenzívních studiích jednotlivých osob a svým charakterem není určena pro zobecňující výzkumy se širším výběrem pokusných osob.

XIII.5. S o c i o m e t r i e

Sociometrické posuzovací techniky jsou velmi populární a často užívané v sociální psychologii. Je to metoda již poměrně stará;

jejím zakladatelem je J.L. Moreno, který ji jednak detailně propracoval, jednak svým typickým způsobem popularizoval a dokonce vytvořil velmi jednostranný kult této techniky jako univerzálního výzkumného, ale i nápravného prostředku. Odhlédneme-li od tohoto extrémního stanoviska, zůstane racionální jádro, které je dobře použitelné. Sociometrickou technikou se zjišťují vzájemné vztahy mezi osobami v malé sociální skupině. Jednotliví členové skupiny se navzájem posuzují, a to většinou z hlediska vzájemných sympatií či antipatií. Od všech členů skupiny se žádá, aby určili, koho ze skupiny by si zvolili jako spolupracovníka, jako spolubydlícího v koleji, koho by určili jako nejnadanějšího atd. Může se případně zaznamenávat i negativní posudek, tj. určení toho, s kým by člověk nechtěl spolupracovat atd. Z výsledků se sestaví sociogram /buď ve formě grafu či matice/ a dále se pak zpracovává buď pouhým náhledem anebo pomocí přesnějších statistických metod.

Pro práci se sociometrickou technikou stanovil Moreno /1953/ šest základních podmínek, a to:

a/ posuzovaná skupina osob musí být přesně určena a pozitivní či negativní volby se omezují jen na její členy;

b/ lze vybírat či odmítnout libovolný počet členů;

c/ je třeba stanovit přesné kritérium, podle kterého se vybírá;

d/ získaná data se využívají záměrně k restrukturalizaci skupiny a osoby jsou o tom informovány;

e/ pozitivní i negativní volby je možno provádět neveřejně;

f/ kritéria výběru, která jsou stanovena v podobě sociometrických otázek, musí být snadno srozumitelná pro všechny členy skupiny.

Ve skutečnosti se užívá mnoha variant či úprav původních Morenových podmínek, a to hlavně v bodě b/. Nejčastější změny bývají tyto: Počet voleb je omezen /většinou jen na jednu osobu hodnocenou pozitivně a jednu negativně/. Při volbě většího počtu osob se uvádí jejich pořadí, při menším počtu se může případně provést i párové srovnání. Členové skupiny se neznají a sleduje se postupné vytváření jejich vztahů. Byly vyvinuty techniky pro postupný výběr osob, takže je možno podrobit sociometrickému šetření i velké skupiny osob /obyvatele určité obce apod./. Používá se i zpětná sociometrie, při které každá pokusná osoba má odhadnout, jak byla posouzena ostatními; zpětných variant bylo navrženo několik a užívají se ke srovnání odhadu vlastní pozice ve skupině se skutečností. Pokusná osoba může mít k dispozici určitý počet bodů a libovolně je přidělí ostatním členům; může je třeba všechny přidělit osobě, kterou volí na prvním místě, nebo je v určitém poměru rozděluje několika osobám atd. Někdy se kombinuje sociometrická volba s tříděním do kategorií. Pokusná osoba zvolí např. 3 členy skupiny, kteří jí stojí nejblíže, a ostatní utřídí do 3 kategorií: také bych ho přijal - neutrální - nepřijal bych ho.

Původní sociometrická šetření se omezovala na grafické znázorňování vztahů sociogramem; postupem doby se však vypracovaly podrobnější a objektivnější techniky. V současné době existují zhruba čtyři cesty pro analýzu sociometrických dat:

a/ Zdokonalování sociogramů detailnějšími grafickými metodami.

b/ Vytváření různých indexů, kterými se popisují získaná data jednak pro zachycení pozice jednotlivců, jednak pro charakterizování celé skupiny, případně jejích podskupin.

c/ Statistické techniky, kterými se zjišťuje, zda počty pozitivních či negativních voleb je možno považovat za náhodné či nikoliv.

d/ Maticová algebra, aplikovaná na sociometrická data.

Literární přehled i základní informace o uvedených metodách lze najít ve studii Tentové /1970/. V ní je zachycen i současný stav poznatků o reliabilitě a validitě sociometrických údajů. Za zmínku též stojí snaha aplikovat sociometrii i ve výzkumu velkých skupin osob; tyto úpravy původní techniky jsou popsány ve studiích Rapoport, Horwath /1961/ a Foster, Horwath /1971/.

Sociometrickou technikou se nebudu detailněji zabývat, protože jsou k dispozici důkladné studie jak původní / Petrušek 1969/, tak přeložené /Kerlinger, 1972, kap. 31/, ve kterých jsou podrobně uvedeny přednosti i slabiny této techniky.

Kapitola XIV.

ÚVOD DO MULTIDIMENZIONÁLNÍHO ŠKÁLOVÁNÍ

XIV.1. Ú v o d

Při práci se škálovacími technikami, které jsme zatím probírali, jsme většinou předpokládali, že podněty, které sledujeme, mají určitou vlastnost, podle které je můžeme posuzovat: uchazeč A je lepší než uchazeč B, podnět C je dvakrát tak chutný než D, situace E je riskantnější než situace F či G atd. Dále se předpokládá, že uvažovaná vlastnost je sledována z hlediska jediné dimenze, a právě proto můžeme srovnávané podněty jednoznačně seřadit. V řadě případů jsme však museli konstatovat, že tomu tak není. Tak jsme se třeba setkávali s netranzitivními volbami při párovém srovnávání ($A > B$, $B > C$, $C > A$) či s položkami, které nebylo možno užít v škálogramu. Tyto nepravidelnosti či přímo i rozpory svědčí o tom, že předpoklad jediné dimenze, podle které se škáluje, může být mnohdy zjednodušující či dokonce přímo neplatný. Od poloviny padesátých let pozorujeme proto snahu vytvořit nové škálovací techniky, které by nebyly vázány předpokladem existence jediné posuzované dimenze.

Multidimenzionální škály jsou v podstatě rozšířením dříve pojednávaných škálovacích metod na situace, kdy posuzované podněty se mohou současně měnit ve dvou či více dimenzích. Typická situace, která je řešena multidimenzionálním škálovánám, může být vyjádřena takto: je dána množina podnětů, které se od sebe liší neznámým počtem dimenzí. Úkolem je určit: a/ minimální počet těchto dimenzí, b/ umístění podnětů na každé ze stanovených dimenzí. Modely multidimenzionálních škál nahrazují jednodimenzionální kontinuum koncepcí vícedimenzionálního prostoru. Podněty jsou chápány jako body v tomto prostoru, a jsou proto charakterizovány větším počtem údajů /shodným s počtem dimenzí prostoru/.

Základním problémem multidimenzionálního škálování je určení počtu dimenzí. Velký počet dimenzí vede k nepřehledným vztahům mezi podněty, a proto je plně pochopitelná snaha o jejich minimalizaci. Zároveň však vzniká nebezpečí, že přílišná redukce povede k nemístnému zjednodušení skutečnosti. Přesto se však v řadě studií mnozí autoři spokojují se škálami o dvou či třech dimenzích, které jsou dobře interpretovatelné /viz např. Osgoodův sémantický diferenciál, ve kterém se pracuje se třemi dimenzemi/. Interpretace jednotlivých dimenzí je základní otázkou, pokud usilujeme o smysluplný výklad dat, získaných pomocí multidimenzionálního škálování. Situace je zde do jisté míry analogická jako při interpretaci faktorů ve faktorové analýze.

Vlastní vývoj multidimenzionálního škálování začíná v padesátých letech a šíří se ze tří center. První skupina autorů byla inspirována Torgersonem, jehož článek v Psychometrice z r. 1952 se dnes považuje za klasický; 11. kapitola Torgersonovy učebnice psychologického škálování /1958/ byla dlouho považována za základní literární pramen. Druhá skupina se soustřeďuje kolem Coombse. V poslední době je však nejsilnější skupina, která se vytvořila v šedesátých letech kolem Sheparda a Kruskala. Síla této skupiny je kromě jiného i v tom, že většinu svých úvah řeší v souvislosti s vytvářením výpočetních programů pro samočinné počítače. Kromě toho se téměř vždy opírají o empirická data z reálných výzkumů a řeší nejen technické škálovací problémy, ale zároveň se zabývají i systematickým sledováním, zda získané dimenze skutečně a reálně odpovídají výchozím datům. Nebezpečí formalismu je zde značně sníženo.

Literatura, pojednávající o multidimenzionálním škálování, je v poslední době stále četnější. Kromě zmíněné kapitoly z Torgersonovy učebnice existuje též instruktivní úvod ve studii Bealse a kol. /1968/. Pravidelné - ale pro studijí účely stručné - referáty o vývoji multidimenzionálních škálovacích matod jsou uveřejňovány ve čtyřletých intervalech v Annual Review of Psychology /Ekman, Sjöberg, 1965; Zinnes, 1969/. Přehled současného stavu poznatků po stránce technické i aplikační je možno nalézt v dvousvazkovém sborníku, který redigovali Shepard, Romney a Nerlove /1972 a, b/. Řada příspěvků - většinou velmi důkladně propracovaných pro stránce matematické - vychází v časopisech Psychometrika a Journal of Mathematical Psychology. Některé metody multidimenzionálního škálování se blíží faktorové analýze /či některým analogickým metodám, jako je trsová analýza, hierarchická analýza apod./; jejich vzájemnými vztahy se zabýval MacCallum /1974/.

Při pohledu na současnou literaturu o multidimenzionálním škálování se nemůžeme zbavit některých pochybností. V prvé řadě je dnes situace značně nepřehledná; vyvíjí se řada metod, které se vzájemně překrývají a bez přímé zkušenosti s empirickým materiálem si jen těžko činíme obrázek o jejich hodnotě a užitečnosti. Použitý matematický aparát bývá pro psychologa dosti obtížný a bez spolupráce se statistikem se neobejdeme.

V empirických studiích se často setkáváme s názvem použité metody /např. Kruskalův M-D-SCAL program, Carrolův INDISCAL aj./ a obtížně získáváme informace, o jakou metodu se jedná. Řada multidimenzionálních škálovacích postupů je již převedena do programů pro samočinný počítač; program dostane své jméno a číslo a pod touto šifrou je dále uváděn. Podrobnější orientace, nezbytná pro pochopení i smysl použitého postupu, je pak velmi těžká. Snad i z těchto důvodů bývají dnes multidimenzionální škálovací postupy užívány poměrně málo.

Další důvod pro určitou skepsi k současnému stavu vývoje vyplývá i z rozporu, který je již na první pohled nápadný. Na jedné straně se setkáváme s řadou metod, které jsou po stránce matematické důkladně propracovány a často i axiomatizovány. Naproti tomu však jejich využití v empirickém výzkumu - ať již základním či aplikovaném - není příliš přesvědčivé. Jedná se často o studie, které nejsou příliš imponující ani svou problematikou ani zvolenou výzkumnou metodou. Jindy nacházíme pouhé ilustrativní pokusy, které mají čtenáře přesvědčit, že vypracovaná škálovací technika by mohla být pro psychologii přínosná.

XIV.2. D r u h y   d a t ,   v h o d n ý c h   p r o   m u l t i d i m e n z i o n á l n í   š k á l o v á n í

Při výkladu metod multidimenzionálního škálování se znovu vynoří otázka, jaké druhy dat máme k dispozici. Jedno z řešení této metodologicky základní otázky jsme vyložil v souvislosti s Coombsovou teorií. Shepard /1966, 1972/ z této koncepce vychází a rozšiřuje ji o další kategorie. Rozeznává v podstatě 4 třídy dat:

a/ proximitní data, při kterých se zjišťuje blízkost či podobnost jednotlivých podnětů,

b/ dominanční data, při kterých se sleduje, kterému podnětu se dává přednost a je proto výše hodnocen,

c/ profilová data, při kterých se řada podnětů srovnává z více hledisek,

d/ data získaná spojeným měřením, při kterém se posuzuje určitý jev, který je definován dvěma současně se měnícími proměnnými.

XIV.2.a. P r o x i m i t n í d a t a

Tato data mají několik forem:

a/ Čtvercová matice n x n , ve které řádky i sloupce odpovídají n-podnětům. Každé pole matice obsahuje údaj, který vyjadřuje podobnost /případně asociaci, korelaci, interakci, nahraditelnost atd./ mezi patřičnou dvojicí podnětů. Podobnost může být přímo odhadována /např. přímé posouzení podobnosti 2 jevů, odhad stupně pozitivních vztahů mezi různými osobami či skupinami atd./, anebo může být odvozena numericky na základě jiných dat /třeba z dat profilových: podněty jsou navzájem srovnávány z mnoha hledisek a úkolem je pak tyto rozdíly sloučit do jediného ukazatele/. Pojem podobnosti je zde chápán velmi široce a obecně; můžeme sem zahrnout i posouzení vzdálenosti anebo odlišnosti. Někdy se pracuje i s několika proximitními maticemi současně /posouzení stejných n-podnětů za různých podmínek či různými skupinami osob/.

b/ Obdélníková matice n x m , ve které řádky a sloupce odpovídají různým podnětům. Jedná se o porovnání jednoho podnětu z n-členné množiny s jedním podnětem z jiné m-členné množiny. Podobnosti mezi podněty uvnitř jednotlivých množin se nesledují. Jako příklad můžeme uvést srovnávání mezi n-muži a m-ženami z hlediska sympatií či jiných aspektů, srovnávání různých předmětů, vybraných ze dvou typů spotřebního zboží apod. Je jistě možné, aby i za těchto podmínek vznikla čtvercová matice /jestliže m=n/, ale liší se podstatně od matice uvedené sub a/. Srovnávání dat je zde možné pouze v rámci jednotlivých řádků či sloupců výchozí matice, ale nemá smyslu mimo ně. Tato data se pak mohou zpracovávat Coombsovou technikou rozvíjení.

c/ Mohou se objevit i některé další formy proximitních dat, které vycházejí ze srovnávání pouze určitého počtu párů. Tak např. Torgerson /1952/ navrhl metodu triád. Pracuje se s třemi podněty, které se párově srovnávají. Sleduje se relativní frekvence, s jakou byly jednotlivé dvojice podnětů označeny jako navzájem nejvíce podobné. Postupně sa analyzují všechny triády, které je možno z daného počtu podnětů vytvořit.

Vlastní metody, které se při škálování dat tohoto typu užívají jsou několikeré: metrické multidimenzionální škály, nemetrické multidimenzionální škály /zde existuje celá řada publikovaných programů pro počítače/, metoda hierarchických trsů, multidimenzionální škálování individuálních diferencí /tato metoda se v poslední době uvádí jako velmi slibná - berą totiž v úvahu i to, že data získaná od jednotlivých osob mohou být odlišně rozložena podle os - blíže o tom v Carroll a Chang, 1970; Carroll, 1972/.

XIV. 2.b. D o m i n a n č n í   d a t a +/

Tato data se mohou vyskytovat ve dvou různých formách:

a/ Jediná čtvercová matice n x n , ve které řádky a sloupce odpovídají stejným podnětům. V jednotlivých polích je zachyceno, do jaké míry je podnět na patřičné řádce preferován, výše hodnocen, či jinak dominuje nad podnětem v patřičném sloupci. Patří sem např. párové srovnávání podnětů určitou skupinou osob, údaje o tom, jak často jeden hráč vítězí nad druhým ve hře či v jiné formě sociálních kontaktů apod. V této podobě se často předpokládá jednodimenzionální uspořádání podnětů; pouze při složitějších vztazích /mnoho kruhových triád/ se uvažuje o vícedimenzionálním řešení.

b/ Množina m čtvercových matic, ve kterých řádky a sloupce odpovídají stejným n podnětům. Každá matice je pak získána za odlišných podmínek /např. od různých osob, pod vlivem různé instrukce apod./.

