

# Jak postupovat, když jsou naše data nereprezentativní?

## Několik zkušeností a příkladů

### ■ Hynek Jeřábek

- Karlova Univerzita, FSV, Institut sociologických studií
- [hynek.jerabek@gmail.com](mailto:hynek.jerabek@gmail.com)

Konference České sociologické společnosti 1.-3. 2. 2017, Praha

Panel: B7: **Společná metodologická východiska sociologického výzkumu a komerčních výzkumných šetření.**

*moderují: Ing. Hana Huntová (SIMAR), Ing. Petra Průšová (Kantar)*

1

## Východiska tohoto zamyšlení

- Jak vypadá **většina** výzkumů **veřejného mínění**?  
Co zajímá politiky a novináře?  
Jaké jsou **výsledky za celek**?
- Podíl odpovídajících se blíží **50-60%**  
a již **řadu let klesá.**
- Co budeme dělat, až nám **více než polovina dotázaných** odmítne spolupráci?

2

## **Motivy: proč se tím zabývat?**

- Co si počít s **nereprezentativními daty** získávanými **on-line** výzkumy ?
- A co když podíly respondentů za dílčí populace **neodpovídají zastoupení těchto částí v celku?** **Vyřešíme vždy problém vážením?**
- Ve všech těchto případech můžeme poukázat na to, že „**manifestní data** o souboru dotázaných **jako o celku**“ **nejsou spolehlivou oporou** pro naše vědecké výsledky.

3

## **Jak situaci řešit ?**

- Na **dvou příkladech** se pokusím ukázat, že **komparativní metoda** využívající „**modelu elaborace**“ je adekvátní odpovědí na některé z uvedených výzkumných situací.

4

# Jak navrhuji postupovat

- Prozatím jsme na začátku
- Existují dva ověřené způsoby, jak vypovídat o datech, která nejsou reprezentativní.
- **Vysvětlení rozdílů mezi podsoubory**  
(**model komparace**)
- **Zobecnění podobných výsledků** za výrazně rozdílné podsoubory (**model elaborace** – *varianta opakování*)

5

## Dva příklady

### Komparace podsouborů:

#### **RAWAG studie Paula Lazarsfelda 1931/32**

- výsledkem bylo srovnání podsouborů posluchačů;  
podsoubory se významně lišily ve svých preferencích

### Opakování – jeden z modelů elaborace:

#### **On-line komunikační výzkum k 11.září 2001**

- v komunikačním chování velmi rozdílných populací se opakovaly (shodovaly) vzorce pravidelností

6

## RAWAG – studie

- **RAWAG – studie:** první velký empirický výzkum posluchačských populací rozhlasu – provedený v zimě 1931/1932 v Rakousku
- Tato unikátní studie provedená *Paulem Lazarsfeldem* ve Vídni byla poprvé publikována až v roce 1996, tedy 20 let po jeho smrti.

## RAWAG – studie

- 400 000 koncesionářů rozhlasu; dotazníky vyplnilo 38 000 rodin = **110 000 respondentů**
- zjišťováno: **věk, pohlaví, povolání, region**
- 50 druhů pořadů rozhlasu; pro každý typ pořadu - zájem poslouchat = stejně; > **více**; < **méně**
- Pro každý typ pořadu jednoduchý index
- **Index = „% více“ – „% méně“**
- **Příklad:  $I_{opera} = „20\%” - 43\% = -23\%$**   
**chtějí poslouchat méně oper**

Spočteno **7000 indexů** pro různé skupiny a různé pořady

# RAWAG – studie

Zájem o **vážnou** a **lehkou** hudbu podle sociálních skupin posluchačů - mužů RAWAG - studie [Lazarsfeld 1932a: 16]

	HUDBA	
	vážná	lehká
DĚLNÍCI	- 47	+ 59
ŽIVNOSTNÍCI	- 39	+ 53
ÚŘEDNÍCI	- 28	+ 46
INTELEKTUÁLOVÉ	- 17	+ 22
ZEMĚDĚLCI (venkované)	- 39	+ 37

vzorec pravidelnosti: s růstem úrovně rodiny **klesá** zájem o vážnou hudbu; s poklesem úrovně **roste** zájem o lehkou hudbu

