

ZURN6311 DOTAZNÍKOVÝ VÝZKUM: PROCES ODPOVÍDÁNÍ NA POLOŽKY

Lenka Dědková

ZÁVĚREČNÝ PROJEKT

- (úkolů nakonec 7 namísto 9□)
- **ZE SYLABU:**
- **Závěrečná práce:** vypracování návrhu dotazníkového šetření v min. rozsahu 8NS. Návrh bude obsahovat odůvodněnou výzkumnou otázku, detailní a odůvodněný plán sběru dat (typ sběru, cílová populace, rámec populace, strategie oslovování respondentů), zhodnocení jeho limitů a zkonstruovaný dotazník. V projektu je nutné se zaměřit na zvážení možných zdrojů zkreslení a chyb dat a snahu je minimalizovat. Součástí bude rovněž úvaha o možnostech generalizování dat nad rámec vzorku. Součástí projektu **není** reálný sběr dat. Projekt lze zpracovat individuálně nebo ve skupině max. 3 studentů. Projekt je nutné odevzdat do 11. 6. 2021 do odevzdáárny v IS. V případě, že odevzdaný projekt nebude dostatečný, bude možné jej opravit do 30. 6. 2021.

ZÁVĚREČNÝ PROJEKT

- 3 hlavní části (nemusejí být takto pojmenované v práci)
- „Teoretická část:“ základní popis zkoumané problematiky, odůvodnění VO, popis toho, na jaká sub témata/oblasti/vlastnosti se výzkum má zaměřit (nemusí jít do úrovně hypotéz, ale mělo by být podrobnější než jen vágně formulovaná VO), popis a odůvodnění cílové populace
- „Metodologická část:“ popis a odůvodnění zvolené metody designu výzkumu (dotazníkové šetření, mód sběru), popis a odůvodnění výběru vzorku (vč. konkrétních kroků, jak byste vzorek sbírali), zvážení etiky výzkumu
 - I pokud neplánujete pravděpodobnostní výběr, popište výběrový rámec (tj. na koho sampling může dosáhnout)
 - Zásadní zvážit zdroje chyb: koho ve výběrovém rámci nemáme? Koho naše konkrétní forma oslovení míjí, kdo si nepřechte pozvánku? Kdo nám na výzvu k účasti patrně nebude reagovat?
 - A co to znamená pro interpretaci výsledků? Budeme mít výsledky spíš nadhodnocené nebo podhodnocené? Na koho je budeme moci generalizovat?
- „Dotazníková část:“ připravte jádro dotazníku – není potřeba řešit vše, ale:
 - Pozvánku a/nebo úvodní text před dotazníkem
 - Naformulujte položky důležité pro dobrý popis vzorku (ať už to jsou jednotlivci anebo domácnost) vzhledem k vašemu výzkumu
 - Návrh měření alespoň jedné zásadní proměnné – konstrukt, který využívá komplexnější měření – ne jednopoložkové 😊
 - A popište, jak byste finální dotazník před ostrým sběrem ověřovali (kognitivní testování, pilotáž,.. vč. toho, koho a jak byste oslovovali)

ZÁVĚREČNÝ PROJEKT

- Z vašich úkolů by to mohly být např.:
 - Bieliková, Paulenová: seznam sledovaných typů politických participací na FB
 - Doubek: seznam důvodů, proč respondenti ne/sledovali MUNI TV
 - Ficbauerová, Labuda, Moresová: postoje vůči COVID (tady pozor, ani tady ne jen jedna položka typu „Jste pro nebo proti očkování“ - měřte komplexněji nebo rozšiřte postoje na víc oblastí)
 - Kuchta, Suchá: spokojenost ve vztahu
 - Mátlová: seznam důvodů pro pořízení auta
 - Vaňková: škála FOMO
- U některých to znamená hledání a adaptace existujících škál, u jiných tvorba nových