Uvažuje se též o možnosti pracovat s pravoúhlými maticemi m x n , ve kterých řádky a sloupce by odpovídaly podnětům vybraným z různých množin, ale zatím tato možnost nebyla zpracována.

+/Terminologická poznámka:
Dominanční data bývají často v literatuře uváděna jako párové srovnávání. Tento termín by je však neodlišoval od proximitních dat, které jsou většinou získána též párovým srovnáváním. Proto termín dominanční data je vhodnější.

Dominanční data jsou zpracovávána buď tak, jak jsme je popsali v kap. III. pro případ jednodimenzionální škály /s určitými úpravami pro situace, kde teoreticky očekáváme normální rozložení dat/, anebo se použije některá z forem multidimenzionální analýzy mnohonásobných dominančních matic /stručný přehled viz Shepard, 1972/.

XIV.2.c. Profilová data

Objevují se vesměs v obdélníkové matici m x n , kde řádky odpovídají n-podnětům a sloupce m-proměnným /nebo naopak/. V jednotlivých polích je zaznamenána zjištěná hodnota určitého podnětu vzhledem k jednotlivým proměnným. Přitom hodnoty m proměnných pro určitý podnět tvoří jeho profilovou charakteristiku. Jako příklad můžeme uvést skóre v každém z m subtestů u n osob, posouzení několika skupin s ohledem na m proměnných, skubjektivní posouzení n osob /spolupracovníků, podřízených apod./ z hlediska m posuzovacích kritérií atd. Přitom je možno podle účelu šetření zaměnit podněty a proměnné /můžeme sledovat n osob z hlediska proměnných, jako je pracovní iniciativa, dochvilnost, pracovní návyky atd., anebo můžeme stejná data posuzovat opačně, tj. sledovat, jak jsou pracovní iniciativa, dochvilnost, atd. rozloženy u n osob/.

Profilová data bývají nejčastěji zpracována pomocí klasické faktorové analýzy, pro kterou dnes existuje celá řada programů; setkáváme se zde s úzkou vazbou mezi multidimenzionálním škálováním a faktorovou analýzou. V některých případech multidimenzionální škálování může nahradit faktorovou analýzu a z tohoto hlediska se objevují i nové pokusy o reformulaci problematiky rotování ve faktorové analýze /viz Jöreskog 1969/. Existují dále různé nonmetrické formy faktorové analýzy, rozpracovávané hlavně Kruskalem a Shepardem. Jinou technikou jsou různé podoby multidimenzionální škálogramové analýzy dle Guttmana a Lingoese /viz Lingoes, 1963, 1972/. Jistou novinkou je i využití multidimenzionálního škálogramu pro kvantitativní data /Lingoes, 1968/. Pro zpracování profilových dat se používá též trsová analýza, která předpokládá, že jednotlivé profily je možno zařadit do vzájemně se vylučujících a vyčerpávajících kategorií tak, že jakékoliv dva profily patří /či nepatří/ do jedné kategorie tehdy, když jsou si dostatečně podobné /či nepodobné/. Trsová analýza může mít i hierarchickou formu /viz Wallace, 1968/.

XIV.2.d. D a t a z í s k a n á s p o j e n ý m m ě ř e n í m

Tato data se vyskytují v podobě obdélníkové matice $m \times n$, ve které řádky odpovídají n úrovním jedné proměnné a sloupce odpovídají m úrovním jiné proměnné. Každé pole reprezentuje údaj, který vzniká při současném působení obou proměnných na úrovni, dané patřičnou řádkou a sloupcem. Jako příklad lze uvést posuzování různých her či sázek, spojených s různou velikostí a s různou pravděpodobností výhry /či ztráty/, posuzování velikosti objektu při jeho různé vzdálenosti a při různé viditelnosti apod. Tuto formu měření zavedli Luce a Tukey /1964/ a rozšířil ji dále Krantz /1964/. Jejich snahou bylo najít jednodimenzionální škálu pro n úrovní jedné proměnné a zároveň druhou jednodimenzionální škálu pro m

úrovní druhé proměnné tak, aby hodnoty v jednotlivých polích matice byly ve vztahu ke konečným škálovacím hodnotám dle pravidel co možná nejjednodušších. Tento typ dat je možno rozšířit i na matice vícerozměrné, ve kterých se současně sleduje souhra tří či více proměnných, působících na sledovaný proces.

Pro zpracování těchto dat byla v posledních letech navržena řada metod, které umožňují vytvořit jednoduché škálové vyjádření vlivu dvou či více proměnných na posuzovaný podnět. Většina vyvinutých postupů je již zpracována v programech pro počítače.

XIV.3. Základní modely multidimenzionálního škálování

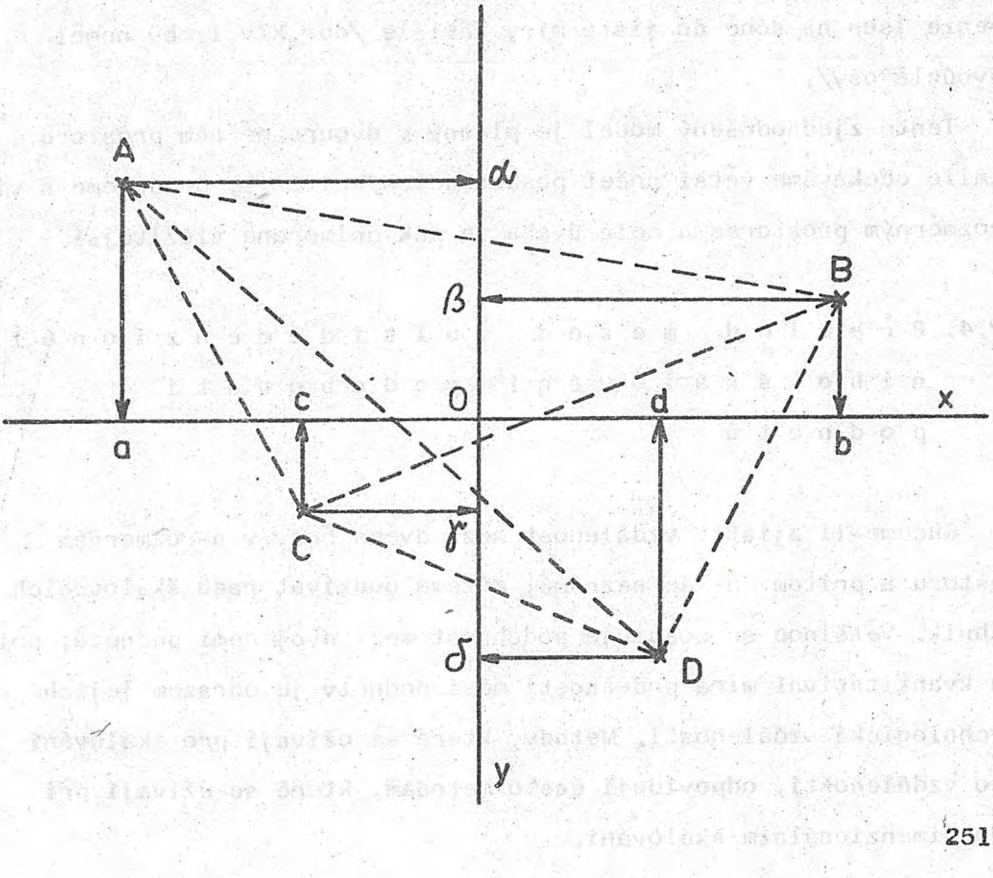
Předpokládejme, že posuzujeme čtyři žáky A, B, C, D a že pro posouzení přicházejí v úvahu dvě dimenze: píle a znalosti. Jako další zjednodušení zavedeme předpoklad, že obě posuzovací dimenze jsou na sobě nezávislé. Rozdíly mezi posuzovanými žáky jsou v obou dimenzích a znázorníme si je na obr. XIV.1.

Jestliže na ose x je zachycen stupeň znalostí jednotlivých žáků, pak jejich pořadí by bylo A, C, D, B /je dáno promítnutím souřadnic těchto bodů na osu x/. Z hlediska píle /osa y/ dostaneme pořadí jiné: A, B, C, D.

Ve skutečnosti však bývá situace složitější. Posuzujeme dané podněty, ale posuzovací dimenze buď nejsou zcela jasné, nebo je nemůžeme přesně stanovit, anebo je posuzovatelé nemohou od sebe oddělit. K dispozici jsou pak většinou údaje o vzdálenosti mezi jednotlivými podněty / v_{ij} /. Při čtyřech podnětech je zřejmě možných šest vztahů /AB, AC, AD, BC, BD, CD/; obecně je počet možných vztahů mezi podněty dán vzorcem $n(n-1)/2$.

OBR. XIV. 1.

ŠKÁLOVÁNÍ 4 PODNĚTŮ V DVOUROZMĚRNÉM PROSTORU.



Vzdálenosti mezi jednotlivými body reprezentují stupeň jejich podobnosti /pokud bychom chtěli zjišťovat vazby dominanční či jiné typy multidimenzionálních dat, pak by byla úvaha vedena poněkud jinak/. I když posuzovatel není instruován, jak má posuzovat, předpokládáme, že bude zvažovat obě dimenze. Vzdálenost mezi oběma body by se měla shodovat s mírou podobnosti, kterou posuzovatel určil. Je ovšem možné, že jedna z dimenzí bude hodnocena více než druhá a ovlivňuje konečný úsudek výrazněji; pak bychom museli obě dimenze vážit různými indexy. Jiná komplikace nastává, když sledované dimenze jsou na sobě do jisté míry závislé /obr.XIV.1. by neměl pravoúhlé osy/.

Tento zjednodušený model je platný v dvourozměrném prostoru. Jakmile očekáváme větší počet posuzovacích kritérií, pracujeme s vícerozměrným prostorem a celá úvaha je pak přiměřeně složitější.

XIV.4. Přehled metod multidimenzionálního škálování podobnosti podnětů

Chceme-li zjistit vzdálenost mezi dvěma body v n-rozměrném prostoru a přitom n je neznámé, můžeme používat řadu škálovacích technik. Většinou se posuzuje podobnost mezi dvojicemi podnětů; přitom kvantitativní míra podobnosti mezi podněty je odrazem jejich psychologické vzdálenosti. Metody, které se užívají pro škálování této vzdálenosti, odpovídají často metodám, které se užívají při jednodimenzionálním škálování.

Řada metod, které se při multidimenzionálním škálování používají, byla naznačena v části XIV.2. Podrobnější popis některých z nich lze najít v monografiích Sheparda a kol. /1972/, Coombse /1964/ či Torgersona /1958/; mnohé další pak ve specializovaných časopiseckých studiích. Pro orientaci některé z používaných metod stručně naznačím; soustředím se pouze na ty metody, které vycházejí z posouzení podobnosti mezi podněty.

Metoda triádových kombinací. Pokusným osobám jsou podněty předkládány v trojicích; úkolem je určit, které dva podnětv jsou si nejvíce podobné a které nejméně. Tak například při posuzování podnětů A, B, C posuzovatel konstatuje: A a C jsou si nejvíce podobné; B a C si jsou nejméně podobné. Z těchto tvrzení je možno odvodit: A je podobnější C než B; B je podobnější A než C; C je podobnější A než B. Postupně se vytvoří všechny možné trojice podnětů; celkem jich je $n(n-1)(n-2)/6$. Zjistí se pak počet případů, ve kterých podnět i je podobnější kterémukoliv z ostatních podnětů než podnět j /existuje i několik jiných postupů/. Z výsledných dat se sestaví matice, která se zpracovává analogicky jako data, získaná metodou párového srovnávání.

Jistou analogií tohoto postupu je i metoda úplných triadických kombinací. Při této variantě se nevysuzují vztahy mezi jednotlivými členy triády jako v předchozím postupu, ale získávají se přímo. Každá triáda je předložena třikrát se žádostí, aby pokusná osoba stanovila, zda podnět A je podobnější podnětu B či C; zda podnět B je bližší podnětu A či C, a konečně zda podnět C je bližší podnětu A či B. Tento postup je jednoduchý, přirozenější, a umožňuje zachytit i určité nepravidelnosti v posuzování /kruhové vazby/.

Metoda multidimenzionálních pořadových stupnic. Od pokusné osoby se žádá, aby k danému podnětu i přirovnala všech ostatních n - 1 podnětů a seřadila je do pořadí co do podobnosti k podnětu i. Tento postup se provede se všemi podněty a od každé osoby se tedy získá celkem n pořadových stupnic. Při srovnávání dvou podnětů se zjišťuje, který z nich je častěji posuzován jako podobnější s některým z ostatních podnětů.

V uvedených postupech se data získávají sice jinou experimentální cestou, ale jejich zpracování je stejné. Získávají se hodnoty $_kp_{ij}$, tj. proporce případů, kdy podnět k je posuzován jako podobnější podnětu i než podnětu j. Příklad numerického postupu je uváděn Torgersonem /1958, str. 263 a dále/.

Jinou metodou, která je jistou analogií metody párových srovnávání, je metoda čtveřic /tetrád/. Vytvoří se všechny párové kombinace čtyř prvků a pak se srovnávají postupně vždy dva páry a zjišťuje se, který z nich je tvořen podobnějšími prvky. I z těchto dat je možno vytvořit matici, která se zpracovává analogicky jako při metodě párového srovnávání.

Multidimenzionální metoda stejných intervalů se v poslední době začíná častěji uplatňovat při výzkumu postojů. Srovnávají se dvě tvrzení a odhaduje se míra jejich shody /např. v diskusi pan A tvrdil "......" a pan B tvrdil"......" Jak dalece se spolu shodují?/. Stupeň shody je pevně zakotven v počátku /úplná shoda/. Shoda mezi různými páry výroků je pak odhadována buď na škále s určitým počtem definovaných kategorií či podle některé z jiných forem posuzovacích škál. Průměrné hodnoty /případně hodnoty mediánu/ jsou interpretovány jako psychologická vzdálenost dvou daných tvrzení.

Multidimenzionální analogie metody následných intervalů bylo použito jak při sledování postojů, tak při zkoumání různých senzorických podnětů. Podnět A se vezme jako standard; může to být třeba výrok, odpovídající určitému postoji. Předpokládejme, že tento postoj zaujímá osoba X a úkolem pokusné osoby je odhadnout, jak dalece bude tato osoba X souhlasit s ostatními výroky. Jako standard se postupně použijí všechny podněty; znamená to tedy, že každý pár podnětů je během doby posuzován dvakrát na kontinuu, které vyjadřuje shodu.

XIV.5. Z á v ě r

Potřeba multidimenzionálního škálování je v současné psychologii zřejmá. V řadě oblastí psychologie pracujeme s podněty, které jsou velmi komplexní, a většinou se neuvažuje o tom, zda by je bylo možno blíže analyzovat a stanovit lépe jejich dimenzionální strukturu. Při podobném analytickém přístupu se využití multidimenzionálních škálovacích metod přímo nabízí. Po určitou dobu byla zábranou numerická složitost některých postupů; samočinné počítače tuto překážku odstranily.

Rozvoj multidimenzionálního škálování je velmi rychlý, snad až překotný a někdy odtržený od reálných výzkumů. Jedná se o problematiku, která je složitá a mnohdy i těžko přehledná. Stejně tak, jako je možno multidimenzionálně škálovat podněty, bylo by možno škálovat i výsledky jednotlivých posuzovatelů a asi by se objevily i některé interakce. Z hlediska složitosti postupů i myšlenkových základů celé problematiky se v poslední době uvažuje o vztazích mezi těmito metodami a faktorovou analýzou - ať již v její metrické či nonmetrické podobě. Z kombinace těchto postupů je možno očekávat v budoucnosti metodicky přínos /viz MacCallum, 1974/. Společný rys obou metod je hledání základních dimenzí, na které je možno redukovat výchozí data. Právě ve sblížení obou metod /hlavně těch forem faktorové analýzy, které rozvíjí Guttman, 1966/ můžeme očekávat další vývoj.

