

MASARYKOVA UNIVERZITA

FAKULTA SOCIÁLNÍCH STUDIÍ

ČÍTANKA PRO KURS PSY 722 — 6. ČÁST

Tento text slouží výhradně jako učební materiál pro studenty kursu „Metody výzkumu v psychologii“ (PSY 722), vyučovaného na Fakultě sociálních studií Masarykovy univerzity v Brně.

Pozorování

Obsah

1	Pozorování chování	1
1.1	Problémy pozorování chování	3
1.1.1	Pozorovatel	3
1.1.2	Validita a spolehlivost	4
1.2	Charakteristiky pozorování	6
1.2.1	Kategorie	6
1.2.2	Jednotky chování	7
1.2.3	Stupeň pozorovatelova vysuzování	8
1.2.4	Obecnost nebo aplikovatelnost	9
1.2.5	Požizování vzorků chování	9
1.3	Posuzovací škály	11
1.3.1	Typy posuzovacích stupnic	11
1.3.2	Slabiny posuzovacích stupnic	13
1.4	Příklady systémů pozorování	15
1.4.1	Měření interakce mezi matkou a dítětem — výzkum Merrillové	15
1.4.2	Posuzování učitelů studenty: dvě studie	15
1.4.3	Měření pedagogického vedení	16
1.4.4	Hodnocení rodičovského chování	17
1.4.5	Časové vzorky dětského chování	17
1.4.6	Pointově-časové vzorkování činností ve třídě	18
1.5	Zhodnocení behaviorálního pozorování	19
1.6	Studijní náměty	20

1 Pozorování chování¹

Jeden univerzitní výzkumný ústav, spolupracující s několika soustavami veřejného školství, studoval chování učitelů a jeho výsledky v práci žáků. Jednou ze zkoumaných hypotéz bylo, že přizpůsobiví a pozorní učitelé budou podněcovat žáky k originálnější a nezávislejší práci. Výzkumná skupina se rozhodla, že užije Ryansova třídního pozorovacího záznamu k měření chování učitelů a studentů.²

Třídní pozorovací záznam je osmnáctipoložkový objektivní pozorovací nástroj, v němž je každá položka dvoupólovým párem popisných adjektiv, jako stranící — nestranný, hrubý — vlídný, neústupný — přizpůsobivý, nudný — podnětný. Záznam se vyvíjel dvěma způsoby: studiem výzkumných zpráv a názorů, které jmenovaly vlastnosti, o nichž se mělo za to, že jsou u učitelů žádoucí a nezbytné k dobrému vyučování, a dále užitím techniky kritických případů. První přístup dal 46 charakteristik, které se zdály relevantními. V technice kritických případů se kladly velkému počtu zkušených učitelů dotazy na specifické chování úspěšných a neúspěšných učitelů ve specifických situacích — tedy jak by se zachoval ten nebo onen v té nebo oné situaci.³ Získalo se přes 500 kritických případů, které se pak redukovaly na seznam 25 dvoupólových zobecněných druhů chování.⁴ Tento seznam byl dále redukován do pracovní formy 18 (druhů) chování.

Využilo se zkušených učitelů jako pozorovatelů; byli pečlivě cvičeni v pozorovací technice a v užívání Záznamu. Mimo to byli ve vhodných intervalech znovu doškoleni. Každý pozorovatel používal rejstříku, obsahujícího popisy nebo behaviorální souhrny adjektivních párů. Tak například „neústupný — přizpůsobivý“ byl popsán jako „rigidní v podrobování rutine“ („neústupný“) a „pružný v přizpůsobování výkladů“ („přizpůsobivý“); „netrpělivý při přerušování a odbíhání“ („neústupný“) a „čelil nevšední situaci ve třídě kompetentně“ („přizpůsobivý“).⁵

Pozorovatelé sledovali velký počet učitelů. Každý učitel byl pozorován různými pozorovateli dvakrát. Jednotkou pozorovacího času byla jedna třídní hodina. Došlo-li mezi dvěma pozorovateli při popisování učitele k rozporům, využilo se třetího pozorovatele. Během vyučovací hodiny si pozorovatelé dělali poznámky. Každá položka záznamu byla zhodnocena na sedmibodové stupnici jako u tohoto páru:

nudný	1	2	3	4	5	6	7	N	podnětný
-------	---	---	---	---	---	---	---	---	----------

N znamená „neutrální“. Záznam byl vyplněn po každém pozorovacím období, po jedné vyučovací hodině.

Vypočítaly se průměry dvou nebo tří záznamových hodnocení. Každý učitel dosáhl tedy skóre u 18 popisných párů. Badatelé provedli četné analýzy s použitím položkových skóre, skóre skupin položek a celkových skóre. Tyto analýzy nás zde nezajímají, s výjimkou toho, že objektivní pozorování chování učitelů dala míry nezávisle proměnné.

Každý člověk pozoruje konání druhých. Díváme se na jiné lidi a nasloucháme jejich hovoru. Usuzujeme, co má ten druhý na mysli, když něco říká, a na základě těchto pozorování vyvozujeme charakteristické znaky, motivace, city a záměry druhých. Říkáme o někom: „Ten umí dobře odhadnout lidi!“, a máme tím na mysli, že má bystrý postřeh při pozorování chování, a že pokládáme za

¹Zdrojem této kapitoly jsou strany 489–509 z knihy *Základy výzkumu chování* od autora F. N. Kerlingera (Praha, Academia, 1972).

²D. Ryans: *Characteristics of Teachers*. Washington, D. C., American Council on Education, 1960, str. 86.

³Viz *ibid.*, str. 79–83. K delšímu a kompetentnímu povšechnému pojednání viz J. Flanagan: *The Critical Incident Technique*. *Psychological Bulletin*, LI/1954, str. 327–358.

⁴Ryans, c. d., str. 82.

⁵*Ibid.*, str. 91.

validní jeho úsudky o tom, co se skrývá za chováním.

Tento každodenní druh pozorování většiny lidí je však pro vědu neuspokojivý. Sociolog musí také pozorovat lidské chování, ale musí být nespokojen s neadekvátností nekontrolovaných pozorování. Hledá spolehlivá a objektivní pozorování, z nichž může čerpat validní závěry. Zachází s pozorováním chování jako s částí postupu měření: podle pravidel přiřazuje čísla předmětům, v tomto případě lidským aktům chování nebo jejich sledům.

1.1 Problémy pozorování chování

V podstatě jsou jen dva způsoby pozorování: můžeme sledovat lidi, jak něco dělají a říkají, a můžeme se také ptát lidí na jejich vlastní jednání a na chování druhých. Hlavními způsoby získávání informací je buď prožívání něčeho přímo, nebo přimět někoho, aby nám řekl, co se stalo. V této kapitole se zabýváme hlavně viděním a slyšením událostí a pozorováním chování, jakož i řešením vědeckých problémů, které vznikají z takového pozorování. Přistoupíme k problému dvěma způsoby:

1. budeme zkoumat hlavní problémy pozorování chování a formulovat jisté poučky a vodítka,
2. budeme zkoumat čtyři nebo pět živých příkladů solidních pozorovacích systémů.

1.1.1 Pozorovatel

Vážným problémem behaviorálního pozorování je sám pozorovatel. Připomeňme, že jednou z obtíží při interviewu byl interviewující, protože byl součástí měrného nástroje. U objektivních testů a stupnic se tento problém téměř nevyskytoval. Při behaviorálním pozorování je pozorovatel jak rozhodnou výhodou, tak i rozhodnou slabinou. Proč? Pozorovatel musí strávit informaci, odvozenou ze svých pozorování, a pak vyvodit dedukce o konstruktech. Pozoruje jisté chování, řekněme dítě bijící jiné dítě. Musí nějak zpracovat toto pozorování a vyvodit dedukci, že toto chování je manifestací konstruktů „agrese“ nebo „agresivního chování“, nebo dokonce „nepřátelství“. Silnou i slabou stránkou tohoto postupu jsou pozorovatelovy schopnosti úsudku. Nebýt úsudku, byl by strojový pozorovatel lepší než lidský (někdy může být lepší). Silnou stránkou je, že pozorovatel umí uvést pozorované chování ve vztah ke konstruktům nebo proměnným studie: spojuje tak chování a konstrukt. Jednou z obtíží měření, které se stále znovu vrací, je překlenutí mezery mezi chováním a konstruktem. Tuto mezeru pomáhají překlenout kompetentní pozorovatelé a dobře konaná pozorování.