## Metodologie studie RAWAG

- Otázky a výpovědi RAWAG studie se týkaly **komparace posluchačských podsouborů**
- Výpověď je tvořena **vzorcem pravidelností**
- Teprve **posloupnost čísel** je výpovědí, **ne jednotlivý číselný údaj**
- Index – 23 říká příliš málo; jen to, že chtějí méně oper v rozhlase
- Posloupnost **-47 -39 -28 -17** pro různá povolání ukazuje **vztah** „oblíby vážné hudby“ a „povolání“
- Výsledky RAWAG studie **vypovídají o vztazích**

# Téma a organizace výzkumu „September 11<sup>th</sup>“

**téma:** okamžité reakce a komunikační chování  
v hodinách následujících po této mimořádné události

- **on-line výzkum** česky, německy, anglicky
- Czech: [www.nms.cz/usa/cz](http://www.nms.cz/usa/cz)
- English: [www.nms.cz/usa/en](http://www.nms.cz/usa/en)
- German: [www.nms.cz/usa/de](http://www.nms.cz/usa/de)
- **sběr dat:** 13. září – 9. října 2001

11

## Složení respondentů - **nereprezentativní**

- **2578 respondentů:** 2090 Čechů, 143 Němců, 88 Slováků, 104 dalších Evropanů, 81 Američanů, Kanadčanů, Britů a Australanů, 72 ze zbytku světa
- 41% žen **59% mužů** (poměr M a Ž všude stejný)
- **mladí** lidé - průměrný věk 31 let
- vysoce **vzdělaní** – 12 a více let školní docházky
- 95% respondentů komunikuje přes **počítač**
- 44% z měst přes 1 mil. obyvatel (**Praha**)

## *Předmět výzkumu & výsledky*

- Co se stalo 11. září?
- Jaké emoce lidé projevovali?
- Jak lidé komunikovali v hodinách následujících bezprostředně po událostech 11.září?

## *Použitá výzkumná metodologie* „Universal Statements“

- Odhalit univerzální názory, postoje a modely komunikačního chování opakované v různých kulturních, civilizačních a jazykových prostředích

## *Findings – prokázaná zjištění*

- Platnost kteréhokoliv tvrzení byla podmíněna jeho opakováním v navzájem rozdílných prostředích.
- Označit něco za **prokázané zjištění** (finding) je možno jen v případě, že se opakovaně prokáže **stejné zjištění v několika podpopulacích**, které se mezi sebou **významně liší**.

## *Co se stalo 11. září? Otevřená otázka*

- Na otevřenou otázku: „**Jak byste jedním slovem označili, co se stalo 11. září?**“ byla nejčastější **emocionální** odpověď „**tragédie**“ „**neštěstí**“ apod. A to i v podsouborech za ČR, EU, USA.



## Uzavřená otázka s nabídkou odpovědí

Variantu „most appropriate“ zvolilo pro slova:

terorismus (*terrorism*) (84%) respondentů a tragédie (*tragedy*) (71%) respondentů.

Teritoriální rozdíly jsou v tabulce 1.

Tato dvě označení „*terrorism*“ & „*tragedy*“ jsou vždy na prvních dvou místech.