Informace o multidimenzionálním škálování, která je zachycena v této kapitole, je jen velmi schematická a pouze naznačuje způsoby uvažování a metodické postupy. Bude užitečné se této problematice systematicky věnovat. Předpokladem bude vytvoření mezioborových týmů pracovníků, kteří by se otázkami využití multidimenzionálního škálování v psychologii systematicky zabývali a získali dostatečné zkušenosti z vlastních experimentálních studií.

Kapitola XV.

SROVNÁVACÍ METODOLOGIE PŘI PSYCHOLOGICKÉM ŠKÁLOVÁNÍ

XV. 1. Ú v o d

V závěrečné kapitole pojednám stručně o některých otázkách srovnávací metodologie a její aplikace v psychologickém škálování. V posledních letech můžeme sledovat vlnu diskusí o základních metodologických principech celé psychologie. Podobné diskuse se objevují vždy v určitých časových intervalech. Většinou signalizují, že obvykle užívaná výzkumná strategie se vyčerpala a začíná brzdit další vývoj vědeckého poznání. Současná psychologie získala ohromné množství dat a dílčích výsledků; kritický rozbor však vede k závěru, že používané výzkumné postupy jsou většinou jednostranné či neúplné. Mohou sice podnítit mnohé další výzkumy, ale pro pokrok našeho poznání již nepřinesou zásadně nové pohledy či teoretické koncepce. Postupně se však objevují nové metodologické postupy, které umožňují stavět nové otázky i vytvářet nové teoretické systémy. Dialektický vývoj vztahů mezi vědeckou metodou a získanými daty, o něž se opírají teoretická zobecnění, má své zákonitosti, které se objevují s větší či menší pravidelností nejen v jednotlivých vědních odvětvích, ale i v různých oblastech společenské praxe.

V dnešní psychologii se stále více diskutuje o problematice zobecňování získaných experimentálních dat. Obrovská masa dílčích poznatků, jejichž množství narůstá se stále větší akcelerací, vedla sice k vyvrácení řady starších rozsáhlých teorií, ale neumožnila vytvořit nové zobecňující systémy. Empirický a ve svých gnoseologických kořenech pozitivistický přístup k výzkumné praxi vedl mnohé autory ke značné skepsi vůči jakémukoliv teoretickému zobecňování poznatků. Naproti tomu v psychologické praxi se zcela mechanicky a nekriticky zobecňují nálezy z jednotlivých dílčích studií na celou populaci či na rozsáhlou množinu jiných situací, které se často značně liší od původního výzkumného projektu. Podobný stav se objevuje v inženýrské psychologii, v psychodiagnostice, v psychoterapii i v jiných psychologických oblastech. Bylo by možno uvádět desítky příkladů neúměrného zobecnění dílčích nálezů, které často proniká až do učebnic, a ve svých důsledcích zkresluje pohled na současný stav psychologických poznatků.

Často se dnes hovoří o externí validitě psychologických nálezů; rozumí se tím zobecnitelnost poznatků z jednotlivých výzkumů na širší oblast jevů či osob. Jednou z cest, jak je možno externí validitu výzkumů výrazně zvýšit, je systematické zavádění srovnávací metody. Rozumíme tím cílevědomý způsob zkoumání, při kterém jednotlivé jevy sledujeme v souvislostech s jinými jevy a poznáváme tak jejich vzájemné vztahy. Srovnávací metodologie může mít v psychologii různé podoby; některé z nich naznačím:

a/ V experimentální situaci je možno sledovat několik relativně nezávislých parametrů /např. latentní dobu reakce, sílu reakce, počet a druh chyb, subjektivní posouzení reakce apod./ a srovnávat zjištěné výsledky; získá se tak mnohem komplexnější pohled na sledovaný jev.

b/ Používat více metod pro měření téhož jevu /např. srovnávat různé dotazníky pro měření stupně neurotičnosti; sledovat úzkost v přirozených podmínkách, v laboratorním pokuse, pomocí dotazníků, využitím posudků od jiných osob atd./; z analýzy výsledků je možno pak získat údaje, společné všem postupům. Získají se tak poznatky, které jsou relativně nezávislé na použitých metodách.

c/ Pracovat s různými skupinami osob /různého věku, profese, vzdělání apod./, s různými experimentátory a případně opakovat stejné výzkumy v různých laboratořích a s různým technickým vybavením. Tímto postupem je možno odhadnout stupeň zobecnitelnosti získaných nálezů.

d/ Sledovat určitý proces v různých situacích /např. v různé denní době, pod vlivem únavy, s různě navozenou motivací, s různou zpětnou informační vazbou, s různou interakcí mezi experimentátorem a pokusnou osobou, atd/. Touto cestou se získají informace o tom, jak dalece je sledovaný proces situačně determinován, či zda je relativně nezávislý na různých situacích.

e/ Srovnávat výsledky šetření s osobnostními charakteristikami jednotlivých pokusných osob a pokusit se tak vysvětlit tu část variance výsledků, která je dána interindividuálními rozdíly mezi lidmi.

f/ Sledovat interakce mezi různými situačními změnami a mezi výkonem různých skupin pokusných osob /různý věk, pohlaví, osobnostní struktura atd./. Tímto postupem je možno odhadnout, do jaké míry je variance výsledků ovlivněna různými situacemi, do jaké míry osobnostními parametry a konečně do jaké míry působí interakce situačních a osobnostních proměnných.

Není možné v této souvislosti pojednávat v plné šíři o problematice srovnávací výzkumné metodologie - zmiňovali jsme se o ní jinde /Břicháček, Brichcín, 1969/. Uvedu zde pouze dvě cesty, které úzce souvisejí s problematikou psychologického škálování.

XV.2. K o n v e r g e n t n í a d i s k r i m i n a č n í v a l i d i t a p s y c h o l o g i c k ý c h z k o u š e k

Na přelomu padesátých a šedesátých let navrhli Campbell a Fiske /1959/ nový postup pro psychometrické experimenty. Řada jejich úvah se dá zobecnit i na další oblasti psychologické metodologie, počítaje v to i psychologické škálování. Oba autoři vyšli z úvah o validitě psychologických testů a stanovili následující pravidla:

a/ Validita musí být konvergentní; při sledování určité vlastnosti musíme získat analogické výsledky, užijeme-li různé a na sobě nezávislé způsoby měření sledované vlastnosti.

b/ Každá zkouška by měla mít diskriminační validitu, tj. její výsledky by měly být nezávislé na výsledcích zkoušek, určených pro měření jiných vlastností či jevů. Pokud by tomu tak nebylo, pak bychom nebyli oprávněni uvažovat o samostatné vlastnosti, kterou v dané zkoušce sledujeme; jednalo by se pak spíše o jakýsi trs vlastností, které spolu těsně souvisejí.

c/ Každá psychologická zkouška je jednotkou, ve které se prolíná jak daná vlastnost, tak použitá metoda, která je většinou nespecifická /dotazník, porovnání apod./. Variabilita získaných výsledků může být způsobena jak obsahovou stránkou sledované vlastnosti, tak reakcí osoby na použitou metodu /a samozřejmě též reliabilitou použité metody/.

d/ Chceme-li zvýšit validitu psychologických nálezů, pak je třeba sledovat a srovnávat výsledky měření několika vlastností pomocí několika metod. Přitom každá ze sledovaných vlastností je měřena všemi užitými metodami.

V základě úvah Campbella a Fiska je myšlenka, že měření jakékoliv jevu je spoludeterminováno metodou, která se používá. Jestliže vlastnosti A a B měříme stejnými metodami /např. pomocí posuzovacích škál/ a zjistíme mezi výsledky vysokou korelaci, pak nelze vyloučit, že toto zjištění je zkreslováno chybami, plynoucími z použité metody /třeba efekt zakotvení apod./. Teprve tehdy, když vlastnost A se měří jednou metodou a vlastnost B jinou, a to zcela nezávislou metodou, můžeme - za předpokladu, že jsme zjistili mezi výsledky vysokou korelaci - tvrdit, že obě vlastnosti jsou spolu velmi těsně vázány.

Představme si nyní data, získaná ve výzkumu, při kterém se postupovalo strategií navrženou Campbellem a Fiskem. Předpokládejme, že měříme tři vlastnosti, značené písmeny A, B, C /třeba introverzi, úzkost a vyrovnanost/ třemi různými metodami: 1, 2, 3 /dotazníkem, laboratorním experimentem a posouzením spolupracovníky/. Získané výsledky jsou shrnuty do korelační matice v tab. XV.1., kterou Campbell a Fiske nazývají maticí mnoha vlastností a mnoha metod /Multitrait - Multimethod Matrix - MTMM/.

Tab. XV.1.

Korelační matice pro sledování konvergentní a diskriminační validity /tři vlastnosti sledovány třemi metodami - umělá data/

Metody		1			2			3		
	Vlast-nosti	A	B	C	A	B	C	A	B	C
1	A	0,96								
	B	0,51	0,89							
	C	0,47	0,43	0,87						
2	A	0,68	0,44	0,43	0,94					
	B	0,44	0,63	0,41	0,61	0,92				
	C	0,43	0,41	0,57	0,42	0,39	0,86			
3	A	0,61	0,46	0,39	0,66	0,32	0,26	0,95		
	B	0,46	0,69	0,36	0,32	0,69	0,20	0,58	0,94	
	C	0,39	0,36	0,61	0,26	0,20	0,58	0,43	0,85	

Campbell a Fiske odvodili, že za ideálních podmínek by pro podobné matice měla platit následující pravidla:

Nejvyšší hodnoty by se měly vyskytnout v hlavní diagonále /jde vlastně o odhad reliability získaných výsledků/.

Nezávislá měření stejné vlastnosti /údaje v diagonálách mezi čárkovaně vymezenými trojúhelníky/ by měla navzájem korelovat výše, než je tomu u měření různých vlastností stejnými metodami. /Tato data jsou vyznačena v tab. XV.1. plnými trojúhelníky / Kdyby tomu tak nebylo, pak by zřejmě výsledky byly v prvé řadě závislé na použité metodě a dostávali bychom jiné výsledky při použití dotazníku, jiné výsledky při šetření laboratorním atd.

Korelace nezávislých měření téže vlastnosti různými metodami by se měly významně lišit od nuly.

Nezávislá měření téže vlastnosti by měla navzájem korelovat výše, než je tomu při měření různých vlastností různými metodami /tato data jsou vyznačena čárkovaným trojúhelníkem/.

Struktura korelací získaných při měření různých vlastností různými metodami /čárkované trojúhelníky/ by měla být zhruba stejná.

Můžeme očekávat, že měření různých vlastností pomocí stejné metody povede k poněkud vyšším korelacím než měření těchto vlastností různými metodami.

Není potřebné v této souvislosti sledovat další úvahy, které jsou zaměřeny na specifickou problematiku psychometrickou. Na okraji pouze poznamenávám, že řada empirických studií, ve kterých byla použita zmíněná metodológie, vedla k podnětným nálezům. Mnohé z nich staví problematiku zobecňování výsledků psychologických výzkumů do nového světla a naznačují oprávněnost metodické skepse.

Zmíním se ještě o další variantě, kterou z myšlenek Campbella a Fiska odovodil Centra /1971/. Pracuje s několika nezávislými skupinami osob /analogie nezávislých metod měření/ a srovnává škálové hodnoty, vypočtené z odpovědí jednotlivých skupin osob v multidimenzionálním posuzovacím prostoru. Jednalo se o posouzení celkového organizačního, myšlenkového i estetického prostředí na různých univerzitách /11 škál/; posouzení prováděli na 22 univerzitách jednak studenti, jednak učitelé a konečně též administrativní pracovníci. Výslednou matici Centra nazývá maticí mnoha skupin a mnoha škál /multigroup - multiscale matrix/. Předpokladem pro tento postup je to, že jednotlivé skupiny jsou s posuzovanými jevy zhruba stejně seznámeny a že jsou dodržena i ostatní pravidla pro aplikaci posuzovacích škál.

Jestliže při podobném postupu zjistíme značnou shodu mezi výsledky jednotlivých skupin, pak můžeme předpokládat, že zvolené škály měří skutečně ty vlastnosti předmětů či jevů, které sledujeme. Pokud by se shoda neprojevila, znamená to, že použité škály nejsou vhodné a nemají obecnější charakter; zachycují spíše subjektivní či skupinově specifický pohled jednotlivých skupin pokusných osob.

Základní srovnávací výzkumný projekt Campbella a Fiska byl modifikován i v dalším směru. Golding a Knudson /1975/ postupují tak, že místo jednotlivých odlišných vlastností sledují konvergenci mezi různými proměnnými, které však jsou věcně i obsahově podobné; při tom se tyto proměnné sledují s použitím několika různých metod. Řadu námětů, jak zpracovávat data, získaná těmito formami experimentálního projektu, uvádí Jackson /1969/.

XV.3. Škálování podobnosti jako základ pro multidimenzionální srovnávací studie

Před dvaceti lety navrhl Ekman /1954/ novou metodu pro rozbor podobnosti podnětů jako základ multidimenzionálního škálování. Ve

XIV. kapitole byl nastíněn myšlenkový postup multidimenzionálního škálování v původní podobě. Zakládal se na zjišťování prostorové vzdálenosti mezi různými podněty. Ekman se neopírá o matematické charakteristiky jednotlivých podnětů ani o jejich umístění v prostoru, ale uvažuje spíše o celkovém subjektivním obsahu podnětů. V přehledové studii /Ekman, Sjöberg, 1965/ pak přímo zavádí pojem "obsahový model multidimenzionálního škálování".

Princip postupu demonstruji na nejstarší Ekmanově studii, která se stala východiskem pro desítky dalších prací, a to jak Ekmanových, tak i jeho žáků. Skupina posuzovatelů /N = 31/ srovnávala párově 14 barevných podnětů o vlnových délkách o rozsahu od 434 mμ do 674 mμ. Posuzovatelé odhadovali stupeň kvalitativní podobnosti mezi podněty na pětibodové škále od 0 /žádná podobnost/ do 4 /shoda/. Výsledky posuzovatelů byly sloučeny a pro každý pár podnětů byla vypočtena průměrná hodnota posouzení podobnosti. Tyto hodnoty pak byly lineárně transformovány na škálu v rozsahu 0,00 až 1,00. Tak vznikla matice podobností, která byla běžným způsobem faktorována. Zjištěné faktory /v uvedené studii dospěl Ekman k pěti faktorům/ byly interpretovány z hlediska teorie barevného vidění.

Teoretici psychologického škálování - jako např. Coombs či Torgerson - však soudili, že Ekmanův postup je příliš komplexní a vede k nadměrnému počtu dimenzí. Větší úpravy navrhli a v mnoha studiích též použili Stone a Coles /1970/, kteří svůj základní výklad publikovali v Bratislavě v časopise Studia psychologica. Vycházejí ze stejného postupu jako Ekman, ale matici podobností považují pouze za průměrná hrubá data, získaná od n posuzovatelů. Vypočítávají mezi nimi korelace a vytvářejí tak matici korelačních podobností /correlational similarity/. Tímto postupem mohou získat jak pozitivní hodnoty, které vyjadřují podobnost , tak negativní hodnoty, které vyjadřují protikladnost či úplnou nepodobnost. Faktorováním takto získané matice mohou získat i bipolární faktory, což u původního postupu Ekmanova nebylo možné. Při novém zpracování dat, získaných Ekmanem pro barevné podněty, stanovili Stone s Colesem /1971/ tři interpretovatelné faktory, které podporovaly poněkud jinou teorii barevného vidění.