Základní slabinou pozorovatele je, že může dělat ze svých pozorování zcela nesprávné dedukce následkem své lidské nedokonalosti. To je velmi vážným problémem. Vezměme dva extrémní případy. Na jedné straně předpokládejme, že nějaký pozorovatel, který je silně zaujat proti církevní školní výchově, pozoruje vyučování v církevní škole. Je pak jasné, že jeho předsudek může snadno znehodnotit pozorovací měrný nástroj. Odvoláme-li se opět na Ryansův třídní pozorovací záznam, může takový pozorovatel snadno posoudit přízpůsobivého učitele jako poněkud nepružného, protože nazírá na učení církevní školy jako na nepružné. Nebo opět může posuzovat skutečně podnětné chování učitele církevní školy jako nudné.

Na druhé straně předpokládejme, že by pozorovatel dovedl být zcela objektivním a že by nevěděl vůbec nic o veřejné ani o církevní výchově. V jistém smyslu by pak každé jeho pozorování bylo nezaujatým, a přece by bylo neadekvátním. Pozorování lidského chování vyžaduje kompetentní znalost onoho chování, dokonce i významu toho chování.

Hlavní obtíží je problém pozorovatelovy dedukce. Je tu však ještě jiný problém: pozorovatel může ovlivnit předměty pozorování pouze již tím, že se stává částí pozorovací situace. Ve skutečnosti a naštěstí to není vážný problém. Vlastně se to stává problémem (těm) nezasvěceným, kteří se snad domnívají, že lidé budou jednat odlišně, dokonce vyumělkovaně, jsou-li pozorováni. Klasickým pedagogickým případem toho druhu je domněnka, že učitelka, která je pozorována, zvláště svými nadřízenými, ukáže své nejlepší stránky. Má se za to, že bude jednat příkladným způsobem, který u ní není nutně obvyklým. To sice může být pravda, ale přehlídí se tu významný moment. Neuvažuje se, že učitelka neumí nebo nedokáže dělat to, co neumí, nebo nedokáže dělat. Nedovede

jednat způsobem, jakým se nenaučila jednat. Neumí být „přízpůsobivou“, abychom užili jednoho z Ryansových adjektiv, jestliže se nenaučila být přízpůsobivou.

Zdá se, že pozorovatelé mají malý účinek na situace, které pozorují.⁶ Zdá se, že se jedinci i skupiny dosti rychle adaptují na přítomnost pozorovatele a jednají, jak by obvykle jednali. To ovšem neznamená, že pozorovatel nemůže mít (na ně) účinek. Znamená to jen tolik, že jestliže si pozorovatel dá záležet, aby byl nenápadným a aby nezbudil u pozorovaných lidí pocit, že jsou posuzováni, pak je pozorovatel jako ovlivňující podnět většinou anulován.

1.1.2 Validita a spolehlivost

Z povrchního hlediska se nezdá nic přirozenějším, pozorujeme-li chování hrubé — vlídné nebo nespravedlivé — spravedlivé, než mít za to, že měříme to, co říkáme, že měříme. A když Ryans dává pozorovateli v glosáři záznamu dosti (ale nikoli příliš) podrobné definice hrubého — vlídného nebo nespravedlivého — spravedlivého, pak je tu pravděpodobně dobrá shoda mezi tím, co se měří, a tím, co se zamýšlelo měřit. Hlavním momentem snad je, že jsou-li jim dány definice, pozorovatelé dovedou aplikovat a také aplikují tyto termíny podobně na tytéž nebo podobné typy chování. Tak například část jeho glosáře pro nespravedlivého — spravedlivého zní: „Věnoval většinu pozornosti jednomu nebo několika žákům“ (nespravedlivý) a „Rozděloval svou pozornost na mnoho žáků“ (spravedlivý).⁷

Klade-li se však na pozorovatele větší interpretační břímě, může tím trpět validita (stejně jako spolehlivost). Příkladem toho je měření Fouriezose, Hutta a Guetzkowa, týkající se potřeb orientovaných na „já“⁸ v konferenčním chování.⁹ Pozorovatelům se dá jednoduchá desetibodová posuzovací stupnice v rozpětí od „Žádného projevu na sebe orientované potřeby“ až do „Veškerého chování sebe-orientovaného typu“; na této stupnici pak naznačují pozorovatelé integrovaná hodnocení egocentricky orientovaných potřeb členů skupiny. Řídí se pěti kategoriemi projevu potřeb a o těchto projevech si během konferencí dělají záznamy. I když se zdá, že se tento systém osvědčil, vzniká dojem, že je tu poněkud velká mezera mezi pozorovanými akty chování a vysouzenými potřebami.¹⁰ Čím více se břímě interpretace klade na pozorovatele, tím je pak větší problém validity. (To ovšem neznamená, že se na pozorovatele nemá klást žádné břímě interpretace.)

Jednoduchým aspektem validity měřítek pozorování je jejich předpovědní kapacita (účinnost). Předpovídají spolehlivě nějaká relevantní kritéria? Potíž je — jako obvykle — v těchto kritériích. Jak zdůrazňují Heyns a Lippitt, jsou nezávislá měřítka týchž proměnných vzácná.¹¹ Můžeme říci, že je pozorovací měřítko učitelova chování validní, protože kladně koreluje s posudky nadřizovaných? Snad bychom mohli mít nezávislé měřítko egocentricky orientovaných potřeb, ale bylo by toto měřítko adekvátním kritériem pro pozorování takových potřeb?¹²

⁶R. Heyns, R. Lippitt: Systematic Observational Techniques. In: Handbook of Social Psychology, red. G. Lindzey, sv. I., Cambridge, Mass., Addison-Wesley, 1954, kap. 10, str. 399.

⁷Ryan, c. d., str. 87.

⁸V originálu „self-oriented needs“ — tedy „potřeby orientované na sebe“, respektive na „vlastní osobu“.

⁹N. Fouriezios, M. Hutt, H. Guetzkow: Self-Oriented Needs in Discussion Groups. In: Group Dynamics, red. D. Cartwright, A. Zander. New York, Harper & Row, 1953, str. 354-360.

¹⁰Poslední výrazy této věty v originále znějí „observed behaviors“ (jde tedy o plurál — „pozorovaná chování“ — překládáme raději pomocí výrazu „akty“, abychom se vyhnuli doslovnému, ale neobratnému „mezi pozorovanými chováními“); a „inferred needs“ — abychom se vyhnuli výrazu „vydedukovanými potřebami“, užíváme v podobných kontextech jako synonyma „vysouzený“ a podobně.

¹¹Heyns, Lippitt, c. d., str. 398.

¹²Pro výborné pojednání o problému kritéria a validity ve vztahu k vyučovacím chování viz Ryans, c. d., kap. 2.

Důležitým klíčem ke studiu validity behaviorálních pozorovacích měřítek by mohla být konstruktová validita. Jestliže jsou proměnné, které se měří pozorovacími postupy, zasazeny v teoretickém rámci, pak by tu měly existovat jisté vztahy. Existují však skutečně? Pracujeme-li například s teoretickým rámcem, založeným na potřebách, můžeme z té teorie dedukovat, že učitelky, které silně podléhají potřebám úlevy a vystřídání, budou dosahovat (v Ryansově pozorovacím systému) vysokého skóre na popudlivém a nejistém, a nízkého na chápajícím. Jestliže platí tyto vztahy, pak je to důkazem pro konstruktovou validitu pozorovacího systému.

Spolehlivost behaviorálních pozorovacích měřítek je záležitostí jednodušší, i když nikterak snadnou. Spolehlivost se obvykle definuje jako souhlas mezi pozorovateli. Čtenář si vzpomene, že se to rovná definici objektivit, uvedené výše. I když tedy celkem vyhovuje alespoň pro praktické účely chápat spolehlivost takto, přece by se mělo širší pojetí spolehlivosti, jak bylo rozvinuto výše, aplikovat na pozorování tak, jak se aplikuje na všechna ostatní měřítka.

Prakticky řečeno, lze tedy spolehlivost pozorování stanovit korelováním pozorování dvou nebo více pozorovatelů. Když se určuje spolehlivost přiřazení (aktů) chování ke kategoriím, užívá se často procenta shody mezi posuzovateli. Avšak jako u všech druhů měření jsou tu také jiné způsoby, jak určovat spolehlivost, například spolehlivost opakováním a spolehlivost určená analýzou rozptylu.

1.2 Charakteristiky pozorování¹³

Prvním a nejdůležitějším zřetelem v kterémkoli pozorovacím systému je znát jasně to, co se pozoruje. To se zdá tak zřejmým, že je to až banální. Podívejme se však, že to není tak zřejmé. Předpokládejme, že studujeme vztah mezi nezávislostí a řešením problémů. Stanovíme hypotézu, že čím nezávislejší dítě, tím víc bude schopno řešit problémů, zůstanou-li ostatní okolnosti stejné. Chceme pozorovat nezávislost, nebo přesněji řečeno, nezávislé chování. Avšak co to je nezávislé chování? Je nezávislým chováním to, když dítě neustále pracuje samo? Nebo je nezávislým chováním, když dítě začíná projekty a hry s jinými dětmi? Vezměme ještě mnohem obtížnější problém: co označujeme demokratickým chováním u učitelky? Je demokratickým to, když je učitelka k dětem milá? Je demokratickou, jestliže organizuje děti do skupin? Co vůbec míníme tím, když říkáme „demokratické chování“? Mělo by být zřejmým, že jasné vědomí toho, co se pozoruje, není tak zřejmé.