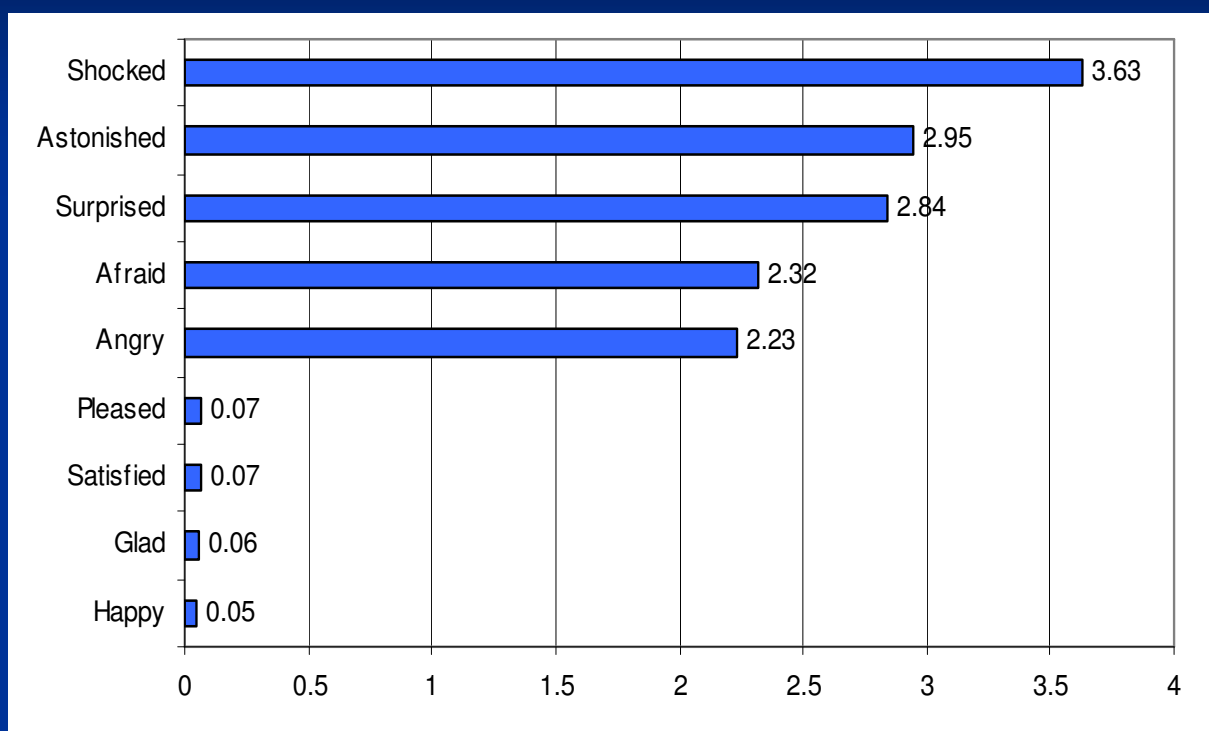
## What happened, In one word.

rank	↓ words	REPRESENTATIVE SAMPLE		ON - LINE SAMPLE			
		choice/selection: 'the most appropriate word'		item evaluation for each word choice from alternatives: % for the variant 'most appropriate'			
		Czech Republic		ON-LINE sum	Czech Rep	Europe	USA Canada
1.	<b>Terrorism</b>	<b>40.3</b>	<b>49.8</b>	<b>84.0</b>	<b>83.9</b>	<b>85.7</b>	<b>86.3</b>
2.	<b>Tragedy</b>	<b>22.7</b>	<b>28.1</b>	<b>71.4</b>	<b>73.6</b>	<b>58.9</b>	<b>83.7</b>
3.	Attack	3.1	3.9	41.4	41.6	37.6	59.2
4.	Disaster	2.0	2.5	40.5	38.7	47.8	52.0
	War attack	-	-	22.9	22.7	19.9	29.8
	Revenge	7.4	9.2	15.5	14.2	19.0	12.5
	War	1.1	1.4	12.1	11.4	12.1	21.3
	<b>Assassination</b>	4.3	5.3	27.1	22.7	<b>49.5</b>	24.0
	Don't know/ hard to say	0.0	0.0				
	Other	19.0	MV				
	None of the above	0.0	0.0				
	Total	100.0	100.0				