Ekehammar /1972/ při srovnávací studii obou postupů dospěl k závěru, že oba navržené modely jsou ve vzájemném vztahu. I když tento vztah není přímo lineární, přece jen vedou oba postupy k analogickým výsledkům.

Jinou úpravu Ekmanova postupu navrhli Lundberg a Devine /1975/ kteří nechávají pokusné osoby odhadovat podobnost mezi podněty na devítibodové škále s rozsahem od +4 do -4. Pozitivní hodnoty zachycují stupeň podobnosti, negativní hodnoty vyjadřují stupeň protikladnosti /"negativní podobnost"/ a 0 znamená, že mezi danými podněty není žádný vztah. Ukázalo se, že dvojice podnětů, které byly posuzovány jako protikladné, by byly v původním postupu řazeny do kategorie 0. /žádná podobnost/. Navržená úprava vede tedy k větší citlivosti škály; ukázalo se, že konečné výsledky se od sebe sice příliš neliší,,ale přece jen dochází k určitému zjemnění, které je psychologicky dobře interpretovatelné.

Ekmanem navržená technika je velmi slibná a byla využita v řadě studií. Hlavní přínos vidím v tom, že umožňuje srovnávat a vyhledávat společné prvky v řadě podnětů. Byly např. srovnávány miniaturní životní situace, do kterých se dostávají studenti při studiu, situace vyvolávající strach, pracovní podmínky, různé známé osobnosti, učitelé a žáci ve škole, tabule Rorschachova testu, škály MMPI, patické formy chování atd. Z tohoto namátkového výčtu řešené problematiky vyplývá, že Ekmanovu techniku je možno použít pro nejrůznější oblasti psychologického výzkumu i praxe. Po stránce numerické je postup dobře prověřen a je dnes standardně užíván.

Nebezpečí této metody vidím v celkovém stupni spolehlivosti výchozích dat. Pětibodová škála podobnosti se mi zdá příliš úzká a při velkém počtu srovnávaných podnětů /v jedné Stoneově studii pokusné osoby srovnávaly dokonce více než 230 párů podnětů/ může snadno dojít k určité stereotypnosti a hodnota výchozích dat se stává spornou. Rozšíření škály pro posouzení podobnosti do rozsahu od 0,00 do 1,00 má opět opačnou nevýhodu. Škála je příliš široká a většina odhadů se grupuje do několika hodnot; i zde vzniká při větším počtu srovnávaných podnětů nebezpečí stereotypnosti /i dalších chyb, o kterých jsem se zmiňoval v kapitole o subjektivních posuzovacích škálách/. Jedna z možných cest pro další úpravu postupu by mohla být dána využitím techniky Q-třídění.

Na okraji ještě poznamenávám, že problematice škálování podobnosti mezi různými jevy se v poslední době věnuje značná pozornost. Přehled stavu bádání je možno nalézt v monografii Grecsonově /1975/.

XV.4. S h r n u t í

Požadavek sledovat a srovnávat výsledky několika zkoušek s použitím více metod se ukazuje jako velmi potřebný a při současném stavu psychologického bádání jako velmi plodný postup. Řada empirických prací, které se v posledních letech v literatuře objevují stále častěji, toto tvrzení podporuje.

V problematice psychologického škálování není zatím srovnávací výzkum jednotlivých škálovacích postupů systematicky prováděn. Existuje sice několik studií ze starší doby, ve kterých se srovnávaly výsledky, získané při použití metody párového srovnávání podnětů s výsledky získanými metodou pořadových stupnic; srovnávaly se též posuzovací škály s pořadovými škálami atd. Řada studií tohoto typu je uvedena v monografii Guilfordově /1954/. Jejich nálezy jsou vcelku uspokojivé; zjišťuje se, že výsledky získané při užití různých škálovacích technik se od sebe příliš neliší. Většinou se však jedná o srovnávání jediného podnětového kontinua, které samo o sobě může být velmi vyhraněné, a proto také snáze škálovatelné /např. chuťové preference, smysl pro humor apod./.

Jako příklad novější srovnávací studie můžeme uvést výzkum Fischera a kol. /1968/, kteří srovnávali techniku párového srovnávání podnětů s technikou sumovaných odhadů; přitom pracovali záměrně se dvěma odlišnými skupinami osob /studenti a dospělí zaměstnanci/. Ukázalo se, že metoda párového srovnávání vedla k vytvoření škály s lepšími psychometrickými vlastnostmi, ale postup je mnohem zdlouhavější a některým osobám nevyhovoval. Rozdíly mezi oběma technikami se projevily výrazněji u skupiny starších zaměstnanců než u studentů. Korelace vytvořených škál /20 bodů/ byla +0,62, což není příliš povzbudivé. Nápadné byly též velké interindividuální rozdíly; u některých osob byla korelace prakticky nulová, u jiných osob vyšší než 0,96.

Při srovnávání méně vyhraněných podnětů či jevů dochází často k výsledkům rozpornějším /viz např. Stoneovu reinterpretaci Ekmanových výsledků či jejich přezkoušení Coombsovými technikami/. Chybí též srovnání některých novějších postupů /Guttman, Coombs aj./ s postupy tradičními. Slibný je naopak návrh Stewartův /1974/ užívat současně faktorovou analýzu a multidimenzionální škálování a srovnáním výsledků dospět k větší obecnosti závěrů.

I když můžeme být v hrubých obrysech vcelku spokojeni s externí validitou nálezů, získaných jednotlivými škálovacími postupy, přece jen je na místě jednak jistá zdrženlivost, jednak doporučení zaměřit výzkum mnohem více na systematické srovnávání škálovacích postupů.

Kapitola XVI.

ZÁVĚR

V předcházejících kapitolách jsem se pokusil vybrat a vyložit základní škálovací metody, které se dnes v psychologii stále častěji používají. Jejich počet je ve skutečnosti mnohem větší, jak ví každý, kdo sleduje odpovídající časopiseckou literaturu. V úvodu jsem naznačil, podle jakých kriterií jsem postupoval při výběru metod, které jsem se pokusil vyložit. V závěru zbývá jen několik obecnějších poznámek.

Proti systematickému používání škálovacích metod v psychologii se objevují skeptické názory. Opírají se o dojem, že jejich pomocí nelze zachytit složité procesy, kterými se psychologie zabývá. Tato námitka má své racionální jádro. Jistě nebudeme složitě škálovat jevy, které můžeme objektivně zachytit a přesně měřit /třeba rychlost pohybu, přesnost pracovního úkonu, počet chyb, kterých se dopustí žáci v určité zkoušce atd./. Naproti tomu však využití škálovacích metod umožňuje mnohem přesněji a dokonaleji vyjádřit jevy, u kterých bychom se jinak museli spokojit s velmi nepřesnými a subjektivně formulovanými názory, odhady či posudky. Škálováním empirických údajů se značně zvýší množství získaných informací, a tím se nesporně zdokonaluje jak teoretické poznámí, tak i praktická činnost.

Tak jako žádná vědecká metoda nemůže být absolutní, ani psychologické škálování nemůže být metodologickou záchranou psychologického výzkumu. Po stránce věcné je přiměřené jen k určitým typům dat a k určitým oblastem zkoumání. Po stránce vývojové není ještě ani zdaleka ukončen vývoj škálovacích metod a můžeme očekávat, že dnešní poznatky i užívané postupy budou dříve i později upraveny, zdokonaleny či dokonce překonány. A konečně po stránce vlastní výzkumné techniky je třeba mít na zřeteli že se při psychologickém škálování může vyskytnout celá řada experimentálních chyb, které vznikají z různých zdrojů. Mohou to být chyby plynoucí a/ z výběru osob, z jejich celkového zaměření, ochoty spolupracovat atd.; b/ z konkréních situačních podmínek, ve kterých se výzkum provádí; c/ z charakteru a vlastností podnětů, které jsou škálovány; d/ z interakce těchto faktorů; e/ z řady náhodných a těžko analyzovatelných vlivů. Pokud si psycholog není vědom těchto úskalí, která se někdy až nebezpečně blíží Scylle a Charybdě, může se snadno dopustit výrazných gnoseologických chyb. I zde - tak jako v celé psychologické metodologii - platí, že racionální a všestranný rozbor celé situace musí předcházet před vlastním výzkumem - aneb je třeba nejdříve užívat našich hlav a teprve pak našich vzorců.

Kdykoliv se v psychologii pracuje s kvantitativními metodami - ať je to při statistickém zpracování získaných dat, ať je to při psychometrickém ověřování různých zkoušek, ať je to v různých oblastech matematické psychologie - hrozí nebezpečí formalismu. Toto nebezpečí platí i při psychologickém škálování - snad o to více, že se jedná o metody, které ještě mnohdy nejsou zcela ustáleny a dostatečně prověřeny pro různé typy výzkumů či pro různé druhy dat. Toto nebezpečí však neplyne z charakteru kvantitativních metod v psychologii, ale z jejich gnosoeologicky chybného využití.

I přes mnohá úskalí, která jsou spojena s využitím škálovacích metod v psychologii, jsem přesvědčen, že se jedná o metody velmi plodné a pro práci psychologa užitečné. Umožňují zachytit, analyzovat a srovnávat subjektivní údaje, které často byly polem, na které se uchylovaly idealisticky laděné ideologie a spekulace. Při psaní této knihy se mi často vynořila jedna z charakteristik Karla Marxe, kterou zachytil P. Lafargue: "... podle jeho názoru byla určitá věda teprve tehdy rozvinuta, když dospěla k tomu, že mohla využívat matematiku" /1961, str. 293/.

Příloha A: PROCENTO PŘÍPADU, KTERÉ LEŽÍ MEZI PRUMĚREM A URČITOU HODNOTOU z /NORMALIZOVANÉ ROZLOŽENÍ/

z	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,0	0,00	0,40	0,80	1,20	1,60	1,99	2,39	2,79	3,19	3,59
0,1	3,98	4,38	4,78	5,17	5,57	5,96	6,36	6,75	7,14	7,53
0,2	7,93	8,32	8,71	9,10	9,48	9,87	10,26	10,64	11,03	11,41
0,3	11,79	12,17	12,55	12,93	13,31	13,68	14,06	14,43	14,80	15,17
0,4	15,54	15,91	16,28	16,64	17,00	17,36	17,72	18,08	18,44	18,79
0,5	19,15	19,50	19,85	20,19	20,54	20,88	21,23	21,57	21,90	22,24
0,6	22,57	22,91	23,24	23,57	23,89	24,22	24,54	24,86	25,17	25,49
0,7	25,80	26,11	26,42	26,73	27,04	27,34	27,64	27,94	28,23	28,52
0,8	28,81	29,10	29,39	29,67	29,95	30,23	30,51	30,78	31,06	31,33
0,9	31,59	31,86	32,12	32,38	32,64	32,90	33,15	33,40	33,65	33,89
1,0	34,13	34,38	34,61	34,85	35,08	35,31	35,54	35,77	35,99	36,21
1,1	36,43	36,65	36,86	37,08	37,29	37,49	37,70	37,90	38,10	38,30
1,2	38,49	38,69	38,88	39,07	39,25	39,44	39,62	39,80	39,97	40,15
1,3	40,32	40,49	40,66	40,82	40,99	41,15	41,31	41,47	41,62	41,77
1,4	41,92	42,07	42,22	42,36	42,51	42,65	42,79	42,92	43,06	43,19
1,5	43,32	43,45	43,57	43,70	43,83	43,94	44,06	44,18	44,29	44,41
1,6	44,52	44,63	44,74	44,84	44,95	45,05	45,15	45,25	45,35	45,45
1,7	45,54	45,64	45,73	45,82	45,91	45,99	46,08	46,16	46,25	46,33
1,8	46,41	46,49	46,56	46,64	46,71	46,78	46,86	46,93	46,99	47,06
1,9	47,13	47,19	47,26	47,32	47,38	47,44	47,50	47,56	47,61	47,67
2,0	47,72	47,78	47,83	47,88	47,93	47,98	48,03	48,08	48,12	48,17
2,1	48,21	48,26	48,30	48,34	48,38	48,42	48,46	48,50	48,54	48,57
2,2	48,61	48,64	48,68	48,71	48,75	48,78	48,81	48,84	48,87	48,90
2,3	48,93	48,96	48,98	49,01	49,04	49,06	49,09	49,11	49,13	49,16
2,4	49,18	49,20	49,22	49,25	49,27	49,29	49,31	49,32	49,34	49,36
2,5	49,38	49,40	49,41	49,43	49,45	49,46	49,48	49,49	49,51	49,52
2,6	49,53	49,55	49,56	49,57	49,59	49,60	49,61	49,62	49,63	49,64
2,7	49,65	49,66	49,67	49,68	49,69	49,70	49,71	49,72	49,73	49,74
2,8	49,74	49,75	49,76	49,77	49,77	49,78	49,79	49,79	49,80	49,81
2,9	49,81	49,82	49,82	49,83	49,84	49,84	49,85	49,85	49,86	49,86
3,0	49,865									
3,1	49,903									
3,2	49,931									
3,3	49,952									
3,4	49,966									
4,5	49,977									
3,6	49,984									
3,7	49,989									
3,8	49,993									
3,9	49,995									
4,0	49,997									
5,0	49,99997									

Příloha B: TABULKA KRITICKÝCH HODNOT χ^2

Stupně volnosti	Hladina významnosti α			
	0,10	0,05	0,02	0,01
1	2,71	3,84	5,41	6,64
2	4,61	5,99	7,82	9,21
3	6,25	7,82	9,84	11,34
4	7,78	9,49	11,67	13,28
5	9,24	11,07	13,39	15,09
6	10,65	12,59	15,03	16,81
7	12,02	14,07	16,62	18,48
8	13,36	15,51	18,17	20,09
9	14,68	16,92	19,68	21,67
10	15,99	18,31	21,16	23,21
11	17,28	19,68	22,62	24,73
12	18,55	21,03	24,05	26,22
13	19,81	22,36	25,47	27,69
14	21,06	23,69	26,87	29,14
15	22,31	25,00	28,26	30,58
16	23,54	26,30	29,64	32,00
17	24,77	27,59	31,00	33,41
18	25,99	28,87	32,35	34,81
19	27,20	30,14	33,69	36,19
20	28,41	31,41	35,02	37,57
21	29,62	32,67	36,34	38,93
22	30,81	33,92	37,66	40,29
23	32,01	35,17	38,97	41,64
24	33,20	36,42	40,27	42,98
25	34,38	37,65	41,57	44,31
26	35,56	38,89	42,86	45,64
27	36,74	40,11	44,14	46,96
28	37,92	41,34	45,42	48,28
29	39,09	42,56	46,69	49,59
30	40,26	43,77	47,96	50,89

Pozn.: Při větším počtu stupňů volnosti se počítá hodnota
$t = \sqrt{2\chi^2} - \sqrt{2n-1}$, kde n je počet stupňů volnosti.