Je tedy nutno definovat dosti přesně a nedvojznačně to, co se má pozorovat. Měříme-li zvědavost, musíme říci pozorovateli, co to je zvědavé chování. Měří-li se kooperativnost, musíme pozorovateli nějak sdělit, jak se kooperativní chování odlišuje od kteréhokoli jiného druhu chování. To znamená, že musíme poskytnout pozorovateli nějakou formu operační definice proměnné, která se měří; musíme tu proměnnou definovat behaviorálně.

1.2.1 Kategorie

Základním praktickým úkolem pozorovatele je přiřadit akty chování ke kategoriím. Kategorie Ryansova záznamu jsou definovány adjektivními páry. V jiném obsáhlém systému pro pozorování chování ve třídě se užívá osmi širokých behaviorálních kategorií nebo dimenzí.¹⁴ Například dvě z nich jsou diferenciací a klima učitele. Diferenciací v sobě zahrnuje způsoby chování, zaměřené na brání nebo nebrání zřetele k individuálním rozdílům žáků. Klima učitele zahrnuje v sobě učitelovy způsoby chování, které určují nebo ovlivňují citové a sociální ovzduší třídy. Tyto široké kategorie jsou dále rozděleny do podkategorií, které naznačují, kde podél měřené dimenze se má dotyčné chování umístit. Tak se například diferenciací rozděluje do 10 podkategorií od „jednoduchá práce — bez přispění učitele“ až „diferencovaná pracovní schopnost a zájmová základna — za individuálního přispění učitele“.¹⁵ Během pozorovacího období si může pozorovatel povšimnout, že se učitel vůbec nepokouší pomáhat kterýmkoli jedincům a že všichni žáci dělají tutéž práci. Přiřadil by tedy tento soubor behaviorálních aktů do první podkategorie.

Z dřívější stati o rozkladu připomeňme, že (všechny) kategorie musí být vyčerpávající a vzájemně se vylučující (tj. nesmí se překrývat — pozn. překl.). Aby se splnila podmínka vyčerpanosti, musí se nejprve definovat U, totiž univerzum projevů chování, které se mají pozorovat. V některých pozorovacích systémech není těžké to učinit. V jiných je to velmi obtížné. V systému Cornella, Lindvalla a Saupeho bylo U chování všech učitelů a žáků ve třídě. Ale to je příliš velký a nejasný řád. A tak tito badatelé rozčlenili U do podsouborů (skupin): A: diferenciací; S: sociální organizace; C: žákovská iniciativa; D: obsah; E: rozmanitost; F: kompetence; G: klima učitele; H: klima žáka.¹⁶ Pečlivá studie tohoto členění ukáže, že lze většinu projevů chování ve třídě zahrnout pod jednu z těchto rubrik. Cornell, Lindvall a Saupe dále rozčlenili výše uvedených osm podsouborů U na další podsoubory podsouborů. Jedno z těchto členění bylo popsáno výše.

¹³V originále plurál: „Characteristics of observations“.

¹⁴F. Cornell, C. Lindvall, J. Saupe: An Exploratory Measurement of Individualities of Schools and Classrooms. Urbana, Ill., Bureau of Educational Research. University of Illinois, 1952.

¹⁵Ibid., str. 53.

¹⁶Ibid., str. 53, 54.

Mohlo by to znít, jako by se musely všechny možné projevy nebo způsoby chování¹⁷ definovat a pozorovat. Tak tomu není. V mnohých případech se může výzkumník zajímat pouze o proměnnou sociální klima (třídy nebo skupiny). V takovém případě by U zahrnovalo chování, podchycené dvěma podsoubory Cornelia, Lindvalla a Saupeho v klimatu učitelů a klimatu žáků. Tyto zahrnují položky jako „Respektoval žákův názor“, „Snažil se chápat žákovo hledisko“, „Nadměrně opravoval nebo kritizoval“.¹⁸ Jinými slovy, U může být velké, střední nebo malé, podle výzkumného problému a výzkumných cílů. Je možno se pohybovat v rozpětí od velmi široké proměnné jako učitelova efektivnost a zahrnout přitom všechny Ryansovy kategorie až k poměrně úzké proměnné jako agresivnost, která bude třeba vyžadovat jen dvě nebo tři podkategorie, jako fyzické a verbální ubližování jinému dítěti a brání předmětů jinému dítěti.

1.2.2 Jednotky chování

Je stále ještě nevyřešeným problémem, jakých jednotek užívat při měření lidského chování. Zde jsme postaveni před rozpor mezi požadavky spolehlivosti a validity. Teoreticky můžeme dosáhnout vysokého stupně spolehlivosti použitím malých, snadno pozorovaných a snadno zaznamenaných jednotek. Je možno se pokusit definovat chování zcela operačně sestavením velkého počtu behaviorálních aktů a tím pak lze obvykle dosáhnout vysokého stupně přesnosti a spolehlivosti. Ale takovým postupem je možno také redukovat ono chování tak, že se již příliš nepodobá tomu chování, které jsme měli v úmyslu pozorovat. A tak jsme ztratili validitu.

Na druhé straně lze užít širokých „přirozených“ definicí a dosáhnout snad vysokého stupně validity. Mohli bychom instruovat pozorovatele, aby pozorovali spolupráci, a definovat kooperativní chování jako „přijímání podnětů, myšlenek a navazování styků ze strany druhých osob; harmonicky spolupracovat s ostatními k (jistým) cílům“, nebo použít nějaké podobné, dosti široké definice. Jestliže měli pozorovatelé zkušenosti se skupinami a rozumějí-li skupinovým procesům, pak by se dalo očekávat, že při použití této definice dokáží validně odhadnout chování jako kooperativní a nekooperativní. Takováto široká, dokonce i vágní definice umožňuje pozorovateli zachytit, dovedli to, plný rozsah¹⁹ kooperativního chování. Dovoluje však také, aby se do percepce pozorovatelů vetřela značná dvojnásobnost interpretace, snižující tím spolehlivost.

Někteří badatelé, kteří jsou ve svém přístupu orientováni silně operačně, trvají na vysoce specifických definicích u pozorovaných proměnných. Budou třeba uvádět seznam četných specifických aktů chování, které má pozorovatel pozorovat. Žádné jiné by se nepozorovaly a nezaznamenávaly. Podobné extrémní přístupy mohou produkovat vysokou spolehlivost, ale mohou také opomenout část důležitého jádra pozorovaných proměnných. Předpokládejme, že se pro kooperativnost sestaví deset specifických typů chování. Dále předpokládejme, že se universum možných (druhů) chování skládá ze 40 nebo 50 typů. Zřejmě se pak zanedbají důležité aspekty kooperativnosti. I když se to, co se měří, měří spolehlivě, může to být zcela triviální nebo alespoň částečně irelevantní vzhledem k proměnné kooperativnosti.

To je problém molárního — molekulárního v jakémkoli postupu měření v sociálních vědách. Molární přístup si bere větší behaviorální celky jako jednotky pozorování. Jako pozorovací cíle se mohou specifikovat kompletní interakční jednotky. Verbální chování se může rozčlenit na kompletní výměny sdělení mezi dvěma nebo více jedinci, nebo na celé odstavce či věty. Naproti tomu molekulární přístup si bere menší výseky chování jako jednotky pozorování. Může se zaznamenat každá výměna sdělení nebo jen částečná výměna. Jednotkami verbálního chování mohou být slova nebo

¹⁷ „Projevy nebo způsoby chování“ je naše rozvedení výrazu „behaviors“.

¹⁸ Ibid, str. 54.

¹⁹ Doslova „plnou příchutí“ nebo „veškerou vůni“: v originále bylo „full flavor“.

krátká rčení. Molární pozorovatel bude vycházet z obecné široce definované proměnné, jak bylo uvedeno výše, a bude brát v úvahu a zaznamenávat rozmanité (druhy) chování pod jednou rubrikou. Spoléhá se na svou zkušenost a interpretaci významu dějů, které pozoruje. Naproti tomu molekulární pozorovatel se snaží zatlačit svou vlastní zkušenost a interpretaci z pozorovacího obrazu. Zaznamenává to, co vidí, a nic více.