PROCES ODPOVÍDÁNÍ

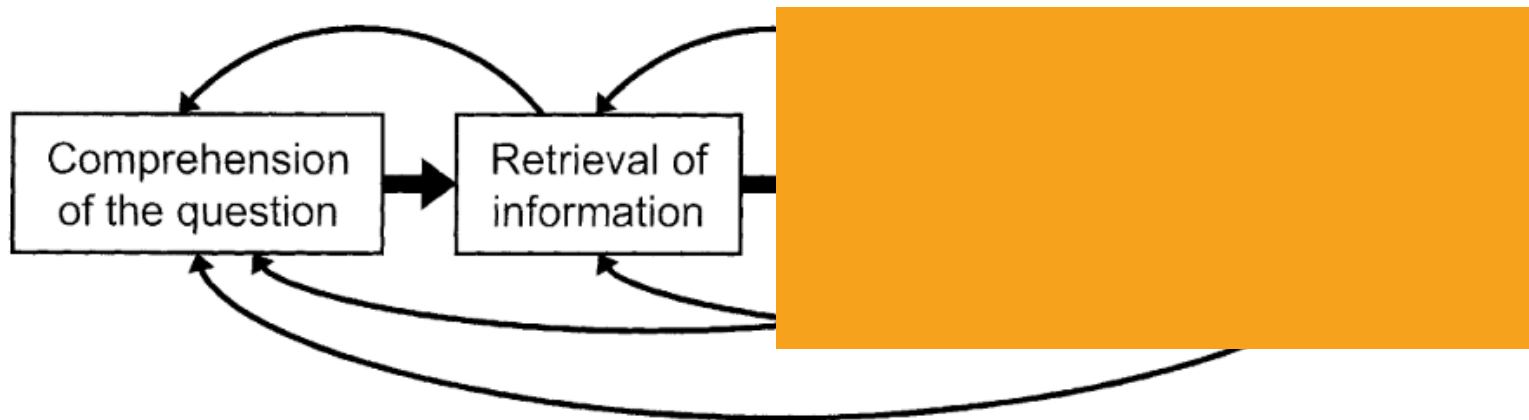
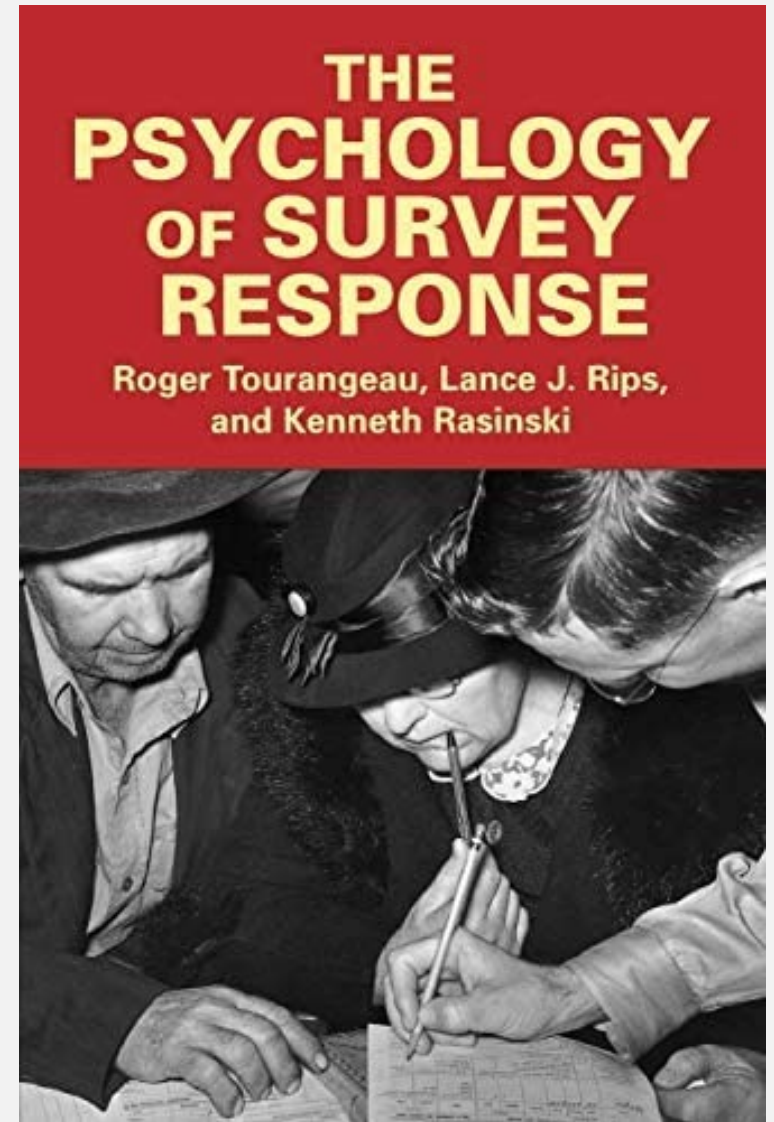


Figure 7.1 A simple model of the survey response process.