Příloha C: POŘADOVÉ HODNOTY TRANSFORMOVANÉ DO C-ŠKÁLY PŘI RŮZNÉM POČTU ŘAZENÝCH PODNĚTŮ

Počet řazených podnětů	Hodnoty C-škály								
	1	2	3	4	5	6	7	8	9
10	..	10	9	7-8	5-6	3-4	2	1	..
11	..	11	9-10	8	5-7	4	2-3	1	..
12	..	12	10-11	8-9	6-7	4-5	2-3	1	..
13	13	..	11-12	9-10	6-8	4-5	2-3	..	1
14	14	..	12-13	9-11	7-8	4-6	3-3	..	1
15	15	14	13	10-12	7-9	4-6	3	2	1
16	16	15	13-14	11-12	7-10	5-6	3-4	2	1
17	17	16	14-15	11-13	8-10	5-7	3-4	2	1
18	18	17	15-16	12-14	8-11	5-7	3-4	2	1
19	19	18	16-17	12-15	9-11	5-8	3-4	2	1
20	20	19	16-18	13-15	9-12	6-8	3-5	2	1
21	21	20	17-19	14-16	9-13	6-8	3-5	2	1
22	22	21	18-20	14-17	10-13	6-9	3-5	2	1
23	23	22	19-21	15-18	10-14	6-9	3-5	2	1
24	24	22-23	20-21	15-19	11-14	6-10	4-5	2-3	1
25	25	23-24	20-22	16-19	11-15	7-10	4-6	2-3	1
26	26	24-25	21-23	17-20	11-16	7-10	4-6	2-3	1
27	27	25-26	22-24	17-21	12-16	7-11	4-6	2-3	1
28	28	26-27	23-25	18-22	12-17	7-11	4-6	2-3	1
29	29	27-28	23-26	18-22	13-17	8-12	4-7	2-3	1
30	30	28-29	24-27	19-23	13-18	8-12	4-7	2-3	1
31	31	29-30	25-28	20-24	13-19	8-12	4-7	2-3	1
32	32	30-31	26-29	20-25	14-19	8-13	4-7	2-3	1
33	33	30-31	27-30	21-26	14-20	8-13	4-7	2-3	1
34	34	31-33	27-30	21-26	15-20	9-14	5-8	2-4	1
35	35	32-34	28-31	22-27	15-21	9-14	5-8	2-4	1
36	36	33-35	29-32	23-28	15-22	9-14	5-8	2-4	1
37	37	34-36	30-33	23-29	16-22	9-15	5-8	2-4	1
38	37-38	35-36	30-34	24-29	16-23	10-15	5-9	3-4	1-2

Příloha C: p o k r a č o v á n í

Počet řazených podnětů	Hodnoty C-škály								
	1	2	3	4	5	6	7	8	9
39	38-39	36-37	31-35	24-30	17-23	10-16	5-9	3-4	1-2
40	39-40	37-38	32-36	25-31	17-24	10-16	5-9	3-4	1-2
41	40-41	38-39	33-37	26-32	17-25	10-16	5-9	3-4	1-2
42	41-42	39-40	33-38	26-32	18-25	11-17	5-10	3-4	1-2
43	42-43	39-41	34-38	27-33	18-26	11-17	6-10	3-5	1-2
44	43-44	40-42	35-39	27-34	19-26	11-18	6-10	3-5	1-2
45	44-45	41-43	36-40	28-35	19-27	11-18	6-10	3-5	1-2
46	45-46	42-44	37-41	29-36	19-28	11-18	6-10	3-5	1-2
47	46-47	43-45	37-42	29-36	20-28	12-19	6-11	3-5	1-2
48	47-48	44-46	38-43	30-37	20-29	12-19	6-11	3-5	1-2
49	48-49	45-47	39-44	30-38	21-29	12-20	6-11	3-5	1-2
50	49-50	46-48	40-45	31-39	21-30	12-20	6-11	3-5	1-2
51	50-51	47-49	40-46	32-39	21-31	13-20	6-12	3-5	1-2
52	51-52	47-50	41-46	32-40	22-31	13-21	7-12	3-6	1-2
53	52-53	48-51	42-47	33-41	22-32	13-21	7-12	3-6	1-2
54	53-54	49-52	43-48	33-42	23-32	13-22	7-12	3-6	1-2
55	54-55	50-53	44-49	34-43	23-33	13-22	7-12	3-6	1-2
56	55-56	51-54	44-50	35-43	23-34	14-22	7-13	3-6	1-2
57	56-57	52-55	45-51	35-44	24-34	14-23	7-13	3-6	1-2
58	57-58	53-56	46-52	36-45	24-35	14-23	7-13	3-6	1-2
59	58-59	54-57	47-53	36-46	25-35	14-24	7-13	3-6	1-2
60	59-60	55-58	47-54	37-46	25-36	15-24	7-14	3-6	1-2
61	60-61	56-59	48-55	38-47	25-37	15-24	7-14	3-6	1-2
62	61-62	56-60	49-55	38-48	26-37	15-25	8-14	3-7	1-2
63	61-63	57-60	50-56	39-49	26-38	15-25	8-14	4-7	1-3
64	62-64	58-61	50-57	39-49	27-38	16-26	8-15	4-7	1-3
65	63-65	59-62	51-58	40-50	27-39	16-26	8-15	4-7	1-3
66	64-66	60-63	52-59	41-51	27-40	16-26	8-15	4-7	1-3
67	65-67	61-64	53-60	41-52	28-40	16-27	8-15	4-7	1-3
68	66-68	62-65	54-61	42-53	28-41	16-27	8-15	4-7	1-3

Příloha C: p o k r a č o v á n í

Počet řazených podnětů	Hodnoty C-škály								
	1	2	3	4	5	6	7	8	9
69	67-69	63-66	54-62	42-53	29-41	17-28	8-16	4-7	1-3
70	68-70	64-67	55-63	43-54	29-42	17-28	8-16	4-7	1-3
71	69-71	64-68	56-63	44-55	29-43	17-28	9-16	4-8	1-3
72	70-72	65-69	57-64	44-56	30-43	17-29	9-16	4-8	1-3
73	71-73	66-70	57-65	45-56	30-44	18-29	9-17	4-8	1-3
74	72-74	67-71	58-66	45-57	31-44	18-30	9-17	4-8	1-3
75	73-75	68-72	59-67	46-58	31-45	18-30	9-17	4-8	1-3
76	74-76	69-73	60-68	47-59	31-46	18-30	9-17	4-8	1-3
77	75-77	70-74	61-69	47-60	32-46	18-31	9-17	4-8	1-3
78	76-78	71-75	61-70	48-60	32-47	19-31	9-18	4-8	1-3
79	77-79	72-76	62-71	48-61	33-47	19-32	9-18	4-8	1-3
80	78-80	73-77	63-72	49-62	33-48	19-32	9-18	4-8	1-3
81	79-81	73-78	64-72	50-63	33-49	19-32	10-18	4-9	1-3
82	80-82	74-79	64-73	50-63	34-49	20-33	10-19	4-9	1-3
83	81-83	75-80	65-74	51-64	34-50	20-33	10-19	4-9	1-3
84	82-84	76-81	66-75	51-65	35-50	20-34	10-19	4-9	1-3
85	83-85	77-82	67-76	52-66	35-51	20-34	10-19	4-9	1-3
86	84-86	78-83	67-77	53-66	35-52	21-34	10-20	4-9	1-3
87	85-87	79-84	68-78	53-67	36-52	21-35	10-20	4-9	1-3
88	86-88	80-85	69-79	54-68	36-53	21-35	10-20	4-9	1-3
89	86-89	81-85	70-80	54-69	37-53	21-36	10-20	5-9	1-4
90	87-90	81-86	71-80	55-70	37-54	21-36	11-20	5-10	1-4
91	88-91	82-87	71-81	56-70	37-55	22-36	11-21	5-10	1-4
92	89-92	83-88	72-82	56-71	38-55	22-37	11-21	5-10	1-4
93	90-93	84-89	73-83	57-72	38-56	22-37	11-21	5-10	1-4
94	91-94	85-90	74-84	57-73	39-56	22-38	11-21	6-10	1-4
95	92-95	86-91	74-85	58-73	39-57	23-38	11-22	5-10	1-4
96	93-96	87-92	75-86	58-74	40-57	23-39	11-22	5-10	1-4
97	94-97	88-93	76-87	59-75	40-58	23-39	11-22	5-10	1-4

Příloha C : p o k r a č o v á n í

Počet řazených podnětů	Hodnoty C-škály								
	1	2	3	4	5	6	7	8	9
98	95-98	89-94	77-88	60-76	40-59	23-39	11-22	5-10	1-4
99	96-99	90-95	78-89	60-77	41-59	23-40	11-22	5-10	1-4
100	97-100	90-96	78-89	61-77	41-60	24-40	12-23	5-11	1-4

Příklad užití tabulky C

Chceme určit transformovanou pořadovou hodnotu osoby, která je umístěna na 15. místo z 52 řazených osob. Ve sloupci "počet řazených podnětů" vyhledáme hodnotu 52. V odpovídající řádce hledáme pořadovou hodnotu 15 a zjišťujeme, že pořadové hodnoty v rozmezí 13-31 odpovídají na C-škále hodnotě 6.

Kap. XVII. D o d a t e k

Od odevzdání rukopisu uplynula určitá doba, ve které se objevily nové náměty pro celou oblast psychologického škálování. /viz např. Jarminckij, 1977/. Využívám proto příležitosti, kterou mi ochotně poskytl vydavatel, k tomu, abych v dodatkové kapitole se zmínil o některých novinkách, které považuji za důležité. Připojená kapitola má dvě části; v prvé upozorním na problematiku škálování a měření z hlediska teorie neostrých množin; druhá obsahuje doplňky k některým škálovacím metodám, popsaným v předešlých kapitolách.

XVII.1. N e o s t r é m n o ž i n y a p s y c h o l o g i c k é š k á l o v á n í

V psychologickém škálování se velmi často objevují výroky, které jsou obsahově neurčité a nejsou ostře ohraničeny. Setkáváme se třeba s tvrzením: inteligence osoby X je značně nadprůměrná, rodinné poměry jsou dosti neurovnané, žák není příliš pilný apod. V posledních letech se objevil pokus vytvořit formální systém, který by umožňoval zpřesnit práci s neurčitými pojmy a vytvořit kvantitativní teorii pro jejich analýzu. Postup se opírá o pojem rozmazaných či neostrých množin /fuzzy sets, nečotkyje, razmytyje množestva/, který vytvořil v polovině šedesátých let L.A.Zadeh /1965/; v poslední době se začíná používat i ve společenských vědách /viz Nowakowska, 1976, Břicháček, 1978/.

V klasické teorii se množinou rozumí soubor předmětů. Je zároveň dáno pravidlo, podle kterého můžeme rozhodnout, které předměty do dané množiny patří a které nikoliv. Rozhodovací pravidlo je jednoznačné a proto také hranice množiny jsou ostré a nesporné. Můžeme např. jednotlivé školní děti rozdělit na žáky s vyznamenáním a žáky s horším prospěchem, dospělé osoby můžeme dělit na muže a ženy atd. V praxi se však setkáváme často s takovými množinami, jejichž hranice nejsou zcela ostré /jsou "rozmazané"/ a u mnoha prvků nemůžeme jednoznačně rozhodnout, zda do určité množiny patří či nikoliv. Právě touto problematikou se zabývá teorie neostrých množin.

Neostré množiny vzniknou třeba tříděním osob podle vlastností, ve kterých jsou rozdíly mezi lidmi kontinuitní ať již v jedné či několika dimenzích /např. velcí - malí, nadaní - nenadaní apod./. O osobě, která měří 190 cm, můžeme říci, že je velká; o osobě vysoké 160 cm řekneme, že je malá, ale osobu vysokou 175 cm budeme zařazovat obtížně. Řešení je buď stanovit přesné dělítko /třeba všecky osoby větší než 172 cm zařadíme mezi velké/ a vytvořit tak ostře definovanou množinu anebo se spokojit s neostrým ohraničením obou množin a zjišťovat spíše míru "příslušnosti" určitého prvku do nich. Přitom si vypomáháme často semanticky nepřesně definovanými výrazy jako třeba "je poměrně velký", "není příliš velký", "je spíše velký než malý" apod.

Klasická teorie množin předpokládá, že libovolný předmět x buď patří do množiny M /obvykle zapisujeme $x \in M$/ anebo do ní nepatří /$x \notin M$; x není prvek množiny M/. Teorie neostrých množin se zabývá takovými jevy, u nichž míra příslušnosti prvku x do množiny M může nabývat jakékoliv hodnoty v rozmezí intervalu $\langle 0,1 \rangle$. Jde přitom o vyjádření míry pravdivosti určitého tvrzení jako např., že žák x je či není nadaný.

Formální výklad základních operací s neostrými množinami zde není možno uvádět. Je popsán v řadě studií Zadehových /viz bibliografie sestavená Gainesem, 1977/ a ve zjednodušené podobě pro potřeby psychologů Herschem a Garamazzem /1976/. V této souvislosti půjde pouze o naznačení některých důsledků teorie neostrých množin pro problematiku psychologického škálování. V literatuře zatím tyto otázky nebyly řešeny a proto se jedná o předběžné úvahy, které si nečiní nárok ani na úplnost ani na nespornost.

Ve škálovací praxi jde často o to najít vhodnou míru přesnosti, s jakou budeme v dané situaci pracovat. Asi by nebylo přiměřené klást otázku, zda ve třetí třídě pražských ZDŠ je počet dětí matematicky nadaných sudý či lichý. Jestliže nemáme k dispozici přesný nástroj, který by jednoznačně a reliabilně detekoval matematicky nadané děti, pak budeme jejich počet pouze hrubě odhadovat a postavená otázka není přiměřená. Zjišťovaná data jsou alespoň o řád méně přesná než byla formulace otázky. Zadeh /1976/ dokonce soudí, že přesnost vyjadřování a komplexnost problematiky jsou často v inversním vztahu. Jak roste komplexnost problému, tak klesá možnost jeho analýzy ve zcela přesných termínech. Soudím však, že právě proto je třeba hledat prostředky, které by umožňovaly vhodně řešit i ty problémy, které jsou pro precizní kvantitativní analýzu prakticky nepřístupné. V této souvislosti se obvykle hovoří o aproximativním myšlení; Brušlinský /1976/ považuje přechod k tomuto uvažování za základní předpoklad pro další pokrok o teorii lidského myšlení.

Při psychologickém škálování se často vytvářejí kategorie buď předem /např. při posuzovacích škálách/ anebo jako výsledek vlastní škálovací procedury /technika následných intervalů, sumovaných odhadů aj./ Přitom jejich hranice většinou nejsou a ani nemohou být ostré. Proto je třeba v těch případech, kdy určitý podnět je na škále umístěn těsně u dělícího bodu, uvažovat o jeho přináležitosti do určité kategorie s větší opatrností než u těch podnětů, které leží na škále v blízkosti středu určitého intervalu. Pojetí neostrých škálových intervalů se do jisté míry blíží úvahám o intenzitní složce Guttmanova škálogramu /viz kap. VIII.3./

Teorie neostrých množin upozornila též na problematičnost a informační vágnost takových výrazů jako "příliš", "poněkud", "značně" atd., které se při popisu různých bodů na škále často vyskytují /hlavně u posuzovacích škál/. Zadeh odvodil z teoretických úvah o neostrých množinách kvantifikátory pro řadu podobných výrazů /je velmi velký, není velmi malý apod./. Jeho koncepty vcelku potvrdil i rozsáhlý empirický výzkum Hersche a Caramazza /1976/, kteří uváděli do vztahu vágní výrazy vyjadřující velikost s různě velkými geometrickými obrazci.

Při škálování vlastností osobnosti je však problematika složitější. V orientačním výzkumu 73 odborných psychologů odhadovalo kolik bodů na škále introverze a extroverze by musela dosáhnout osoba, kterou by bylo možno označit za velmi extrovertovanou či mírně introvertovanou, nebo ne příliš extrovertovanou apod. /celkem 16 různých formulací - podrobnosti viz Břicháček 1978/. Ukázalo se, že pozitivní formulace jednotlivých výroků jsou posuzovány celkem homogenně; na jejich základě může vzniknout škála, která se blíží intervalové stupnici. Výsledky jsou shrnuty v Tab. XVII.1.