1.2.3 Stupeň pozorovatelova vysuzování

Systémy pozorování se liší v jiné důležité dimenzi, totiž v dávce vysuzování, dedukce, požadované od pozorovatele. Molekulární systémy vyžadují poměrně málo vysuzování. Pozorovatel si prostě všímá toho, že jedinec něco dělá nebo říká. Tak například nějaký systém může od pozorovatele vyžadovat, aby si znamenal každou interakční jednotku, kterou je možno definovat jako jakoukoli verbální výměnu sdělení mezi dvěma jedinci. Dojde-li k výměně sdělení, zaznamená se; nedojde-li k ní, nezaznamená se. Nebo nějakou kategorií může být „Bije jiné dítě“. Pokaždé, když dítě bije jiné dítě, zaznamená se to. V takových systémech se nedělají žádné dedukce, jestliže je ovšem vůbec možno uniknout dedukcím (například „bije“). Zaznamenává se pokud možno co nejvěrněji čisté chování.

Systémy pozorování s tak nízkými stupni pozorovatelské dedukce jsou vzácné. Většina systémů vyžaduje jistého stupně vysuzování. Badatel bude třeba dělat výzkum o chování výchovné komise a rozhodne se třeba, že se na jeho problém hodí analýza s nízkou měrou vysuzování. Užije pak třeba položek pozorování jako: „Navrhuje opatření“, „Přerušuje jiného člena komise“, „Polemizuje o něčem“, „Klade otázku“, „Dává příkaz inspektorovi“, „Dělá návrh“ a podobně. Protože jsou takové položky poměrně nedvojznačné, měla by být spolehlivost tohoto systému pozorování značná.

Systémy, které vyžadují od pozorovatele vyšších stupňů dedukce, jsou častější a ve většině výzkumů pravděpodobně užitečnější, zvláště ve většině pedagogických výzkumů. Vysoce úsudkový systém pozorování dává pozorovateli značované kategorie, které vyžadují větší nebo menší interpretace pozorovaného chování. Výše zmíněný systém pozorování Fouriezose, Hutta a Guetzkowa má dosti vysoké nároky na pozorovatelské vysuzování. Ačkoli se podává a definuje pět oblastí potřeb, musí pozorovatel vždy vysoudit, že nějaké určité chování je v potřebách orientováno egocentricky. Tak například jednou z oněch pěti potřeb je dominování. Definuje se jako pokusy nějakého jedince dávat najevo intelektovou nadřazenost nad jinými jedinci. Autoři říkají: „Dominování v sociálních situacích zahrnuje pokusy ovládat a řídit v sociální situaci. . . Vůdce zde vládne s malým zřetelem pro potřeby skupiny. . .“²⁰ A pak říkají, že se dominování ukazuje tam, kde jedinec odmítá vyslechnout argumenty proti svým vlastním představám a snaží se přinutit skupinu, aby šla za jeho plány. Pro pozorovatele však zdůrazňují nutnost rozlišovat mezi dominováním, založeným na potřebách ego, a dominováním, založeným na situačních požadavcích. Je zřejmé, že tento systém je vysoce deduktivní. Aby se však čtenář nedomníval, že to nutně narušuje spolehlivost, poznamenejme, že se testové koeficienty spolehlivosti rozkládaly od 0,67 do 0,96. K dosažení tak uspokojivé spolehlivosti zdůrazňují autoři nutnost cvičit pozorovatele.

Není možno dělat rovnou generalizace o relativních přednostech systémů s různými stupni vysuzování.²¹ Snad nejlepší radou pro nováčka je zaměřit se na střední stupeň vysuzování. Příliš vágní kategorie s příliš malou specifikací toho, co se má pozorovat, kladou na pozorovatele nadměrné břímě. Různí pozorovatelé mohou příliš snadno dávat témuž chování různé výklady. Naopak příliš specifické kategorie, i když omezují mnohoznačnost a nejistotu, se mohou stávat příliš rigidními

²⁰Fouriezos, Hutt, Guetzkow, c. d., str. 684.

²¹Heyns a Lippitt pojednávají o dimenzi vysuzování v pozorováních v každém z několika dosti úplně popsaných sociálně psychologických systémů pozorování. Heyns, Lippitt, c. d.

(ztrnulými) a nepružnými, dokonce i triviálními. Spíše než cokoli jiného by měl student studovat různé osvědčené systémy a věnovat přitom zvláštní pozornost kategoriím chování a definicím (instrukcím), připojeným k těmto kategoriím jako vodítka pro pozorovatele.

1.2.4 Obecnost nebo aplikovatelnost

Systémy pozorování se značně liší ve své obecnosti nebo stupni aplikovatelnosti na jiné výzkumné situace než na ty, pro které byly původně sestaveny. Některé systémy jsou zcela obecné: jsou sestaveny pro použití u mnohých různých výzkumných problémů. Jedním takovým obecným systémem je dobře známá Balesova analýza skupinové interakce.²² Je to systém nízkého (stupně) vysuzování, v němž veškeré verbální i nonverbální chování, podle předpokladu ve kterékoli skupině, lze kategorizovat do jedné z dvanácti kategorií: „projevuje solidarnost“, „souhlasí“, „žádá o názor“ atd. Oněch dvanáct kategorií je seskupeno do tří větších souborů: sociálně–emotivně kladného, sociálně–emotivně záporného a úkolově neutrálního. Ryansův systém byl zřejmě zamýšlen pro obecné užití v kterékoli třídě. Také systém Cornelia, Lindvalla a Saupeho je obecný.

Některé systémy však byly konstruovány pro použití ve zvláštních výzkumných problémech a mohou nebo nemusí být obecně aplikovatelné na jiné situace. Takovým je zřejmě systém Fourieze, Hutta a Guetzkowa. A přece jej lze aplikovat i na jiné situace než jen na usnášecí konference. Po nějaké adaptaci by se ho dalo pravděpodobně použít v pedagogickém výzkumu — třeba pro zasedání výchovné komise. Nebo by se snad s jeho pomocí daly pozorovat dětské skupiny a kluby. Po vhodné revizi a adaptaci lze pravděpodobně aplikovat většinu systémů, sestavených pro specifické výzkumné problémy, také na jiné výzkumné problémy.

1.2.5 Pořizování vzorků chování

Poslední charakteristika způsobů pozorování, totiž pořizování vzorků, přesně řečeno není charakteristikou. Je způsobem získávání pozorování. Dříve než se ve skutečném výzkumu užije systému pozorování, musí se rozhodnout, kdy a jak se bude systém aplikovat. Jestliže se mají pozorovat způsoby chování učitelů ve třídě, jak se bude pořizovat vzorek těchto způsobů chování? Budou se pozorovat všechny specifikované způsoby chování v jedné vyučovací hodině? To právě dělá Ryans. Nebo se budou specifikované vzorky specifikovaných způsobů chování pořizovat systematicky nebo náhodně? Jinými slovy, musí se přikročit k sestavení a využití nějakého druhu výběrového plánu.

Jsou dva aspekty vzorkování chování: pořizování vzorků událostí a pořizování časových vzorků.²³

Vzorků událostí jsme se vlastně dotkli dříve, když se pojednávalo o jednotkách pozorování. Stručného vysvětlení však potřebuje další význam, spojený s tímto termínem. Pořizování vzorků událostí je výběrem pro pozorování integrálních behaviorálních dějů či událostí dané třídy.²⁴

Příklady integrálních událostí jsou záchvaty zlosti, rvačky a hádky, hry, citové příhody, výměny verbálních sdělení o specifických předmětech hovoru, situace problémových řešení, interakce ve třídě mezi učiteli a žáky atd. Vyšetřující, který sleduje události, musí buďto vědět, kdy k těmto událostem dochází, a být přítomen, když k nim dojde, jako je tomu v případě událostí ve třídě, anebo čekat, až nastanou, jako v případě hádek.

²²R. Bales: *Interaction Process Analysis*. Cambridge, Mass., Addison–Wesley, 1951.

²³Viz H. Wright: *Observational Child Study*. In: *Handbook of Research Methods in Child Development*, red. P. Mussen, New York, Wiley, 1960, kap. 3. Pro úplný přehled časového vzorkování viz R. Arrington: *Time Sampling in Studies of Social Behavior: A Critical Review of Techniques and Results with Research Suggestions*, *Psychological Bulletin*, XL/1943, str. 81–124.

²⁴Wright, c. d., str. 104.

Vzorky událostí mají tři přednosti: za prvé jsou tyto události přirozenými životními situacemi, a tak obsahují sobě vlastní validitu, kterou obvykle nemají časové vzorky. Za druhé, integrální událost obsahuje v sobě kontinuitu chování, kterou nemají rozdrobenější behaviorální akta časových vzorků. Jestliže pozorujeme situaci řešení problému od počátku do konce, sledujeme tu přirozenou a úplnou jednotku individuálního a skupinového chování. Tím dosahujeme ucelené a větší realistické jednotky individuálního nebo sociálního chování. Jak jsme viděli v dřívější kapitole, když se pojednávalo o přirozených experimentech a terénních výzkumech, mají naturalistické situace důraznost a blízkost k psychologické a sociální realitě, což experimentům obvykle chybí. Třetí přednost vzorků událostí spočívá v důležité charakteristice mnohých behaviorálních událostí: jsou totiž někdy řídké a vzácné. Tak se například můžeme zajímat o rozhodování, prováděná na administrativních a legislativních zasedáních. Nebo se můžeme zajímat o poslední krok při řešení problému. Proměnnou mohou být také disciplinární metody učitelů. Takové a mnohé jiné události jsou poměrně řídké. Jako takové se mohou snadno propást časovým vzorkováním; vyžadují proto vzorky událostí.