Groves et al., s.
218



- Ve studijních materiálech

KROK 3. POSUZOVÁNÍ

- Práce s tím, co jsme vylovili z paměti, tvorba odpovědi
 - Je to relevantní?
 - Co to znamená pro odpověď?
 - Musím (a chci) lovit dál?
 - Je potřeba to nějak doplnit? (odhadnout to, co si nepamatuju přesně)?

STRATEGIE POSUZOVÁNÍ FAKTICKÝCH OTÁZEK

- (Někdy máme informaci hotovou ve tvaru, který vyžaduje otázka a odpověďový formát – pak jen paměť)
- Pokud se ptáme na události, které jsou **málo časté a dobře zapamatovatelné**, pak je respondent typicky **spočítá** (tj. individuálně si je vybaví a sečte + někdy nadhodnotí pro zapomenuté události)
 - Kolikrát za život jste měl/a nějakou zlomeninu?
 - Podobných otázek v našich výzkumech ale moc nebývá
 - Čím častější a čím méně významné, tím je tato strategie obtížnější (až nemožná)
- Nebo hodnotí pro něj obvyklou, typickou míru (**rate-based estimation**) – zvláště pro události, které bývají **pravidelné**
 - Spočítá za kratší úsek, vynásobí počtem měsíců/týdnů
 - Kolikrát měsíčně nakupujete potraviny?
 - Kolik káv vypijete za běžný týden?
 - Často vede k nadsazování (u událostí které napříč časem fluktuují)
- Nebo hodnotí vágní míru, guess (**impression-based estimation**) - „eeee, není to moc časté, tak třeba 3krát“

TABLE 5.1 Strategies for Answering Frequency Questions in Surveys

Type of Information Used	Strategy
Information about specific events	<p><i>Recall-and-count</i> (episodic enumeration): Recall each event and count the events to get the total number</p> <p><i>Recall and count by domain</i> (additive decomposition): Recall and count events separately by domain</p> <p><i>Recall-and-extrapolate</i> (rate estimation): Recall a few events to estimate a rate and then project that rate over the reference period</p>
Exact tally	<i>Tally</i> : Recall current tally of events
Generic information	<p><i>Retrieved rate</i>: Retrieve existing information about rate</p> <p><i>Recommended rate</i>: Retrieve information about the recommended rate and then adjust upward or downward</p>
General impression	<p><i>Guess</i> (rough approximation; direct estimate)</p> <p><i>Context-influenced estimate</i>: Use the value given by the middle response category as an anchor and adjust it based on the impression</p>

Note: Alternative names for a strategy are given in parentheses.

POSUZOVÁNÍ HODNOTÍCÍCH OTÁZEK

- Postoje, vlastnosti, sebehodnocení, schopnosti, hodnoty...
- Pokud jsou stabilní, často o nich mluvíme, přemýšlíme, jsou výrazné, identifikujeme se jejich prostřednictvím:
 - Lze si jen vybavit odpověď (existující informace v paměti)
 - Jsou v paměti snadno dostupné
- Hodně jich ale stabilních není, 2 důvody:
 - Mění se v čase (přirozeně)
 - Nebo to jsou jen vágní postoje/hodnocení, které pro nás nejsou důležité, které reálně nemáme, které tvoříme až v momentě položení otázky („nonattitudes“)
 - I na věci, na které nemají postoj, lidé v dotazníku obvykle odpoví (pokud nemají možnost dát „nemám názor“)