Tab.č.XVII.1. Posouzení semanticky neurčitých pojmů v dimenzi introverze-extroverze /škála v rozsahu 0-24 bodů; 73 posuzovatelů/

	Velmi značně introvertovaný	Velmi introvertovaný	Introvertovaný	Mírně introvertovaný	Mírně extrovertovaný	Extrovertovaný	Velmi extrovertovaný	Velmi značně extrovertovaný
Průměr	2,32	3,55	6,32	8,70	14,33	16,56	19,89	22,22
Směrodatná odchylka	1,77	2,47	3,50	2,16	1,83	2,97	2,87	1,79

Jestliže však sledujeme negativní formulace stejných výroků, získáme výsledky zcela odlišné. Posouzení jednotlivých výroků je mnohem variabilnější a víceznačné. Nevytvoří se žádná použitelná škála. Výsledky jsou shrnuty v Tab. XVII.2.

Tab.č.XVII.2. Posouzení negativních semanticky neurčitých pojmů v dimenzi introverze - extroverze /škála v rozsahu 0-24 bodů; 73 posuzovatelů/

	Není velmi značně introvertovaný	Není velmi introvertovaný	Není introvertovaný	Není mírně introvertovaný	Není mírně extrovertovaný	Není extrovertovaný	Není velmi extrovertovaný	Není velmi značně extrovertovaný
Průměr	10,97	9,54	12,66	8,36	13,72	9,63	13,93	12,67
Směrodatná odchylka	4,70	2,49	2,91	4,11	5,01	3,50	2,68	5,10

I bez bližší analýzy, která je provedena jinde /Břicháček, 1978/, je zřejmé, že negativní formulace výrazně "rozostřují" hranice jednotlivých kategorií a prakticky znemožňují rozumnou komunikaci mezi lidmi. Na nebezpečí, plynoucí z užívání negativních informací, bylo poukazováno v psychologii již mnohokrát. Je zřejmé, že i v psychologickém škálování je třeba se jim vyhýbat a jednotlivé posuzovací kategorie popisovat pozitivně.

Celkově přináší úvahy o neostrých množinách do problematiky škálování v psychologii pozměněný pohled na charakter výsledných škál. V mnoha případech bude ohraničení bodů na škále neostré a z tohoto faktu mohou vyplynout některé chyby v uvažování a v interpretaci výsledků. Bude jistě na místě sledovat vývoj aplikace neostrých množin /a vůbec celé fuzzy-logiky/ na metodologickou problematiku měření a odvodit z toho poučení pro potřebu psychologického výzkumu i praxe.

XVII.2. D o d a t k y k n ě k t e r ý m š k á l o v a c í m m e t o d á m

Při průběžném sledování literatury se objevuje neustále řada nových podnětů. Poměrně nejrychlejší vývoj je možno sledovat ve dvou oblastech: jednak stále přibývající počítačové programy pro nejrůznější škálovací techniky i pro jejich různé aspekty, jednak roste počet prací, zabývajících se multidimensionálním škálováním. První oblast je příliš technická a vázaná na určitý typ počítače, který je k dispozici; druhá oblast přesahuje rámec orientační informace, obsažené ve XIV. kapitole. Soustředím se jen na ty podněty, které přinášejí metodické zdokonalení pro běžné škálovací metody.

XVII.2.a. Nové poznatky o pořadových stupnicích

V poslední době se výrazněji objevil námět na vytvoření jakéhosi kompromisu mezi metodou párového srovnávání podnětů a metodou pořadových stupnic. Hovoří se pak o metodě mnohonásobných pořadových stupnic /Straton, 1975, Tziner a Zahal, 1977/. Víme již, že při větším počtu podnětů je jejich párové srovnávání prakticky neproveditelné anebo je spojené s mnoha metodickými úskalími. Řazení podnětů do pořadí je sice poměrně rychlé, ale s rostoucím počtem podnětů se snižuje reliabilita získaných dat a to hlavně tehdy, když srovnávané podněty jsou si velmi podobné. Toto nebezpečí je možno snížit tím, že systematicky v několika blocích srovnáváme různě vytvořené podmnožiny podnětů.

Postupuje se tak, že n-podnětů, z kterých chceme vytvořit škálu, se rozdělí do podmnožin o k-prvcích. Počet prvků může být určen buď důvody věcnými anebo vyplývá automaticky ze zvoleného projektu. Obvykle se uvádí, že nejvýhodnější situace nastává, když se pracuje s řazením 3 - 5 prvků. Při stanovení výzkumného postupu se obvykle užívá projekt s vyrovnanými neúplnými bloky. Platí o nich tyto požadavky: a/ každý podnět se objeví v celém experimentálním projektu stejně často, b/ každá dvojice podnětů se objeví společně v rámci bloků stejně často, c/ pořadí podnětů v rámci jednotlivých bloků se musí systematicky měnit. Uveďme jeden z možných příkladů uspořádání 7 podnětů do podmnožin o 3 prvcích; schema je v Tab.č.XVII.3.

Znamená to, že posuzovatel nejprve řadí podle své preference podněty A, B, C, pak uvažuje o trojici D, A, E, následuje F,

Tab.XVII.č.3. Projekt vyrovnaných neúplných bloků pro srovnání sedmi podnětů.

Blok	Podněty		
1	A	B	C
2	D	A	E
3	F	G	A
4	B	F	D
5	G	E	B
6	C	D	G
7	E	C	F

G, A atd. Snadno je možno ověřit, že každý podnět se objeví třikrát a to pokaždé na jiném místě v pořadí prezentace a konečně každý pár podnětů se vyskytuje společně jedenkrát. Můžeme proto pro každou osobu sestavit komplexní matici párových preferencí a případně ověřovat, zda se objevily kruhové triády /i když u trojice, které se spolu vyskytují v rámci jednotlivých bloků nemohou vzniknout./

Je možno volit i některé jiné postupy pro sestavení mnohonásobných pořadových stupnic. Může třeba nastat situace, kdy záměrně chceme, aby některé dvojice podnětů se vyskytly spolu v rámci různých bloků častěji než jiné dvojice. Nebo se spokojíme s tím, že některé páry podnětů se spolu nevyskytnou vůbec a matice dat nebude kompletní. V těchto případech můžeme užít některou z forem projektů částečně vyrovnaných neúplných bloků. Podrobnosti o jejich sestavení a vyhodnocení získaných dat je možno najít ve studii Stratonově /1975/, ve které jsou uvedeny i návody pro výpočet vnitřní konzistence získaných dat /analogie Kenddallova koeficientu zeta/. Uvedený autor též srovnával mnohonásobné pořadové stupnice s jednoduchými pořadovými stupnicemi a zjistil, že je mezi výsledky značná shoda; u mnohonásobných stupnic se však prokazuje větší reliabilita při opakovaném šetření i větší vnitřní konzistence dat. Tento nález se potvrdil u tří velmi odlišných skupin osob i při dvou různých typech podnětů.

Je proto možno uvedenou metodu škálování doporučit k běžnému užívání. Pokud můžeme užít některou variantu projektu vyrovnaných neúplných bloků, měli bychom jí asi dávat přednost jak před párovým srovnáváním podnětů, tak před jednoduchými pořadovými stupnicemi.

V posledních létech se též objevily nové úvahy o aplikaci faktorové analýzy na pořadové škály. Studie Kruskala a Sheparda /1974/ je považována za první teoretické řešení nonmetrické faktorové analýzy. Prokázali, že za určitých podmínek /zanedbatelná náhodná chyba/ je možno dojít iterativním faktorovým postupem k vhodné metrické reprezentaci. Jejich postup rozšířili Woodward a Overall /1976/ i na situace, kde náhodná chyba měření hraje velkou roli. Vzhledem k tomu, že s navrženou variantou faktorové analýzy není k dispozici příliš mnoho zkušeností, nelze ji zatím doporučit k přímému využití v praxi, ale spíše jen k pokusnému užití. Bude však na místě systematicky sledovat další vývoj této problematiky.

V zaměřených studiích, které se v poslední době zabývaly metodickou problematikou pořadových škál, se znovu objevila otázka velikosti chyby vyplývající z pořadí prezentace. Rozumí se tím poměrně stabilní vliv pořadí, ve kterém jsou podněty předkládány k posouzení. Do jisté míry se prosazuje i v pořadí, které sestaví posuzovatel. Podněty na prvých místech v seznamu bývají spíše přeceňovány na úkor podnětů, které se objevují až na konci. Velikost této chyby je těžko předvídatelná a závisí nejen na obsahové stránce podnětů, ale i na jejich homogenitě a na délce celé řady /Rand a Wagner, 1975/. Z dalších metodických studií /Wagner, 1976, Wagner a Daubney, 1976/ vyplývá, že této chybě podléhají výrazněji nezacvičení posuzovatelé či případně posuzovatelé, kteří nejsou dostatečně motivováni. Wagner a spol. doporučují, aby se pořadí předkládaných podnětů systematicky měnilo /podle některého z vyvážených projektů/; vliv pořadí se tím sice nevyloučí, ale rozloží se zhruba stejně na všechny podněty. Variance výsledků se tím sice zvýší, ale systematická chyba se rozloží.

XVII.2.b. Vyhodnocování výsledků, získaných metodou sumovaných odhadů

Cooper /1976/ se zabýval analýzou dat, získaných při použití Likertových škál. Při odvození nového statistického testu vycházel z pravděpodobnosti distribuce jednotlivých odpovědí; předpokládá se, že volba kterékoliv kategorie má stejnou pravděpodobnost. Jde o to zjistit, zda získané výsledky jsou distribuovány do jednotlivých kategorií rovnoměrně či nikoliv.

Předpokládejme následující příklad: 30 osob posuzovalo určitý film na pětibodové škále. Předpokládáme, že vzdálenost mezi body na škále je stejná a že při posuzování se může každá škálová kategorie objevit stejně často. Po sloučení všech posudků dostaneme následující výsledek:

	1	2	3	4	5	
Zcela špatný	0	2	9	12	7	velmi dobrý

Znamená to tedy, že žádný posuzovatel nepovažoval film za zcela špatný, 2 jej považují za spíše špatný, 9 jej posoudilo jako průměrný atd. Výsledkům se přikládají váhy od 1 do 5 /jsou uvedeny pod frekvenční tabulkou/. Cooper odvodil, že pravděpodobnost náhodného rozložení posudků můžeme testovat podle vztahu

$$z = \frac{S - \frac{N(r+1)}{2}}{\sqrt{\frac{N(r^2-1)}{12}}}$$

kde N = počet posuzovatelů /=30/

r = počet posuzovaných kategorií /=5/

S = celkové skóre, které se získá součtem všech vážených výsledků; v našem případě $S = 1{\cdot}0 + 2{\cdot}2 + 3{\cdot}9 + 4{\cdot}12 + 5{\cdot}7 = 114$

Hodnota z se tedy vypočte takto:

$$z = \frac{114 - \frac{30 \cdot 6}{2}}{\sqrt{\frac{30 \cdot 24}{12}}} = 3{,}1$$

což je hodnota velmi významná /viz tabulky hodnot z/. Znamená to tedy, že skupina posuzovatelů nehodnotila film náhodně, ale považuje ho za spíše dobrý.

Pro menší počet pozorovatelů vyvinul Cooper tabulky, z kterých je možno vyčíst, zda zjištěné distribuce se liší od distribuce náhodné či nikoliv. Tabulky pro různá N a různá r by byly příliš rozsáhlé, ale je možno je snadno vypočíst samočinným počítačem. Zjednodušená tabulka významných hodnot je připojena k citované práci. Cooperův postup je velmi rychlý a užitečný; jeho užití na empirický materiál se dobře osvědčil.

XVII.2.c. D o d a t e k k m e t o d ě p á r o v é h o s r o v n á v á n í p o d n ě t ů

Při výkladu metody párového srovnávání podnětů /kap.III.2./ jsem upozornil, že škálovací proceduru nelze realizovat, jestliže při posouzení některého páru se objeví naprostá shoda mezi posuzovateli /tj. 100% preferuje podnět x před podnětem y/. Obvykle se pak postupuje tak, že podobná data se vyloučí a pracuje se s neúplnou maticí. V poslední době se objevilo určité řešení /Krus a Krus, 1977/, které se opírá o Mc Nemarův /1947/ postup pro testování rozdílů mezi závislými frekvencemi či proporcemi.

Postup demonstruji na zjednodušeném příkladě. Předpokládajme, že 50 posuzovatelů posuzuje párově 3 podněty /A, B, C/ a zjistí se výsledky, které jsou shrnuty v Tab.č.XVII.4.

Tab.č.XVII.4. Výchozí matice a její transformace na P matici

Výchozí matice dat $(f_{i\,j})$ při párovém posuzování 3 podnětů				Transformace výchozí matice na p - matice $(p_{i\,j} = f_{i,j} / N)$			
	A	B	C		A	B	C
A		50	40	A	0,00	1,00	0,80
B	0		20	B	0,00		0,40
C	10	30		C	0,20	0,60	

Je zřejmé, že všech 50 posuzovatelů dalo přednost podnětu B před A, 40 posuzovatelů preferuje C před A a konečně 30 volí raději B než C. V pravé části tabulky jsou výchozí údaje transformovány na matici pravděpodobnostních hodnot. Z P-matice se získávají hodnoty z; v našem příkladě není přímý postup možný.

Hodnota z pro srovnání podnětů A, B by byla . Proto se hodnoty z nevyhledávají přímo v tabulkách pro z-hodnoty, ale korigují se podle postupu, odvozeného Mc Nemarem a to buď přímo z frekvenčních hodnot dle vztahu

$$z = \frac{b - c}{(b-c)^{1/2}}$$

anebo z P-matice

$$z = \frac{b - c}{[(b+c) I N]^{1/2}}$$

V obou vzorcích symboly b, c značí jednotlivé hodnoty, které leží symetricky ve výchozí matici. Vyjdeme-li z matice frekvenčních hodnot, pak se jedná o následující páry : 50 a 0; 40 a 10; 20 a 30. Po výpočtu hodnoty z pak stanovíme běžným způsobem konečnou škálu; postup je zachycen v Tab.č.XVII.5.

Tab.č.XVII.5. Matice z_{ij}, vypočtená z Tab.č.XVII.4. a stanovení škálových hodnot.

	A	B	C
A	0,00	7,07	4,24
B	-7,07	0,00	-1,41
C	-4,24	1,41	0,00
$\sum$	-11,31	8,48	2,83
$\bar{x}$	-3,77	2,83	0,94
H	0,00	6,60	4,71

$$(H = \bar{x} + 3,77)$$

Při volbě tohoto postupu je však třeba upozornit, že není přímo srovnatelný s postupem, uvedeným v kap. III. Mc Nemarovým postupem se získávají vyšší hodnoty; relativní poloha jednotlivých bodů na škále zůstává při obou postupech sice zhruba stejná, ale absolutní hodnoty jsou odlišné. Není proto možné přímo srovnávat výsledky, získané oběma postupy.

XVII. 2.d. Behaviorálně zakotvené škály jako varianta posuzovacích škál

Při používání posuzovacích škál /hlavně pro účely průmyslové psychologie/ se po řadu let objevuje námitka, že sledované procesy či podněty bývají složité a proto snaha posoudit je na jednodimenzionální škále nemusí být přiměřená. Kromě toho se často žádá, aby posuzovatel vystupoval spíše v roli hodnotitele než v roli pozorovatele. Mnohé posuzovací škály mají různou hodnotu spolehlivosti často jako důsledek toho, zda se posuzuje specifické chování /bylo rozhodnutí správné, reakce pohotová apod./ či vysuzovaný rys osobnosti. Před časem se objevil pokus vytvořit behaviorálně zakotvené škály, které by umožňovaly posoudit činnost člověka z hlediska většího počtu dimenzí. Postup byl navržen již před delší dobou /Smith, Kendall, 1963/; během doby postupně krystalizoval a ustálil se na určité podobě. Byl kriticky analyzován Schwabem a spol. /1975/ pro účely posuzování pracovních činností.