Požizování časových vzorků je výběrem behaviorálních jednotek pro pozorování v různých bodech v čase. Jednotky pozorování se mohou zvolit systematickými nebo náhodnými způsoby tak, aby byly reprezentativními pro definované univerzum chování. Dobrým příkladem je učitelské chování. Předpokládejme, že se studují vztahy mezi jistými proměnnými jako učitelova pozornost, spravedlivost a iniciativa na jedné straně a žákova iniciativa a kooperativnost na druhé straně. Můžeme vybrat náhodné vzorky učitelů a pak vzít časové vzorky jejich behaviorálních aktů. Tyto časové vzorky mohou být systematické: tři pětiminutová pozorování v určených dobách v průběhu — řekněme — každé z pěti hodin vyučování, přičemž to budou první, třetí a pátá vyučovací hodina jednoho dne a druhá a čtvrtá vyučovací hodina příštího dne. Anebo tyto časové vzorky mohou být náhodné: pět pětiminutových pozorovacích období, vybraných náhodně ze specifikovaného univerza pětiminutových období. Jak je zřejmo, existuje mnoho způsobů, jak se mohou časové vzorky sestavovat a vybírat. Jako obvykle musí být způsob, jakým jsou takové vzorky vybírány, jejich délka a jejich počet, pod zřetelem výzkumného problému.²⁵

Časové vzorky mají důležitou výhodu v tom, že zajišťují badatelům získání reprezentativních vzorků chování. To však platí pouze pro (druhy) chování, které se vyskytují dosti často. Chování, která se vyskytují zřídka, mají vysokou pravděpodobnost, že uniknou ze vzorkovací sítě, pokud se nevyberou vzorky v širokém rozmezí. Tak například tvořivé chování, soucitné anebo nepřátelské chování může být velmi řídké. Přece však je časové vzorkování význačným příspěvkem k vědeckému studiu lidského chování.

Jak vyplývá z výše uvedeného, trpí časové vzorky nedostatkem kontinuity, nedostatkem přiměřeného kontextu a snad i přirozenosti. To platí zejména tehdy, užívá-li se malých jednotek času a chování. A přece není důvodu, proč by se událostní a časové vzorkování nemohlo někdy kombinovat. Kdybychom studovali recitace ve třídě, mohli bychom si vyčlenit náhodný vzorek vyučovacích hodin jednoho učitele v různých dobách a pozorovat všechny recitace během vzorkovaných období, přitom pak dávat pozor, abychom pozorovali každou recitaci v její úplnosti.

²⁵V poutavé studii vůdcovství a moci vlivu skupiny u malých dětí podotýká Merei, že by časové vzorkování ukázalo jen vůdce, jak dávají příkazy, a skupinu, jak poslouchá, kdežto prodloužená pozorování by ukázala vnitřní fungování příkazování a poslouchání. F. Mere: *Group Leadership and Institutionalization*, Human Relations, II (1949), str. 23–39.

1.3 Posuzovací škály²⁶

Dosud jsme hovořili pouze o pozorování aktuálního chování. Pozorovatel naslouchá a dívá se přímo na předměty pod daným zřetelem. Sedí ve třídě a pozoruje interakce mezi učitelem a žákem a mezi žáky navzájem. Nebo může sledovat skupinu dětí a naslouchat jí, jak řeší problém za jednosměrně průhledným sklem. Existuje však ještě jiná třída behaviorálního pozorování, o které je třeba se zmínit. Tento typ pozorování se bude nazývat zapamatovaným chováním nebo postřehnutým chováním. Je vhodné uvažovat o něm pod tématem posuzovacích stupnic.

Při měření zapamatovaného nebo postřehnutého chování obvykle předložíme pozorovateli systém pozorování ve formě nějakého druhu škály a žádáme ho, aby zhodnotil nějaký předmět v jedné nebo více charakteristikách, přičemž ten předmět není přítomen. Za tím účelem musí pozorovatel dělat své odhady a posudky na základě minulých pozorování nebo na základě svých postřehů o tom, jaký je pozorovaný předmět a jak se bude chovat. A právě pomocí posuzovacích stupnic je vhodné měřit jak aktuální chování, tak i postřehnuté či zapamatované chování.

Posuzovací škála je psychologickým měrným nástrojem, který vyžaduje od posuzovatele nebo pozorovatele, aby přiřadil posuzovaný předmět ke kategoriím nebo kontinuím, jimž se přidělila čísla. Z psychologických měrných nástrojů lze posuzovací škály snad nejčastěji všude najít, a to pravděpodobně proto, že je lze zdánlivě snadno konstruovat a — což je důležitější — že se jich dá snadno a rychle použít. Bohužel je zdánlivá snadnost konstrukce klamnou a na snadnost použití se těžce doplácí, a to nedostatkem validity, zaviněným četnými zdroji předpojatých a zkreslujících tendencí, které vnikají do posuzovacích měr. Přesto však mohou být posudky nesmírně cennými, užívá-li se jich promyšleně, dovedně a pečlivě.

1.3.1 Typy posuzovacích stupnic

Existují čtyři nebo pět typů posuzovacích stupnic. O dvou typech se pojednávalo v předcházející kapitole; byly to zaškrťávací seznamy a nástroje nucené volby. Nyní budeme uvažovat pouze o třech typech a jejich charakteristikách. Jsou to: kategoriální posuzovací škála, numerická posuzovací škála a grafická posuzovací škála. Jsou si velmi podobné, liší se vlastně v podrobnostech. Kategoriální posuzovací škála předkládá pozorovateli nebo posuzovateli několik kategorií, z nichž si vybere tu, která nejlépe charakterizuje chování nebo vlastnost posuzovaného předmětu. Předpokládejme, že se posuzuje učitelčino chování ve třídě. Jednou z posuzovaných vlastností bude třeba pohotovost. Kategoriální položkou by mohlo být:

- Jak je pohotová? (Jedno z toho zatrhnout)

– velmi pohotová

²⁶Výborné pojednání o posuzovacích stupnicích viz J. Guilford: *Psychometric Methods*. 2. vyd. New York, McGraw-Hill, 1954, kap. 11. Eysenck dovedně shrnul nejdůležitější podstatné práce, zvláště faktorově-analytické a příbuzné práce o posuzování: H. Eysenck: *The Structure of Human Personality*. New York, Wiley, 1953, kap. 11. Ačkoli byly již dříve v této knize zmínky o posuzovacích škálách, nepojednávalo se o nich systematicky. Při čtení toho, co následuje, by měl mít studující na paměti, že posuzovací stupnice jsou vlastně objektivními škálami. Jako takové by se byly mohly zařadit do 27. kapitoly. Pojednání o nich se však vyhradilo této kapitole, protože se pozornost 27. kapitoly soustřeďovala hlavně na měřítka, na která odpovídala nebo reagovala zkoumaná osoba, podrobená měření. Naproti tomu jsou posuzovací škály měřítka jedinců a jejich reakci, charakteristik a (projevů) chování ze strany pozorovatelů. Kontrast je tedy mezi tím, jak zkoumaná osoba vidí samu sebe, a tím, jak ji vidí ostatní (pozn. překl. doslova: Kontrast je tedy mezi zkoumanou osobou, jak vidí samu sebe, a zkoumanou osobou, jak ji vidí druzí). Je také důležité si povšimnout, že se posuzovacích stupnic užívá i k měření psychologických předmětů, produktů a podnětů jako psaní náčrtů (konceptů), essayů, protokolů interviewů a projekčních testových materiálů.