## Jaké emoce lidé prožívali 11. září?

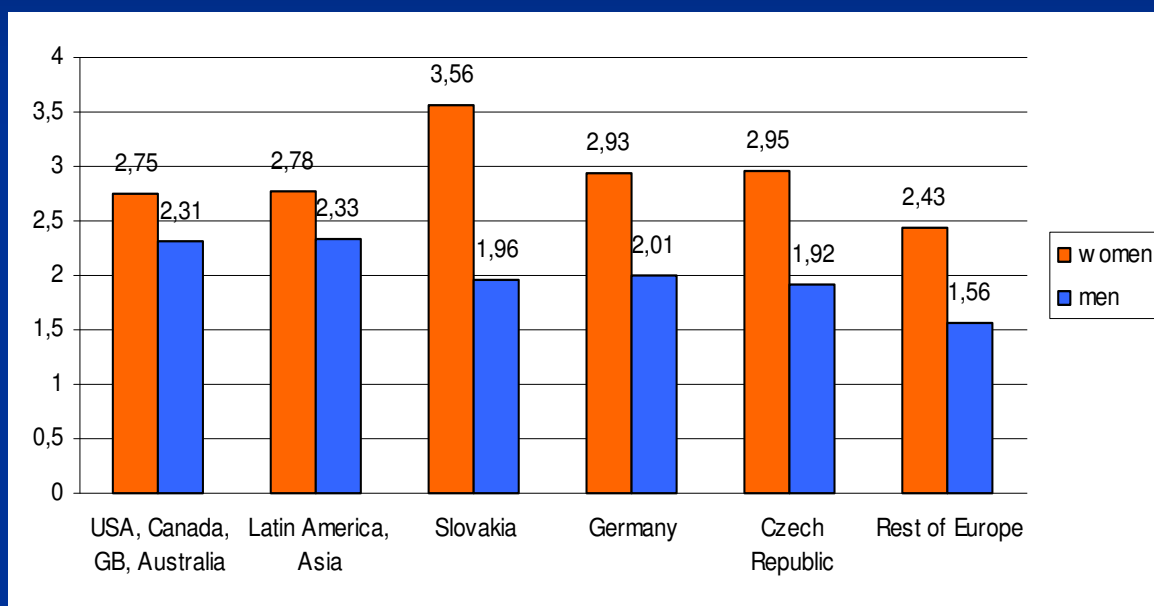
- Nejčastějším označením prožívaných emocí bylo: „**Byl/a jsem šokován/a**“ „I was shocked“ - 81%.
- Tato emoce byla **nejčastěji** uváděná **ve všech regionálních** souborech: ČR, Slo, Něm, EU, USA+Can+Austr.

### ■ How did people emotionally experience it?



## Vyjádření pocitu strachu (be afraid)

podle zemí (srovnání ženy a muži)



## Jak lidé o 11. září diskutovali ?

- 99% of the internetové populace o události hovořilo s přáteli a příbuznými.
- A to ve všech částech světa !!
- Potvrdili jsme hlavní hypotézu výzkumu: „**Události 11. září se staly důležitým podnětem pro konverzaci lidí na celém světě.**“

## *Jak dlouho lidé diskutovali?*

- **90%** všech respondentů hovořilo face to face s přáteli a rodinou **více než hodinu**;
- také **75%** Američanů a Kanadánů hovořilo o událostech **více než hodinu** (i když diskutovali méně);
- **Mezilidská osobní komunikace** o této emocionálně vypjaté události byla bezpochyby jednou **z nejdůležitějších aktivit** toho dne, odpoledne i večera ve většině rodin na celém světě.

23

## *Hovořili jste o události ve více skupinách?*

- **74%** respondentů odpovědělo '**Ano**'. Byla to typická odpověď pro Evropany, Němce (**83%**), Čechy a Slováky (**74%**), obyvatele zemí Latinské Ameriky (**76%**);
- Méně to platilo pro **Američany** a obyvatele **Asie**, avšak i zde to bylo **přes 50%** případů.

24

# Závěry:

- Naše závěry z výzkumu se **nemusí zabývat vždy celkem**
- Podstatnou výpovědí, mohou být **významné rozdíly nebo podobnosti** skryté uvnitř získaných dat

## *Dva modely práce s nereprezentativními daty*

- Model **komparace**  
a z ní vyplývající vzorce pravidelností  
**Vysvětlení rozdílů mezi podsoubory**
- Model **elaborace** a jeho varianta  
„opakování“ (**replikace**)  
**Zobecnění podobných výsledků**  
za výrazně rozdílné podsoubory

# *Diskuse*

Děkuji Vám za pozornost !

**Vaše otázky, připomínky, reakce, návrhy,  
výhrady... ?**

Kontakt: [hynek.jerabek@gmail.com](mailto:hynek.jerabek@gmail.com)