STRATEGIE POSUZOVÁNÍ HODNOTÍCÍCH OTÁZEK

- Pokud odpověď nemáme hotovou v paměti nebo je zastaralá, pak 3 základní cesty:
- **Impression-based**
 - Využití vágních dojmů a obecných pocitů („*eeee, jo, asi jsem spíš pro, proč ne...*“) – stává se při nízké motivaci, nedostatku času na přemýšlení, nedostupných informací v paměti nebo malé znalosti objektu postoje
- **General attitudes or values**
 - „top→down“ proces – začneme od obecnějších postojů a od nich odvodíme postoj ke konkrétní otázce
 - Otázka „Rozpoznám kvalitní informace o zdraví“ – *jasně, poznám obecně kvalitní informace, takže určitě i ty o zdraví*
- **Specific beliefs or feeling about target**
 - „bottom→up“ proces – obráceně než předchozí

POSUZOVÁNÍ NA ŠKÁLÁCH

- Rozsah rating scales
 - Uvedeným rozsahem respondentovi implicitně říkáme, jakou odpověď čekáme
 - Častý efekt je, že respondent využije heuristiku – nabízenou škálu vezme jako rozsah, kdy na krajní hodnoty postaví extrémní chování/hodnocení, doprostřed hodnotu průměrného, typického člověka a od toho se ve svém odhadu odrazí (*jsem na internetu víc nebo míň než typický člověk? Asi o něco málo víc → dám o něco vyšší hodnotu než je půlka škály. Považuji se za člověka, co je na internetu fakt hodně? Dám se na konec škály.*)
 - U frekvencí je dobré postavit kategorie odpovědí tak, aby střední možnost byla co nejbližší (očekávanému) průměru cílové populace
 - U hodnotících dvousměrných střed na „neutrálu“

POSUZOVÁNÍ NA ŠKÁLÁCH

- Rozsah rating scales
 - U frekvencí nebo délky trvání (jak často, kolik času..) širší rozsahy vedou k vyšší prevalenci vyšších hodnot
 - Kolik času strávíte na internetu během běžného pracovního dne?

- | | |
|---------------|---------------|
| • Méně než 1h | • Méně než 1h |
| • 1-2h | • 1-2h |
| • 2-3h | • 2-3h |
| • 3-4h | • 3-4h |
| • 4-5h | • 4-5h |
| • 5h a více | • 5-6h |
| | • 6-7h |
| | • 7-8h |
| | • 8h a více |

- (k tomu, jak volit rozsah se ještě vrátíme za chvíli)