- pohotová
- nepohotová
- naprosto nepohotová

Poněkud jinou kategoriální položkou by mohlo být:

- Jak je důmyslná? (Jedno z toho zatrhnout)

- nesmírně důmyslná
- velmi důmyslná
- důmyslná
- nedůmyslná
- velmi nedůmyslná
- naprosto nedůmyslná

Ještě jiná forma užívá zhuštěných popisů. Taková položka by mohla vypadat asi takto:

- Je vynalézavá? (Jedno z toho zatrhnout)

- Vždy vynalézavá; nikdy jí nechybějí nápady
- Má dobré nápady
- Někdy postrádá nápady
- Velmi nevynalézavá; zřídka má nápady

Numerické posuzovací škály lze snad nejsnáze konstruovat a používat jich. Poskytují také čísla, kterých lze přímo užít ve statistickém rozboru. Mimoto se mohou blížit intervalovému měření, protože v mysli pozorovatele mohou čísla představovat stejné intervaly.²⁷ Kteroukoli z předešlých kategoriálních stupnic je možno rychle a snadno přeměnit na numerickou posuzovací škálu prostým připojením čísel před každou z kategorií. K výše uvedené položce pohotovosti lze připojit čísla 3, 2, 1, 0 nebo 4, 3, 2, 1. Možnosti volby v položce dirmyslnosti by mohly mít před sebou čísla 5, 4, 3, 2, 1, 0. Vhodnou metodou numerického posuzování je užít u každé položky téhož numerického systému, řekněme 4, 3, 2, 1, 0. To je ovšem systém užívaný u postojových stupnic sumačního posuzování. U posuzovacích stupnic je však pravděpodobně lépe dávat jak verbální popisy, tak i čísla.

V grafických posuzovacích stupnicích se kombinují čáry nebo proužky s popisnými výroky. Takové stupnice mají mnoho obměn: vertikální dílcové linie, spojitě linie, nevyznačované linie, linie, rozčleněné na vyznačené stejné intervaly (jak znázorněno výše) a jiné. Tyto jsou pravděpodobně nejlepší z obvyklých forem posuzovacích stupnic. V mysli pozorovatele vštěpují a ustalují kontinuum. Vnukají představu stejných intervalů. Jsou jasné a snadné k porozumění a užívání. Guilford je trochu přeceňuje, když říká: „Předností grafických posuzovacích stupnic je mnoho, jejich vad málo“, ale myslel to dobře.²⁸

²⁷Guilford, c. d., str. 264.

²⁸Ibid., str. 268.

1.3.2 Slabiny posuzovacích stupnic

Ratingy mají dvě vážné slabiny; jedna z nich je zevnější, druhá vnitřní. Zevnější vadou je to, že je lze zdánlivě tak snadno konstruovat a používat jich, takže se jich užívá bez rozlišování, často bez jakýchkoli vědomostí o jejich vnitřních vadách. Nebudeme se zde zdržovat zmínkami o mnohých chybách, které se mohou vloupat do neobratné konstrukce a užívání posuzovacích stupnic. Raději varujeme čtenáře před tím, aby se jich chápal pro všechny a kterékoli potřeby měření. Měli bychom si nejprve položit otázku: Je nějaký lepší způsob, jak měřit mé proměnné? Jestliže ano, uijme ho. Jestliže nikoli, pak studujme vlastnosti dobrých posuzovacích stupnic, pracujme se svědomitou péčí a podrobme ratingové výsledky empirické zkoušky a přiměřenému statistickému rozboru.²⁹

Vnitřní vadou posuzovacích stupnic je jejich náchylnost ke konstantní či předpojaté chybě. To pro nás ovšem není nové. Setkali jsme se s tímto problémem v předešlé kapitole, když jsme uvažovali o reaktivním sklonu. U posuzovacích škál to však zvláště ohrožuje validitu. Konstantní chyba nabývá několika forem, z nichž nejpronikavější je pověstný haló efekt. Je to tendence posuzovat předmět v konstantním směru celkového dojmu předmětu. Každodenními případy tohoto efektu jsou: domnívat se, že je někdo inteligentní, protože s námi souhlasí; domnívat se, že je nějaký člověk charakterní, protože, ho máme rádi; vzdávat velkou chválu republikánským prezidentům a odsuzovat demokratické prezidenty.

Efekt se často projevuje v měření, zvláště u posuzovacích škál. Profesor hodnotí kvalitu eseje výše, než by si zasloužil, protože má rád zkoušeného. Nebo může posuzovat druhou, třetí a čtvrtou otázku výše (nebo níže), než jak by si zasloužily, protože první otázka byla dobře zodpovězena (nebo chabě zodpovězena). Dalším případem efektu je učitelovo hodnocení dětských výkonů, pokud je ovlivňováno učenlivostí nebo nedostatkem poddajnosti dětí. V posuzování jedinců na ratingových stupnicích má posuzování jedné vlastnosti tendenci ovlivnit ratingy jiných vlastností.

Je nesmírně těžko vyhnout se haló efektu. Zdá se, že je haló efekt zvláště silný u rysů, které nejsou jasně definovány, které nejsou snadno pozorovatelné a které jsou morálně důležité.³⁰

Dvěma důležitými zdroji konstantní chyby jsou chyba přísnosti a chyba shovívavosti. Chyba přísnosti je celkovou tendencí posuzovat všechny jedince ve všech vlastnostech příliš nížko. To je ten strohý známkovatel, člověk, který říká: „V mých hodinách se nedávají jedničky!“ Chyba shovívavosti je opačnou celkovou tendencí, totiž posuzovat příliš vysoko. To je ten dobrák, který má každého rád, a tato vřelost se obráží v jeho posudcích.

Zneklidňujícím zdrojem nevalidity v posudcích je chyba centrální tendence, celková tendence vyhýbat se všem extrémním soudům a posuzovat stále jen středem posuzovací stupnice. Projevuje se zejména tehdy, když posuzovatelé nejsou obeznámeni s předměty, které se posuzují.

Jsou ještě jiné, méně důležité typy chyb, které se tu již nebudou probírat. Důležitější bude, jak se vyrovnat s těmi, které byly vyjmenovány výše. Je to složitá věc, o které zde nelze pojednávat. Čtenář se odkazuje na Guilfordovu kapitolu v Psychometrických metodách, kde se podrobně pojednává o mnohých opatřeních, jak čelit chybám.³¹

²⁹ Guilfordova rada je neocenitelná: *ibid.*, str. 264–268 a 293–296.

³⁰ *Ibid.*, str. 279.

³¹ Se systematickými chybami je možno se do jisté míry vypořádat statistickými prostředky. Guilford vypracoval důmyslnou metodu, užívaje analýzy rozptylu. Jeho základní myšlenkou je, že se z celkového rozptylu posudků extrahují rozptyly poplatné zkoumaným osobám, posuzovatelům a vlastnostem. Posudky se pak korigují. Když se posuzují jedinci pouze v jedné vlastnosti, je snadnější metodou dvojná analýza rozptylu (korelovaných skupin). Také spolehlivost lze snadno vypočítat. Použití analýzy rozptylu k určení spolehlivosti, jak jsme se dověděli výše, bylo příspěvkem Hoytovým. Ebel aplikoval analýzu rozptylu na spolehlivost posudků. Viz Guilford, c. d., str. 280–288, 383, 395–397; R. Ebel: Estimation on the Reliability of Ratings, *Psychometrika*, XVI (1951), str. 407–424.

V behaviorálním výzkumu se může a mělo by se užívat posuzovacích stupnic. Jejich neoprávněné, nouzové a neinformované užívání bylo právem odsouzeno. To však by nemělo znamenat všeobecné odsouzení. Posuzovací stupnice mají přednosti, které je činí cennými nástroji vědeckého výzkumu. Ty byly shrnuty Guilfordem: vyžadují méně času než jiné metody; jsou celkem zajímavé a pro pozorovatele snadné k použití; mají velmi široký rozsah aplikace; lze jich užívat při velkém počtu vlastností.³² Mohlo by se dodat, že je možno jich použít jako doplňků k jiným metodám. To znamená, že jich lze užít jako nástrojů na pomoc behaviorálním pozorováním a také ve spojení s jinými objektivními nástroji, s řízenými rozhovory (interviewy), a dokonce i s projektivními měřítky.

³²Guilford, c. d., str. 297.

1.4 Příklady systémů pozorování

O třech nebo čtyřech systémech pozorování jsme se zmínili již dříve. Níže je shrnuto několik jiných behaviorálních systémů pozorování, aby se studujícímu pomohlo získat cit pro rozmanitost systémů, které jsou možné, a pro způsoby, jakými se takové systémy konstruují a jak se jich používá. Mimoto může studující získat další porozumění toho, kdy je behaviorální pozorování na místě, nebo jinými slovy, kdy výzkumný problém vyžaduje přímého měření chování. Shrnuty jsou také systémy a posuzovací škály zapamatovaného chování.

1.4.1 Měření interakce mezi matkou a dítětem — výzkum Merrillové³³

Jsouc vedena snahou měřit proměnné kontakt mezi matkami a dětmi, specifická ovládnutí dětského chování a facilitace a inhibice dětského chování, sestavila Merrillová zajímavý a užitečný systém pozorování. Matka i dítě byly pozorovány v experimentální místnosti, vybavené dětskými hračkami a jednosměrně průhlednou zástěnou. Matkám se řeklo, aby si představily, že jsou doma se svými dětmi na půl hodiny bez zaměstnání. Netušily, že jsou ony samy pozorovány, protože se jim řeklo, že se výzkum týká chování dítěte ve hře.