Vývoj behaviorálně zakotvených škál má několik kroků:

a/ Stanovení kritických bodů. Osoby, které dobře znají sledovanou činnost, popisují specifické příklady účinného či neúčinného chování; do jisté míry tato fáze připomíná Flanaganovu techniku kritických případů.

b/ Dimenze činnosti. Poznatky z předcházejícího bodu shrne nyní skupina jiných odborníků do menší množiny dimenzí, které předběžně definují. Obvykle se užívá 5 - 10 dimenzí. Přitom je snaha vyhnout se komplexním a globálním dimenzím či vlastnostem, které se špatně popisují.

c/ Revize dimenzí. Jiná skupina odborníků je seznámena s dimenzemi, stanovenými v bodě b/ a jejich úkolem je přiradit k nim výchozí /bod a/ konkrétní příklady chování. Stanoví se kriterium shody mezi řazením v bodě b/ a při revizi /obvykle kolem 80%, ale při složitých činnostech se připouští i 50% jako dostatečné/. Dále se pak pracuje pouze s těmi položkami, které byly stabilně zařazovány.

d/ Škálování událostí /položek/. Další skupina odborníků nyní posuzuje, jak dalece popsané příklady chování jsou ukázkou vhodné či nevhodné činnosti. Obvykle se volí škála o 7 - 9 bodech. Průměrné posouzení stanoví stupeň, jak dalece daný příklad reprezentuje správné chování. Směrodatná odchylka umožňuje odhadnout shodu mezi posuzovateli; obvykle se žádá, aby nebyla větší než 1.50 /při sedmibodové škále/.

e/ Konečná verze škály. Skupina příkladů /5 - 7/, které odpovídají požadavkům, uvedeným sub b/ a c/ pak popisuje a do jisté míry i behaviorálně definuje jednotlivé dimenze chování. Konečná škála se skládá z několika podškál /většinou graficky vyjádřených/ a úkolem pokusné osoby je umístit posuzovaný jev na všechny dimenze.

Výhoda behaviorálně zakotvených škál je v tom, že se na jejich vývoji podílí řada odborníků: jejich názory jsou konfrontovány různými cestami a tím se jejich objektivita zvyšuje. Výsledné škály jsou záměrně orientovány na různé formy činnosti a programově se snaží vyhnout posuzování vysouzených osobnostních vlastností. Není zanedbatelné ani to, že takto vyvinuté škály se mohou užívat i pro účely výcviku nových zaměstnanců, při rekvalifikačních kursech atd. Řadu výzkumů, které poukazují na výhody i úskalí tohoto typu posuzovacích škál uvádí Schwab a spol. /1975/.

Jistě ani tento typ posuzovacích škál není bez úskalí. Poměrně pracný postup, který předpokládá větší počet znalců, není vždy proveditelný. Tak jako u ostatních metod, které se opírají o mínění expertů, i zde stojí otázka, zda jsou výsledky aplikovatelné do širší oblasti zkoumání /analogické činnosti, analogické profese atd./. Nabízí se však i další možnost využití behaviorálně zakotvených škál a to jejich kombinace s Guttmanovým škálogramem. Je možné, že právě tato kombinace by mohla přinést další metodický pokrok. Pokud by se zdařilo zakotvit jednotlivé dimenze v několika bodech, znamenalo by to značné zjemnění celé posuzovací procedury.

XVII. 2.e. D o p l n ě k k e š k á l o g r a m o v é a n a l ý z e

Zdá se, že Guttmanovým škálám se v poslední době věnuje větší pozornost. Množí se studie jak s jejich systematickým využitím pro různé výzkumné účely, tak se zaměřením pro další zjemňování celé procedury.

V prvé řadě se objevuje snaha domyslet a technicky dořešit pravidla dichotomizování odpovědí na kladné a záporné a přitom maximálně snížit počet chyb, které přitom vznikají. Podnětná je studie Wimberleyova /1976/, který navrhl dvě nové alternativy pro koeficient reprodukovatelnosti /viz kap. VIII.3.a./. Srovnává několik možných způsobů, jak získané odpovědi rozdělit. Proti formálním kriteriím Guttmanovým či jeho následníků, kteří postupují tak, aby koeficient reprodukovatelnosti byl co nejmenší, navrhuje postupy, které se opírají spíše o věcný rozbor získaných dat a z jejich logiky odvodit dělící bód na škále. Zatím však nejsou dostatečné podklady k tomu, aby bylo možno rozhodnout, zda navržené postupy jsou skutečně vhodnější; na první pohled se to zdá pravděpodobné, ale přece jen by bylo dnes předčasné pokoušet se o konečné rozhodnutí.

Zájem o jednodimenzionální škály vede i k vytváření některých škálovacích technik, které sice vycházejí z Guttmanových úvah, ale rozvíjejí alternativní postupy. Patří sem v prvé řadě teorie řazení /ordering/, která rozšiřuje použití škálogramu i na takové položky, které sice leží na jediné dimenzi, ale není mezi nimi lineární vztah. Základní myšlenky i přehled dosavadních zkušeností uvádí Bart /1976/; postup však není ještě zcela ustálen a proto jej zatím neuvádím.

L i t e r a t u r a

Balcar, E., Kožený, J.: Slova a čísla: Studie ke škálování verbálních kvantifikátorů. Československá psychologie, 1975, 19 447-458.

Barlett, C.J., Edgerton, H.A.: Stanine Values for Ranks for Different Number of Things Ranked. Educational and Psychological Measurement, 1966, 26, 287-290.

Beals, R., Krantz, D.H., Tversky, A.: Foundations of Multidimensional Scaling. Psychological Review, 1968, 75, 127-142.

Běljajev, E.V.: Problemy sociologičeskogo izměrenija. Voprosy filosofii, 1967, 7.

Berka, K.: Kritický rozbor operacionalistického pojetí měření. Filosofický časopis, 1971, 19, 209-221.

Berka, K.: Škály měření. ÚFS ČSAV, Praha, 1972.

Brogden, H.E., Taylor, E.K.: The Theory and Classification of Criterion Bias. Educational and Psychological Measurement, 1950, 10, 159-186.

Brown, S.R.: Bibliography on Q-techique and its Methology. Perceptual and Motor Skills, 1968, 26, 587-613 /Monograph Supplement 4-V26/.

Břicháček, V.: Rozbor experimentálně navozených rozhodovacích procesů. Československá psychologie, 1968, 12, 468-471.

Břicháček, V., Brichcín, M.: Problèmes méthodologiques de l'analyse multivariée des activités psychomotrices. Bulletin de Psychochologie, 1969, 12, 792-798.

Břicháček, V., Hampejsová, O.: Neparametrické techniky statistického hodnocení v psychologickém výzkumu II. Psychologické štúdie, Bratislava, SAV, 1963, 4, 221-247.

Campbell, D.T., Fiske, D.W.: Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. Psychological Bulletin, 1959, 56, 81-105.

Campbell, J.P., Dunnette, M.D., Arvey, R.D., Hellervik, L.V.,: The Development and Evaluation of Behaviorally Based Rating Scales. Journal of Applied Psychology, 1973, 57, 15-22.

Cantril, H.: A Study of Aspirations. Scientific American, 1963, 208, 41-45.

Carroll, J.D.: Individual Differences And Multidimensional Scaling, in: Shepard, R.N., et al.: Multidimensional Scaling, Vol. I., New York, Seminar Press, 1972. 105-156.

Carroll, J.D., Chang, J.J.: Analysis of Individual Differences in Multidimensional Scaling via an N-way Generalisation of "Eckart -Young" Decomposition. Psychometrika, 1970, 35, 283-319.

Centra, J.A.: Validation by the Multigroup-Multiscale Matrix. Educational and Psychological Measurement, 1971, 31, 675-683.

Clauss, G.: Zur statistischen Auswertung von Befragungsergebnissen. Probleme und Ergebnisse der Psychologie, 1962. III./IV., 127-136.

Clauss, G.: Zur Methodik von Schätzskalen in der empirischen Forschung. Probleme und Ergebnisse der Psychologie, 1968, 26, 7-53.

Coombs, C.H.: Psychological Scaling without a Unit of Measurement. Psychological Review, 1950, 57, 145-158.

Coombs, C.H.: Theory and Methods of Social Measurement, in: Festinger, L., Katz, D.: Research Methods in the Behavioral Sciences, New York, Dryden Press, 1953, 471-535.

Coombs, C.H.: On the Use of Inconsistency of Preferences in Psychological Measurement. Journal of Experimental Psychology, 1958, 55, 1-7.

Coombs, C.H.: A Theory of Data. Psychological Review, 1960, 67, 143-159.

Coombs, C.H.: A Theory of Data. New York, Wiley, 1964.

Coombs, C.H., Pruitt, D.: Components of Risk in Decision Making: Probability and Variance Preferences. Journal of Experimental Psychology, 1960, 60, 265-277.

Cosgrove, D.J.: Diagnostic Rating of Teacher Performance. Journal of Educational Psychology, 1959, 50, 200-204.

Coyne, L.; Holzman, P.S.: Three Equivalent Forms of a Semantic Differential Inventory. Educational and Psychological Measurement 1966, 26, 665-674.

Dember, W.N., Earl, R.W.: Analysis of Exploratory , Manipulatory and Curiosity Behavior. Psychological Review, 1957, 50, 514-518.

Domas, S., Tiedeman, D.V.: Teacher Competence: An Annotated Bibliography. Journal of Experimental Education, 1950, 19, 101-218.

Dunn-Rankin, P., King, F.J.: Multiple Comparisons in Simplified Rank Method of Scaling. Educational and Psychological Measurement, 1969, 29, 315-329.

Edwards, A.L.: On Guttman's Scale Analysis. Educational and Psychological Measurement, 1948, 8, 313-318.

Edwards, A.L.: Techniques of Attitude Scale Construction. New York, Wiley, 1957.

Edwards, A.L., Kenney, K.C.: A Comparison of Thurstone and Likert Techniques of Attitude Scale Construction. Journal of Applied Psychology, 1946, 30, 72-83.

Edwards, A.L., Kilpatrick, F.P.: A Technique for the Construction of Attitude Scale. Journal of Applied Psychology, 1948, 32, 374-384.

Edwards, A.L., Kilpatrick, F.P.: Scale Anylysis and the Measurement of Social Attitudes. Psychometrika, 1948, 13, 99-114.

Ekehammar, B.: Comparative Study of Some Multidimensional Vector Models for Subjective Similarity. Scandinavian Journal of Psychology, 1972, 13, 190-197.

Ekehammar, B.: Two Multidimensional Methods Applied to Two Types of Stimuli. Perceptual and Motor Skills, 1972, 34, 535-542.

Ekman, G.: Dimensions of Color Vision. Journal of Psychology, 1954, 38, 467-474.

Ekman, G.: A Direct Method for Multidimensional Ratio Scaling. Psychometrika, 1963, 28, 33-41.

Ekman, G., Hosman, J., Lindström, B.: Smoothness and Preference: A Study of Quantitative Relations in Individual Subjects. Journal of Experimental Psychology, 1965, 70, 18-26.

Ekman, G., Sjöberg, L.: Scaling. Annual Review of Psychology, 1965, 16, 451-474.

Ertel, S.: Die Standardisierung eines Eindrucksdifferentials. Zeitschrift für experimentelle und angewandte Psychologie, 1965, 12, 22-58 /a/.

Ertel, S.: Weitere Untersuchungen zur Standardisierung eines Eindrucksdifferentials. Zeitschrift für experimentelle und angewandte Psychologie, 1965, 12, 177-208 /b/.

Esser, U.: Skalierungsverfahren. in: Friedrich, W.: Methoden der marxistisch-leninistischen Socialforschung. Berlin, DVW, 1970, 184-242.

Faverge, J.M.: Sur la mesure en psychologie. Journal de Psychologie, 1954, 47-51, 417-430.

Fechner, G.T.: Elemente der Psychophysik. Leipzig, Breitkopf, Härtel, 1860.

Fischer, S.T., Weiss, D.J., Dawis, R.V.: A Comparison of Likert and Pair Comparisons Techniques in Multivariate Attitude Scaling. Educational and Psychological Measurement, 1968, 28, 81-94.

Foster, C.C., Horwath, W.J.: A Study of a Large Sociogram III. Reciprocal Choice Probabilities as a Measure of Social Distance. Behavioral Sciences, 1971, 16, 429-433.

Golding, S.L., Knudson, R.M.: Multivariable-Multimethod Couvergence in the Domain of Interpersonal Behavior. Multivariate Behavioral Research, 1975, 10, 425-448.

Goodenough, W.H.: A Technique for Scale Analysis. Educational and Psychological Measurement, 1944, 4, 179-190.

Green, B.F.: Attitude Measurement. in: Lindzey, G.: Handbook of Social Psychology, Vol. 1, Cambridge, 1954, 335-369.

Greenberg, M.G.: A. Modification to Thurstone's Law of Comparative Judgement to Accomodate a Judgement Category of "Equal" or "No Difference". Psychological Bulletin, 1965, 64, 108-112.

Guilford, J.P.: Psychometric Methods. New York, Mc-Graw - Hill., 1954.

Guilford, J.P.: Preparation of Item Scores for the Correlations between Persons in a Q-factor Analysis. Educational and Psychological Measurement, 1963, 23, 13-22.

Gutjahr, W.: Zur Skalierung psychischer Eigenschaften. Probleme und Ergebnisse der Psychologie, 1968, 23, 9-40.

Gutjahr, W.: Die Messung psychischer Eigenschaften. Berlin, DVW, 1972.

Guttman, L.: A Basis for Scaling Qualitative Data. American Sociological Review, 1944, 9, 139-150. Reprint in: Fishbein, M.: Readings in Attitude Theory and Measurement, New York, Wiley, 1967, 96-107.

Guttman, L.: The Cornell Technique for Scale and Intensity Analysis. Educational and Psychological Measurement, 1947, 7, 247-279.

Guttman, L.: The Principal Components of Scalable Attitudes. in: Lazarsfeld, P.F.: Mathemathical Thinking in Social Sciences, Glencoe, Illinois, 1954.

Guttman, L.: Order Analysis of Correlation Matrices. in: Cattell, R.B.: Handbook of Multivariate Experimental Psychology, Chicago, McNally, 1966, 439-458.

Guttman, L., Suchman, E.A.: Intensity and Zero Point for Attitude Analysis. American Sociological Review, 1947, 12, 55-67. Reprint in: Fishbein, M.: Readings in Attitude Theory and Measurement, New York, Wiley, 1967, 267-276.

Hall, R.F.: Application of Unfolding Theory to the Measurement of Attitudes. Educational and Psychological Measurement, 1970, 30, 621-637.

Harris, G.G.: Bibliography - Multidimensional Scaling /1960-1971/. Bell Laboratories, Rep. No. 205, 1972.

Hays, W.L.: Quantification in Psychology. Brooks, Belmont, California, 1967.

Hofstätter, P.R.: Einführung in die Sozialpsychologie. Stuttgart, Kröner, 1959.

Itelson, L.B.: Matematičeskije i kibernetičeskije metody v pedagogike. Moskva, Prosvěščenije, 1964.

Jackson, D.N.: Multimethod Factor Analysis in the Evaluation of Convergent and Discriminant Validity. Psychological Bulletin, 1969, 72, 30-49.

Jöreskog, K.G.: A General Approach to Confirmatory Maximum Likelihood Factor Analysis. Psychometrika, 1969, 34, 183-202.

Kelly, G.A.: Non-parametric Factor Analysis of Personality Theories. Journal of Individual Psychology, 1963, 19, 115-147.