KROK 4. ZODPOVĚZENÍ

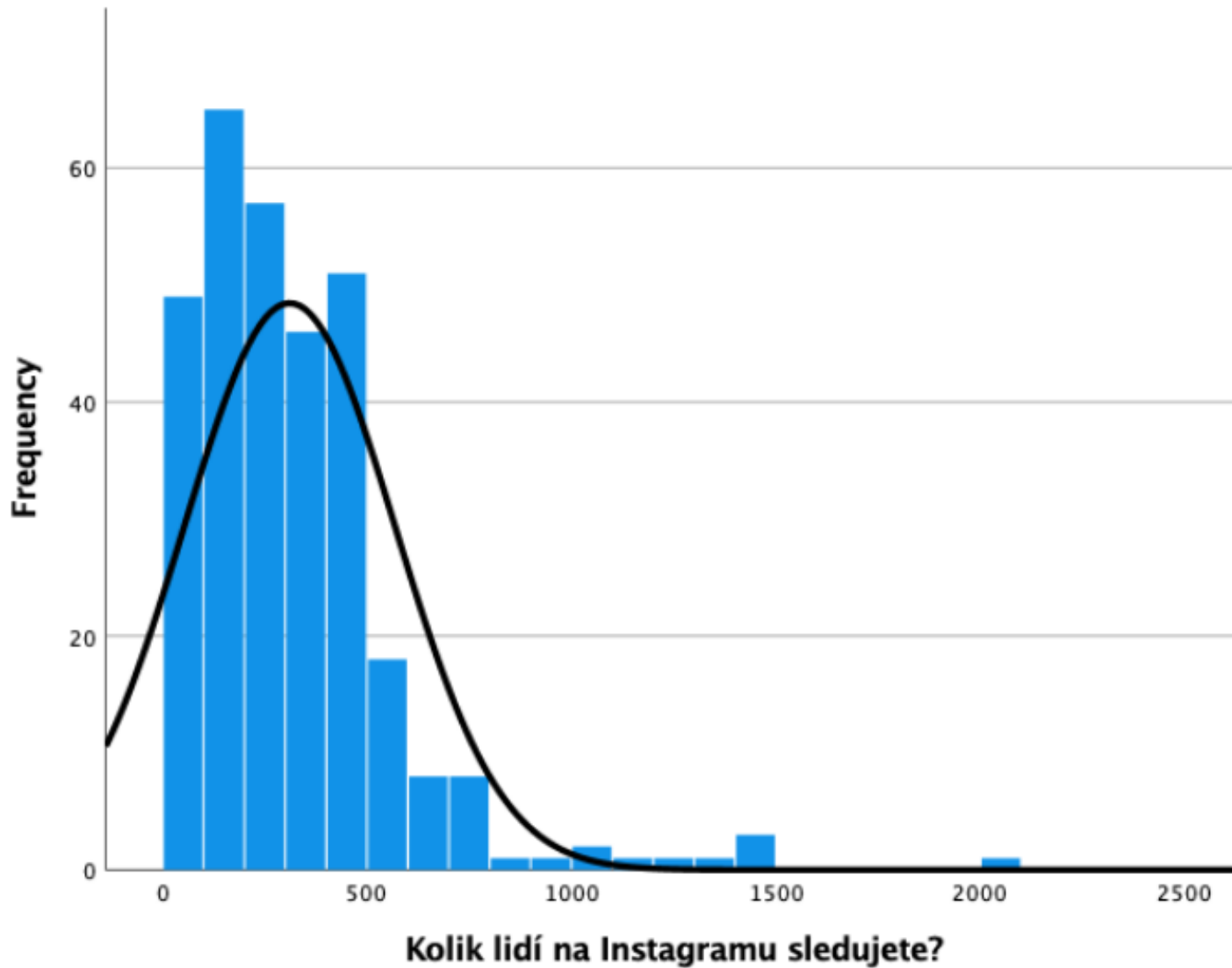
- Proces, kdy z „vnitřní“ odpovědi děláme dotazníkovou odpověď
 - V závislosti na nabídnutých možnostech odpovědí – formát odpovědí - **jsme schopni odpovědět?**
 - I na tom, jak jsme odpovídali na předcházející otázky (**efekt pořadí** – další hodina)
 - I na tom, zda (pravdivou) odpověď poskytnout chceme - **ochota**

FORMÁT: OTEVŘENÉ POLOŽKY

- Jednodušší na vymýšlení (protože neřešíme možnosti odpovědí) → méně zkreslení odpovědí ze strany výzkumníků
- Ale – respondenti do nich musejí investovat víc energie (a to neradi dělají, takže u nich bývá vyšší míra missingů) a i následná analýza stojí víc úsilí
 - Překlepy a slovní doplnění i pokud má otevřená odpověď být prosté číslo („12 minut“, „6x“, „asi deset“)
 - Slovní odpověď → kvalitativní kódování (různě jemné, různě kvalitní, ale vždy víc úsilí než kategorie odpovědí)

FORMÁT: OTEVŘENÉ POLOŽKY

- **Při otevřených otázkách na číselná fakta (počty, frekvence)**
 - Respondenti sami často přemýšlí v kategoriích („*bylo to tak něco mezi 20 a 30*“) a pak reportují **zaokrouhlené číslo**, které do intervalu spadá („*...tak dám 25*“)
 - Časté násobky 5, 10, 25, 50, 100...
 - Čím vyšší číslo reportují, tím pravděpodobněji využijí zaokrouhlování
 - Čím dál v minulosti se staly počítané události, tím spíš zaokrouhlování
 - Před analýzou otevřené položky často stejně rekódujeme do kategorií, protože v případě teoreticky neohraničených počtů nebo ohraničených vysokých počtů mívají výrazně nenormální rozložení s mnoha „dírami“ – tvar nevhodný pro analýzu
 - **Ohraničené**: např. den má 24h, týden 7 dní... - otevřený formát ok
 - **Ohraničené, ale moc široké** – limit přátel na FB (5000) – otevřený formát spíš nedoporučujeme
 - **Neohraničené, ale s nízkou očekávanou prevalencí** – např. počet účtů na sociálních sítích, počet pravidelně sledovaných zpravodajských serverů,.. - otevřený formát ok
 - **Neohraničené s potenciálně vysokou prevalencí u většiny respondentů** - otevřený formát spíš



- Naše data z minulé hodiny – počty fotek na SNS za posledních 6 měsíců.