Chování matky se zaznamenávalo každých pět vteřin, přičemž se užívalo dosti složitýho znaménkového systému, založeného na kategoriích, odvozených z výše zmíněných proměnných. Například S znamená, že M (matka) strukturuje, tj. užívá nepřímých prostředků, aby podnítila nebo ovlivnila dítě; P znamená, že M poučuje nebo podává informaci, aby zvýšila vědomosti dítěte; I znamená, že matka interferuje, zasahuje do aktivity dítěte, aby ji zastavila. Každá kategorie je doprovázena jasnými definicemi a příklady. Kategoriální systém je obsáhlý: zahrnuje všechny možné případy chování, které by se mohly vyskytnout v sezení při hře. Průměrná spolehlivost pro pět sezení byla 0,88.

Merrillová provedla čistý, účinný a přitom prostý experiment, užívajíc systému pozorování. Rozdělila matky do experimentálních a kontrolních skupin a řekla ve druhém půlhodinovém sezení matkám experimentální skupiny, že se hravý potenciál jejich dětí dosud nerealizoval, čímž navodila motivaci podnítit dítě, aby se snažilo „lépe“. Matkám v kontrolní skupině se řeklo, že se bude postupovat stejně jako dříve. Kdyby byl pozorovací postup dosti citlivý, měl by ukázat výsledek navozené motivace. A to ukázal.

1.4.2 Posuzování učitelů studenty: dvě studie

Posuzování učitelů studenty náleží do kategorie „pozorování zapamatovaného chování“, o kterém se pojednávalo již dříve. K tomuto účelu se konstruovalo mnoho nástrojů. Do nedávna však učitelé, kteří užívali takových stupnic, měli malou představu o tom, co právě měří. Coffman se pokusil přijít na to, co jedna taková škála měří.³⁴ Užil pětibodové grafické posuzovací škály, sestávající z 18 rysů (a druhů) chování. Každá podstupnice používala pěti popisných výroků. Dva z oněch 18 rysů a průvodních popisů jsou:

Příprava na třídní schůzky. Třídní schůzky velmi pečlivě plánovány. Obvykle dobře připraveny. Příprava často nepřiměřená. Malá příprava. Žádná příprava.

³³B. Merrill: A Measurement of Mother — Child Interaction, *Journal of Abnormal and Social Psychology*, XLI (1946), str. 37–49.

³⁴W. Coffman: Determining Students' Concepts of Effective Teaching from Their Ratings of Instructors, *Journal of Educational Psychology*, XLV (1954), str. 277–286.

Organizace kursu. Každá lekce dobře organizována. Většina lekcí dobře organizována. Trochu organizace, ale ne vždy jasné. Velmi málo organizace. Žádná organizace.³⁵

Faktorovou analýzou posudků o 55 učitelích našel Coffman čtyři faktory, které nazval vcítění, organizace, přesný — upravený — normální a verbální obratnost. Podobná, ale složitější studie učitelské výkonnosti užila popisného zaškrtávacího seznamu, aby vypočítala škálové hodnoty a preferenční a diskriminační indexy u 150 popisných výroků.³⁶ Konstruovalo se deset souborů po čtyřech výrociích, přičemž se užívalo výsledků předešlých rozborů. Posuzovatelé pak seřadili do pořadí výroky v každém souboru. To znamená, že student, když posuzoval učitele, vybíral v každé čtveřici tu položku, která se na daného učitele nejvíce hodila, dal před ni „1“. Pak měl zvolit další nejvhodnější položku a dát před ni „2“. Před nejméně aplikovatelnou položku se ovšem umístila „4“. Zde je jedna ze čtveřice:

Soubor *i*

- Má vždy učební materiál připraven. (3)
- Dobře ovládal předmět. (1)
- Povzbuzoval studenty, aby přišli na odpovědi. (4)
- Pravidla a stanovy spravedlivé.³⁷ (2)

(Čísla v závorkách se vztahují na čtyři faktory, zasazené do každé čtveřice: 1. znalost a organizace učební látky; 2. přiměřenost vztahů se studenty ve třídě; 3. přiměřenost plánů a postupů ve třídě; 4. nadšení při práci se studenty.)

1.4.3 Měření pedagogického vedení

Nástrojem značné potencionální užitečnosti ve studiu pedagogické administrace je tzv. dotazník popisu a chování vedoucího (LBDQ) přizpůsobený Halpinem pro pedagogické využití u školních ředitelů.³⁸ Dotazník LBDQ je nástrojem Likertova typu o 80 behaviorálních položkách, na které odpovídají pozorovatelé v pětibodové škále. Pozorovatel se instruuje, aby popsal ředitelovo chování tak, jak je zná, tím že odpovídá na každou z 80 položek: „Jedná vždy (často, občas, zřídka, nikdy) tímto způsobem“. Hlavními dimenzemi tohoto nástroje jsou zasvěcovací struktura a ohleduplnost. Dvě položky první kategorie jsou:

Kritizuje chabou práci.

Dává členům (učitelského) sboru na vědomí, co se od nich očekává.

Dvě z položek ohleduplnosti jsou:

Najde si čas, aby vyslechl členy sboru.

V důležitých věcech si zajistí souhlas sboru dříve, než začne jednat.

³⁵Tamtéž, str. 278. Toto byly dvě z nejlepších položek, jak se ukázalo faktorovou analýzou.

³⁶D. Cosgrove: Diagnostic Rating of Teacher Performance, Journal of Educational Psychology, L (1959), str. 200–204.

³⁷Ibid., str. 202.

³⁸A. Halpin: The Leadership Behavior of School Superintendents. Chicago, Midwest Administration Center, University of Chicago, 1956. Dotazník LBDQ je uveden na str. 91–95. Viz také A. Halpin: The Leadership Behavior and Combat Performance of Airplane Combat Performance of Airplane Commanders, Journal of Abnormal and Social Psychology, XLIX (1954), str. 19–22.

Jedinečným rysem Halpinova výzkumu je dvojí využití dotazníku LBDQ. Nejprve se ho použilo k tomu, aby se poznalo skutečné chování ředitelů: jak se ředitelé opravdu chovají; to je LBDQ reálný. Za druhé ho bylo užito, aby se poznalo, jak by se ředitelé měli chovat; to je LBDQ ideální. Každá skupina — ředitel, členové sboru a členové pedagogické komise — odpovídala na tento nástroj oběma způsoby. Tak se studovaly různé postřehy skutečných a ideálních ředitelských rolí.

1.4.4 Hodnocení rodičovského chování

Jedním z nejnáročnějších pokusů měřit rodičovské chování v přirozených podmínkách je výzkum Baldwina, Kalhorna a Breesea.³⁹ Pozorovatelé navštěvují rodiny a mezi jiným posuzují chování rodičů na posuzovací stupnici pozorování o 30 položkách. Položky se skládají z dvoupólových adjektivních párů jako kárající — chválící, odmítající — oddaný, bezstarostný — úzkostný, diktátorský — demokratický a tak dále. Chování matky se pozoruje v rodině a posuzuje se na 30 položkách. Každá z položek má pečlivé pozorovatelské návody k tomu, co pozorovat a posuzovat. Přes tyto návody se od pozorovatele vyžaduje dosti vysuzování. Badatelé shledali, že se těchto 30 položek seskupuje do menšího počtu faktorů, z nichž nejdůležitějšími jsou vřelost, demokratičnost a shovívavost. Tak například faktor vřelosti se skládá z těchto položek: podřizování dítěte — dítě středem pozornosti, kárání — chválení, odmítání — přijímání, nepřátelský — laskavý, izolace — úzký kontakt.

1.4.5 Časové vzorky dětského chování

V důležité studii osobnostní struktury čtyřletých a pětiletých dětí, vysouzené ze skutečného chování dětí, využili Cattell a Peterson velmi vhodně časových vzorků behaviorálních aktů.⁴⁰ Pozorovatelé užívali seznamu o 76 položkách, kterými zrakem „přelétávali“ každých pět minut, aby si tak připomínali možné behaviorální akty, které měli pozorovat a zaznamenávat. Pozorovatelé se mísili s jinými dospělými ve velké zahradě a na hřišti a nenápadně konali svá pozorování. Soustředili se plně na jedno dítě po dvě půlhodinová období hry (plus pět minut na každém začátku a konci doby hry).

Tyto položky jsou velmi zajímavé, protože obsahují také položky z jiných podobných výzkumů, a to 40 verbálně definovaných rysů, o kterých bylo předtím zjištěno, že měří osobnosti dětí a jiné položky, které mohou mít k tomu vztah (relevanci). Některé z nich jsou:

Žádá dospělého o pomoc.

Hraje si samo.