Kendall, M.G.: Rank Coorelation Methods, London, Griffin, 1948.

Kerlinger, F.N.: Základy výzkumu chování. Praha, Academia, 1972.

Kilpatrick, F.P., Cantril, H.: Self-Anchoring Scaling. Journal of Individual Psychology, 1960, 16, 158-173.

Kožený, J.: Psychometrické ověření FIRO-B škál na české populaci. Československá psychologie, 1973, 17, 36-53.

Krantz, D.H.: Conjoint Measurement: The Luce-Turkey Axiomatisation and Some Extensions. Journal of Mathematical Psychology, 1964, 1, 248-277.

Krasemann, I.: Skalierungsverfahren. in: Stoljarow, V.: Zur Technik und Methodologie einiger quantifizierender Methoden der soziologischen Forschung. Berlin, Dietz, 1966, 221-376.

Krech, D., Crutchfield, R.S., Ballachey, E.L.: Člověk v společnosti. Bratislava, SAV, 1968.

Kukla, F., Sydow, H.: Zur metrischen Darstellung phänomenaler Ähnlichkeiten bei komplexen Reizmaterial. Zeitschrift für Psychologie, 1970, 178, 111-142.

Kull, I.G.: O někotorom klasse psychometričeskich modelej. Matematičeskaja psichologija I., Tartu, 1974, 29-46.

Kyverjalg, A.A.: Voprosy metodiky pedagogičeskich issledovanij. Tallin, Valgus, 1971.

Lafargue, P.: Vzpomínky na Marxe. Praha, Svoboda, 1945.

Lazarsfeld, P.F.: A. Conceptual Introduction to Latent Structure Analysis. in: Lazarsfeld, P.F.: Mathematical Thinking in the Social Sciences. Glencoe, Illinois, Free Press, 1954, 349-387.

Lazarsfeld, P.F.: Latent Structure Analysis. in: Koch, S.: Psychology: A Study of a Science, Vol. 3, New York, Mc-Graw-Hill, 1959, 476-543.

Lienert, G.A.: Verteilungsfrei Methoden in der Biostatistik. Meisenheim, Hain, 1962.

Likert, R.: A Technique for the Measurement of Attitudes. Archives of Psychology, 1932, No. 140, 44-53. Reprint in: Fishbein, M.: Readings in Attitude Theory and Measurement, New York, Wiley, 1967, 90-96.

Lingoes, J.C.: Multiple Scalogram Analysis: A Set Theoretic Model for Analysing Dichotomous Items. Educational and Psychological Measurement, 1963, 23, 501-524.

Lingoes, L.C.: The Multivariate Analysis of Qualitative Data. Multivariate Behavioral Research, 1968, 3, 61-94.

Lingoes, J.C.: A General Survey of the Guttman - Lingoes Nonmetric Program Series. in: Shepard, R.N. et al.: Multidimensional Scaling, Vol. I. New York, Seminar Press, 1972, 52-67.

Lissitz, R.W., Green, S.B.: Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach. Journal of Applied Psychology, 1975, 60, 10-13.

Loevinger, J.: The Technique of Homogeneous Tests Compared with Some Aspects of "Scale Analysis" and Factor Analysis. Psychological Bulletin, 1948, 45, 507-530.

Luce, R.D.: Individual Choice Behavior. New York, Wiley, 1959.

Luce, R.D., Tukey, J.W.: Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. Journal of Mathematical Psychology, 1964, 1. 1-27.

Lundberg, V., Devine, B.: Negative Similarities. Educational and Psychological Measurement, 1975, 35, 797-807.

Mac Callum, R.C.: Relations between Factor Analysis and Multidimensional Scaling. Psychological Bulletin, 1974, 81, 505-516.

Maxwell, A.E.: The Logistic Transformation in the Analysis of Paired-Comparison Data. British Journal of Mathematical and Statistical Psychology, 1974, 27, 62-71.

Maxwell, A.E., Pilliner, A.E.G.: Deriving Coefficients of Reliability and Agreement for Ratings. British Journal of Mathematical and Statistical Psychology, 1968, 21, 105-116.

Meehl, P.E.: The Congnitive Activity of the Clinician. American Psychologist, 1960, 15, 19-27.

Měření ve společenských vědách /Závěry z diskuse o měření v sociologii/. Sociologický časopis, 1973, 9, 514-519.

Montgomery, H,: Direct Estimation: Effect of Methodological Factors on Scale Type. Göteborg Psychological Reports, 1971, 1, No. 9.

Moreno, J.L.: Who shall survive? New York, Beacon, 1953.

Moreno, J.L.: Die Grundlagen der Soziometrie. Köln, Westdeutscher Verlag, 1954.

Neff, W.S., Cohen, J.: A Method for the Analysis of the Structure and Interal Consistency of Q-sort Arrays. Psychological Bulletin, 1967, 68, 361-368.

Nunnally, J.: The Analysis of Profile Data. Psychological Bulletin, 1962, 54, 311-319.

Orlik, P.: Eine Modellstudie zur Psychophysik des Polaritätsprofils. Zeitschrift für experimentelle und angewandte Psychologie, 1965, 12, 614-647.

Orlik, P.: Eine Technik zur erwartungstreuen Skalierung psychologischer Merkmalsräume aufgrund von Polaritätsprofilen. Zeitschrift für experimentelle und angewandte Psychologie, 1967, 14, 616-650.

Osgood, C.E.: The Nature and Measurement of Meaning. Psychological Bulletin, 1952, 49, 197-237.

Osgood, C.E., Suci, G.J., Tannenbaum, P.H.: The Measurement of Meaning. Urbana, University of Illinois Press, 1957.

Petrusek, M.: Sociometrie. Praha, Svoboda, 1969.

Rapoport, A., Horwath, W.J.: A Study of a Large Sociogram. Behavioral Sciences, 1961, 6, 279-291.

Reuchlin, M.: Meranie v psychológii. in: Fraisse, P.: Kapitoly z experimentálnej psychológie, Bratislava, SPN, 1967, 137-180.

Rorer, L.G.: The Great Response Style Myth. Psychological Bulletin, 1965, 63, 129-156.

Samsonov, J.B.: Některé problémy sociologického měření. Sociologický časopis, 1973, 9, 289-294.

Shepard, R.N.: Metric Structures in Ordinal Data. Journal of Mathematical Psychology, 1966, 3, 287-315.

Shepard, R.N.: A Taxonomy of Some Principal Types of Data and of Multidimensional Methods for their Analysis. in: Shepard, R.N., et al.: Multidimensional Scaling, Vol. I, New York, Seminar Press, 1972, 23-49.

Shepard, R.N.: Representation of Structure in Similarity Data: Problems and Prospects. Psychometrika, 1974, 39, 373-422.

Shepard, R.N., Romney, A.K., Nerlove, S.B.: Multidimensional Scaling, Vol. I, Theory. New York, Seminar Press, 1972a.

Shepard, R.N., Romney, A.K., Nerlove, S.B.: Multidimensional Scaling, Vol. II, Applications. New York, Seminar Press, 1972b.

Scheuch, E.K.: Skalierungsverfahren in der Sozialforschung. in: König, R.: Handbuch der empirischen Sozialforschung, Stuttgart, Enke, 1962, 348-384.

Schmidt, H.D.: Posudzovanie ľudského správania rating-škálami. Bratislava, SPN, 1970.

Siegel, S.: Nonparametric Statistics for the Behavioral Sciences. New York, McGraw-Hill, 1956.

Spence, I., Domoney, D.W.: Single Subject Incomplete Designs for Nonmetric Multidimensional Scaling. Psychometrika, 1974, 39, 469-490.

Stephenson, W.: The Study of Behavior: Q-Technique and its Methodology. Chicago, University of Chicago Press, 1953.

Stevens, S.S.: On the Theorv of Scales of Measurement. Science, 1946, 103, 677-680.

Stevens, S.S.: Mathematics, Measurement, and Psychophysics. in: Stevens, S.S.: Handbook of Experimental Psychology. New York, Wiley, 1951, 1-49.

Stewart, T.R.: Generality of Multidimensional Representations. Multivariate Behavioral Research, 1974, 9, 507-519.

Stone, L.A., Coles, G.J.: Correlational Similarity: the Basis for a New Revised Method of Similarity Analysis. Studia Psychologica, 1970, 12, 258-264.

Stone, L.A., Coles, G.J.: Dimensions of Color Vision Revisited. Journal of Psychology, 1971, 77, 79-87.

Suchman, E.A.: The Logic of Scale Construction. Educational and Psychological Measurement, 1950, 10, 79-93.

Šafář, L.: Škálování. in: Pergler, P.: Vybrané techniky sociologického výzkumu. Praha, Svoboda, 1969, 347-686.

Taylor, J.B.: Rating Scales as Measures of Clinical Judgement. Educational and Psychological Measurement, 1968, 28, 747-766.

Taylor, J.B., Haefele, E., Thompson, P., O'Donoghue, C.: Rating Scales as Measures of Clinical Judgement II: The Reliability of Example-Anchored Scales under Conditions of Rater Heterogeneity and Divergent Behavior Sampling. Educational and Psychological Measurement, 1970, 30, 301-310.

Taylor, J.B., Parker, H.A.: Graphic Ratings and Attitude Measurement: A Comparision of Research Tactics. Journal of Applied Psychology, 1964, 48, 37-42.

Taylor, J.B., Ptacek, M., Carithers, M., Griffin, C., Coyne, L.: Rating Scales as Measures of Clinical Judgment III: Judgments of the Self on Personality Inventory Scales and Direct Ratings. Educational and Psychological Measurement, 1972, 32, 543-557.

Tent, L.: Schätzverfahren in der Unterrichtsforschung. in: Ingenkamp, K.: Handbuch der Unterrichtsforschung, Teil I., Weinheim, Beltz, 1970, 853-999.

Thurstone, L.L.: A Law of Comparative Judgment. Psychological Review 1972, 34, 273-286.

Thurstone, L.L., Chave, E.J.: The Measurement of Attitude. Chicago, Illinois, 1929.

Torgerson, W.S.: Multidimensional Scaling: I - Theory and Method. Psychometrika, 1952, 17, 401-419.

Torgerson, W.S.: Theory and Methods of Scaling. New York, Wiley, 1958.

Torgerson, W.S.: Scaling. in: Sills, D.L.: International Encyclopedi of the Social Sciences. New York, McMillan, 1968, Vol. 14, 25-39.

Váchová, D., Šmolka, P.: Škála pro měření přátelské dimenze interpersonálních vztahů ve věkové skupině 18-25 let. Českolsovenská psychologie, 1975, 19, 67-71.

Walenta, K.: Podstawowe pojecia teorii pomiaru. in: Kozielecki, J.: Problemy psychologii matematycznej, Warszawa, PWN, 1971, 43-64.

Walace, D.: Cluster Analysis. in: Sills, D.L.: International Encyclopedia of the Social Sciences. New York, McMillan, 1968, Vol. 3, 519-524.

Wittenborn, J.R.: Contributions and Current Status of Q-Methodology. Psychological Bulletin, 1961, 58, 132-142.

Zander, R.: Ein Programm zur numerischen Auswertung von Schätzskaler in der empirischen Forschung. Probleme und Ergebnisse der Psychologie, 1972, 41, 85-89.

Zavala, A.: Development of the Forced-Choice Rating Scale Technique. Psychological Bulletin, 1965, 63, 117-124.

Zinnes, J.L.: Scaling. Annual Review of Psychology, 1969, 20, 447-478.

Zdravomyslov, A.G., Jadov, V.A.: Opyt konkretnovo issledovanija otnošenija k trudu. Voprosy filosofii, 1964, 4, 72-84.

Literatura ke kapitole č. XVII.

Bart, W.M.: Some results of ordering theory for Guttman scaling; Educational and Psychological Measurement, 1976, 36, 141-148.

Brušlinskij, A.V.: Psichologija myšlenija i poňatija množestv; in: Rubachin, V.F. et al.: Psichologija i matematika, Moskva, Nauka, 1976.

Břicháček, V.: Neostré množiny a psychiatrická diagnostika; Československá psychiatrie, 1978, v tisku.

Břicháček, V.: Vyjádření dimenze introverze-extroverze pomocí neostrých množin; Československá psychiatrie, 1978, v tisku.

Cooper, M.: An exact probability test for use with Likert-Type scales: Educational and Psychological Measurement, 1976, 36, 647-655.

Gaines, B.R.: The fuzzy decade: A bibliography; International Journal of Man-Machine Studies, 1977, 9, 1-68.

Hersch, H.H., Caramazza, A.: A fuzzy set approach to modifiers and vagueness in natural language; Journal of Experimental Psychology; General, 1976, 105, 254-276.

Jarminckij, J.D.: Problema izmerenija i psychologičeskije metody; Voprosy psychologii, 1977, 22, 115-122.

Krus, D.J., Krus, P.H.: Normal scalling of the unidimensional dominance matrices: the domain referenced model: Educational and Psychological Measurement, 1977, 37, 189-193.

Kruskal, J.B., Shepard, R.N.: A nonmetric variety of linear factor analysis; Psychometrika, 1974, 39, 123-157.

Mc Nemar, Q.: Note on the sampling error of the difference between correlated proportions or percentages; Psychometrika, 1947, 12, 153-157.

Nowakowska, M.: Methodological problems of measurement of fuzzy concepts in social sciences; Behavioral Science, 1976

Rand, T.M., Wagner, E.E.: Generalisability of position error in ranking; Psychological Reports, 1975, 37, 811-814.

Schwab, D.P., Heneman, H.G.: Behaviorally anchored rating scales; Personnel Psychology, 1975, 28, 549-562.

Smith, P.C., Kendall, L.M.: Retranslation of expectations: an approach to the construction of unambiguous anchors for rating scales; Journal of Applied Psychology, 1963, 47, 149-155.

Straton, R.G.: Obtaining paired comparisons data from multiple rank orders using partially balanced incomplete block designs; Educational and Psychological Measurement, 1975, 35, 555-566.

Tżiner, A.E., Zahal, T.C.: Applitude measurement through use of computer constructed questionnaires; Educational and Psychological Measurement, 1977, 37, 241-243.

Wagner, E.E.: The Friedman two-way analysis of variance as a test for ranking error; Educational and Psychological Measurement, 1976, 36, 615-617.

Wagner, E.E., Daubney, J.H.: Position error in rating; Educational and Psychological Measurements, 1976, 36, 825-828.

Wimberley, R.C.: ALAM and ALAS: Questionning error assignments in unidimensional Guttman scaling; Educational and Psychological Measurement, 1976, 36, 361-367.

Woodward, J.A., Overall, J.E.: Factor analysis of rank-ordered data; Psychological Bulletin, 1976, 83, 864-867.

Zadeh, L.A.: Fuzzy sets; Information and Control; 1965, 8, 338-353

Zadeh, L.A.: A fuzzy algorithmic approach to the definition of complex or imprecise concepts; in Bossel H.: Systems theory in the Social Sciences, Birkhäuser, Basel, 1976, 202-282.

Rejstřík jmenný

Rejstřík věcný

Pozn. Dodatková kapitola není v rejstřících zahrnuta

PhDr. Václav Břicháček

Úvod do psychologického škálování

Obálku navrhol akad. maliar Daniel Fischer
Grafická úprava Zoltán Racsko
Jazyková úprava Jana Smolanová
Vydal Psychodiagnostické a didaktické testy, n.p., Bratislava
v roku 1978 ako svoju 15. publikáciu. Vydanie prvé,
počet strán 322.
Zodpovedný redaktor Igor Boor

Tlač vlastná AH 13,98 VH 14,59
Náklad 3000 ex.

Tematická skupina 02/9
Cena 30,- Kčs