V1_open	sV2_freq
0	Nikdy
0	Nikdy
6	Párkrát
17	Aspoň jednou za měsíc
18	Aspoň jednou za měsíc
20	Aspoň jednou za měsíc
30	Aspoň jednou za měsíc
50	Aspoň jednou týdně
50	Aspoň jednou týdně
50	Aspoň jednou týdně
52	denně
70	Aspoň jednou týdně

- Z teoreticky intervalové proměnné zaokrouhlováním vzniká ordinální proměnná

UZAVŘENÉ POLOŽKY

- Jsou pro nás o něco obtížnější na vymýšlení
- Jsou náchylnější zkreslením daným výzkumníky (ne všechny možnosti odpovědí nás napadnou)
- Pro respondenty jednodušší a rychlejší, pro analýzu také
 - (1) s nikým;
 - (2) 1-3;
 - (3) 4-6;
 - (4) 7-10;
 - (5) 11-14;
 - (6) 15-18;
 - (7) více než 18
- **Objektivní frekvence: jak stanovit kategorie, možnosti:**
 - (1) s nikým;
 - (2) 1-5;
 - (3) 6-10;
 - (4) 11-15;
 - (5) 16-20;
 - (6) více než 21
- **Stejně velké kategorie, lineární nárůst:**
 - Odhadni, s kolika lidmi, které neznáš osobně, ses na internetu za posledních 6 měsíců bavil/a:
 - (1) s nikým;
 - (2) 1-2;
 - (3) 2-4;
 - (4) 5-8;
 - (5) 9-16;
 - (6) 17-32;
 - (7) více než 33
 - 4. log 2: (1) s nikým; (2) 1-2; (3) 2-4; (4) 5-8; (5) 9-16; (6) 17-32; (7) více než 33
 - 5. něco: (1) s nikým; (2) 1-5; (3) 6-15; (4) 16 - 30; (5) 31 - 50; (6) více než 51

UZAVŘENÉ POLOŽKY

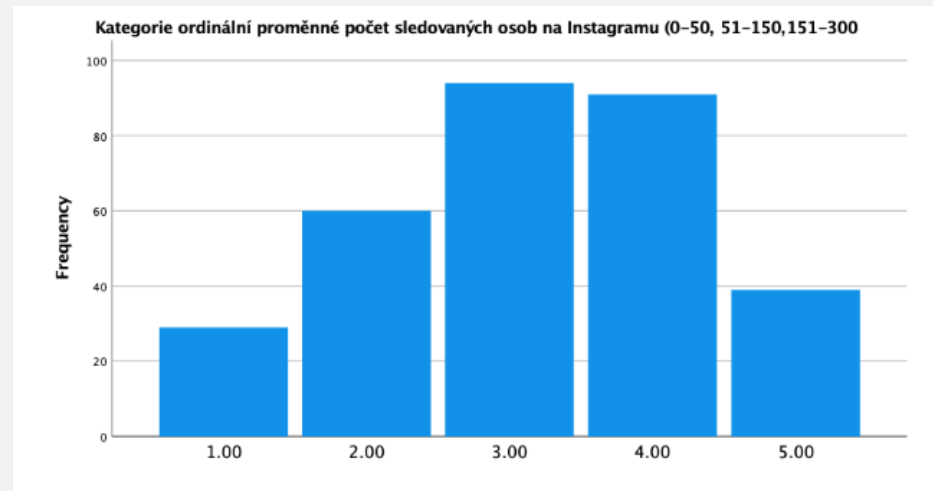
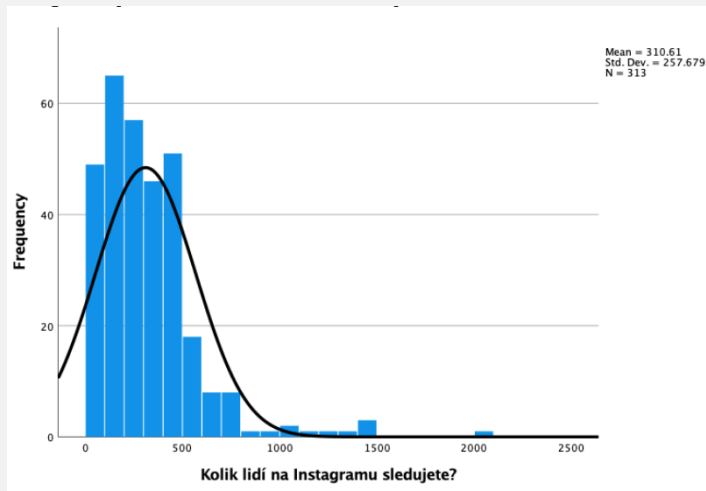
- Objektivní frekvence: jak stanovit kategorie, možnosti:
 - **Stejně velké kategorie, lineární nárůst:**
 - *Odhadni, s kolika lidmi, které neznáš osobně, ses na internetu za posledních 6 měsíců bavil/a:*
 - (1) s nikým;
 - (2) 1-3;
 - (3) 4-6;
 - (4) 7-10;
 - (5) 11-14;
 - (6) 15 - 18;
 - (7) více než 18
 - (1) s nikým;
 - (2) 1-5;
 - (3) 6-10;
 - (4) 11-15;
 - (5) 16-20;
 - (6) více než 21
 - **Postupně se zvětšující kategorie, logaritmický nárůst:**
 - (1) s nikým;
 - (2) 1;
 - (3) 2-3;
 - (4) 4-7;
 - (5) 8-15;
 - (6) 16-31;
 - (7) více než 32
 - **Postupně se zvětšující kategorie bez matematiky☺**
 - (1) s nikým;
 - (2) asi 1 - 2;
 - (3) asi 2 - 4;
 - (4) asi 5 - 8;
 - (5) asi 9 - 15;
 - (6) asi 16 - 30;
 - (7) více než 31