Projevuje vážný nebo úzkostný výraz.

Projevuje zvědavost.

Pomáhá jinému dítěti.

Stojí za povšimnutí, že většina z těchto 76 položek je typu nízkého (stupně) vysuzování, ačkoli tu také jsou položky vyžadující dedukování („Projevuje podezření“, „Dává najevo strach z okolí“).

O propracované analýze dat (faktorové analýze) zde nelze pojednávat. Postačí metodu komentovat. Položkový vzorek vyčerpával velmi široký rozsah možných (druhů) chování a ve svém

³⁹A. Baldwin, J. Kalthorn, F. Breese: Patterns of Parent Behavior, Psychological Monographs, LVIII (1945), č. 3; The Appraisal of Parent Behavior, Psychological Monographs, LXIII (1949), č. 4.

⁴⁰R. Cattell, D. Peterson: Personality Structure in 4–5 Year Olds, by Factoring Observed; Time-Sampled Behavior, Estratto della Rassegna di Psicologia Generale e Clinica, 111 (1958), str. 3–21.

charakteru byl zřetelně behaviorální. Položky tvoří výborný fond, z něhož mohou pedagogičtí badatelé extrahovat položky pro výzkum na specifických výzkumných problémech. Mimoto podává Cattellova a Petersonova analýza alespoň pokusné identifikace toho, co tyto položky měří.

1.4.6 Pointově–časové vzorkování činností ve třídě

Při studiu vztahu mezi studentským chováním a činnostmi ve třídě u různých jedinců rozvinul Kowatrakul slibnou metodu pozorování chování ve třídě, které se má použít ve spojení s tím, co nazývá pointově–časové vzorky.⁴¹ Užilo se šesti kategorií chování: zabraný do trvající práce (úzkolově orientován), orientován na sociální práci (sociální poznámky, které jsou pracovně orientovány), sociálně přátelský (jakékoli sociální poznámky), momentální pasivita, zabraný do práce v jiné akademické oblasti, a zabraný do práce v neakademické oblasti. Pozorovaly se také tři druhy činnosti ve třídě: samostatná sedavá práce, sledování a naslouchání a diskuse. Těchto dvou systémů se užilo ve spojení jednoho s druhým.

Nevšedním rysem Kowatrakulova systému bylo pointově–časové vzorkování. Pozorovatel sleduje zkoumaného jedince tak dlouho, až zaznamená jedno chování. Pak ihned přejde na další zkoumanou osobu. Chování dalšího jedince musí být nezávislé na chování prvního (autor však bohužel neříká, jak se toho má dosáhnout). Tohoto postupu se užívá, dokud se nezaznamená jedna „pointa chování“ každé zkoumané osoby ve skupině. Pak začne nové kolo pozorování. Kowatrakul sděluje, že v každé třídě trvá každé kolo od tří do pěti minut.

Hlášené (koeficienty) spolehlivosti jsou nesmírně vysoké — většina z nich větší než 0,90, některé dokonce až 0,98 (byly to spolehlivosti liché — sudé). Ve zvláštní studii 1689 časových vzorků, konané dvěma pozorovateli, byla 94 procentní shoda. Zřejmě je to vysoce spolehlivá metoda.

Ti, kdo studují pedagogický výzkum, by měli věnovat zvláštní pozornost tomuto systému, protože se dobře hodí, jak se zdá, k pozorováním naznačených proměnných ve třídě. Právě stránka pointově–časového vzorkování si rozhodně zaslouží dalšího výzkumu a užití. Málo systémů dává tak vysoké (výsledky) spolehlivosti. Dále se v tomto případě validita nezdá velkým problémem vzhledem k téměř samozřejmé povaze kategorií a činností, které jsou pod nimi zahrnuty.

⁴¹S. Kowatrakul: Some Behaviors of Elementary School Children Related to Classroom Activities and Subject Areas, *Journal of Educational Psychology*, L (1959), str. 121–128.

1.5 Zhodnocení behaviorálního pozorování

Není vůbec žádných pochyb o tom, že objektivní pozorování lidského chování pokročilo nad prvopočáteční stadium. Jako jiné metodologické pokroky měření, učiněné v posledních deseti až dvaceti letech, byly také tyto pokroky pozoruhodné. V pozorování a hodnocení aktuálního i zapa-matovaného chování se již pociťuje vzrůst psychometrické a statistické nadvlády a kultivovanosti. Společenskovědní výzkum může a měl by těžit z těchto pokroků. Tak například mnoho pedagogic-kých výzkumných problémů zřetelně vyžaduje pozorování: děti ve třídách, setkávající se s učiteli a mezi sebou ve vzájemných interakcích; ředitelé, učitelé, diskutující o školních problémech na schůzích sboru; výchovné komise, které se při plánování programů propracovávají k usnesením. Jak jsme viděli, lze (způsobů) pozorování chování užít v situacích přirozených i experimentálně navozených. Je tu jeden metodologický přístup, který je v podstatě stejný u obou druhů situací.

Abychom shrnuli: badatelé by měli věnovat více pozornosti možnosti měření některých svých proměnných přímým pozorováním chování. Taková pozorování přicházejí v úvahu, když je nevhodné nebo nemožné použít jiných metod. Dokonce se jich musí použít, když jsou proměnné výzkumných studií ve své povaze interaktivní a interpersonální. Stručně řečeno, musíme užít pozorování, když jsou výzkumné proměnné zřetelně behaviorální. Jsou-li nezávislými nebo závislými proměnnými chování ředitelů, učitelů, žáků nebo rodičů, pak by zřejmě bylo vhodnějším pozorovat a měřit tyto druhy chování přímo, než jen se někoho na ně dotazovat, i když takové dotazování může být důležité. Nakonec je nutno mít vždy na paměti, že data a jejich rozbor nejsou o nic lepší než původní (způsoby) pozorování, odkud přicházejí.

1.6 Studijní náměty

Studující by se měl blíže seznámit se dvěma nebo třemi systémy pozorování chování. Zdá se, že ze systémů, zmíněných v této kapitole, budou nejužitečnější monografie Ryansova a Cornelia, Lindvalla a Saupeho. Tyto dva systémy se však omezují na pozorování ve třídě. Mnoho studujících bude mít užitek z širšího výběru a dalších studií. Zde je několik dalších dobrých systémů pozorování chování:

1. H. Anderson: Domination and Socially Integrative Behavior. In: R. Barker, J. Kounin, H. Wright, red.: *Child Behavior and Development*, New York, McGraw–Hill, 1943, kap. XXVII. Tato souhrnná kapitola Andersonovy dobře známé pozorovací práce obsahuje dosti složitý a dobře propracovaný systém pro studování účinků dvou typů chování učitelů na chování žáků.
2. R. Barker, T. Dembo, K. Lewin: Frustration and Regression. *Ibid.*, kap. XXVI. Jiná dobře známá a kompetentní studie, v níž se pozorovala konstruktivnost chování dětí za podmínek frustrace. Pojednává se tu o zajímavých pokusech o validizaci korelováním konstruktivnosti s mentálním a chronologickým věkem.
3. R. Bales: Some Uniformities of Behavior in Small Social Systems, In: G. Swanson, T. Newcomb, E. Hhrtlely, red.: *Readings in Social Psychology*. Opr. vyd., New York, Holt, Rinehart and Winston, Inc., 1952, str. 146–159. Zde se podává dobrý stručný popis Balesova systému.
4. N. Gage, G. Leavitt, G. Stove: Teachers' Understanding of Their Pupils and Pupils' Ratings of Their Teachers, *Psychological Afonographs*. LXIX (1955), celé č. 406. Toto je komplexní studie, užívající ratingů. Viz zvláště 12 položkovou škálu „Náš učitel“ (Our Teacher) (str. 14) k měření postřehů žáků o učitelích. Autoři také užíli formy párovaných srovnání téže stupnice.
5. W. Olson: The Measurement of Nervous Habits. In: W. Dennis: *Readings in Child Psychology*, Englewood Cliffs, N. J., Prentice–Hall, 1951, str. 439–445. Je to výňatek z Olsonovy monografie o tomto předmětu, jedna z průkopnických studií, užívající časového vzorkování.
6. M. Parten, S. Newhall: Social Behavior of Preschool Children. In: Barker, Kounin, Wright, c. d., kap. XXIX. Je to dobrý příklad behaviorálních pozorování a časového vzorkování, kde se užívá dětí jako zkoumaných osob. Pozorovanými proměnnými byly sociální participace a vůdcovství. Výsledky se validizovaly oproti ratingům učitelů.
7. H. D. Schmidt: Posudzovanie ľudského správania rating–škálami. Bratislava, SPN, 1971.
8. Zd. Šafář: Škálování, v.: *Vybrané techniky sociologického výzkutu*. Praha, Svoboda, 1969, str. 347–453.