UZAVŘENÉ POLOŽKY

- Co je nejlepší záleží:
 - **Na očekávaném rozložení a maximálním rozsahu**
 - Ohraničené vs neohraničené
 - Kolika respondentů se týkají extrémní rozsahy
 - **V případě analytických cílů výzkumu: teorie! S čím chci dané číslo dávat do vztahu? Proč?**
 - **Předpokládám čistě lineární vztah**
 - S každým jednotlivým neznámým člověkem na internetu, se kterým se bavím, něco jiného klesá nebo roste
 - Očekávám, že každý jednotlivý člověk stejně silně sníží moji osamělost
 - → lepší lineární kategorie
 - **Nepředpokládám čistě lineární vztah**
 - Čekám, že bude větší rozdíl mezi tím, zda se bavím s nikým nebo s jedním člověkem
 - Ale menší rozdíl udělá to, zda se bavím s osmi nebo devíti lidmi
 - → lepší postupně narůstající kategorie

UZAVŘENÉ POLOŽKY

- **Velikost kategorií** – opět teorie a existující data!
 - Záleží na očekávané frekvenci
 - Volíme tak, aby byly dostatečně zastoupeny jednotlivé kategorie a nevznikaly díry (důležité pro analytický cíl)
 - Volíme smysluplně podle toho, kde čekáme rozdíly
 - Volíme zároveň tak, aby kategorií nebyl příliš velký počet
 - Pro větší čísla používáme zaokrouhlené rozsahy (protože tak lidé přirozeně uvažují)
 - U neohrazených je poslední kategorie typicky XY a víc – cutt off opět volíme podle teorie/dat



0-50
51-150
151-300
301-500
501-2000

JSME SCHOPNI ODPOVĚDĚT?

- **Mít vyčerpávající a nepřekrývající se odpovědi**
 - *Kolik z posledních 4 týdnů chodilo vaše dítě fyzicky do školy?*
 - Jeden
 - Dva
 - Tři
 - Čtyři
- **Pokud my jako výzkumníci nepotřebujeme mít vyčerpávající seznam, ale respondentům může nějaká odpověď chybět, přidáváme „else“ možnost**
 - Co pro vás a váš partnerský život platí?
 - Nezaný/á
 - Ženatý/á
 - Svobodný/á, ale mám partnera/partnerku
 - „Něco jiného,“ „jinak,“ „jinde,“ „jiné“

JSME SCHOPNI ODPOVĚDĚT?

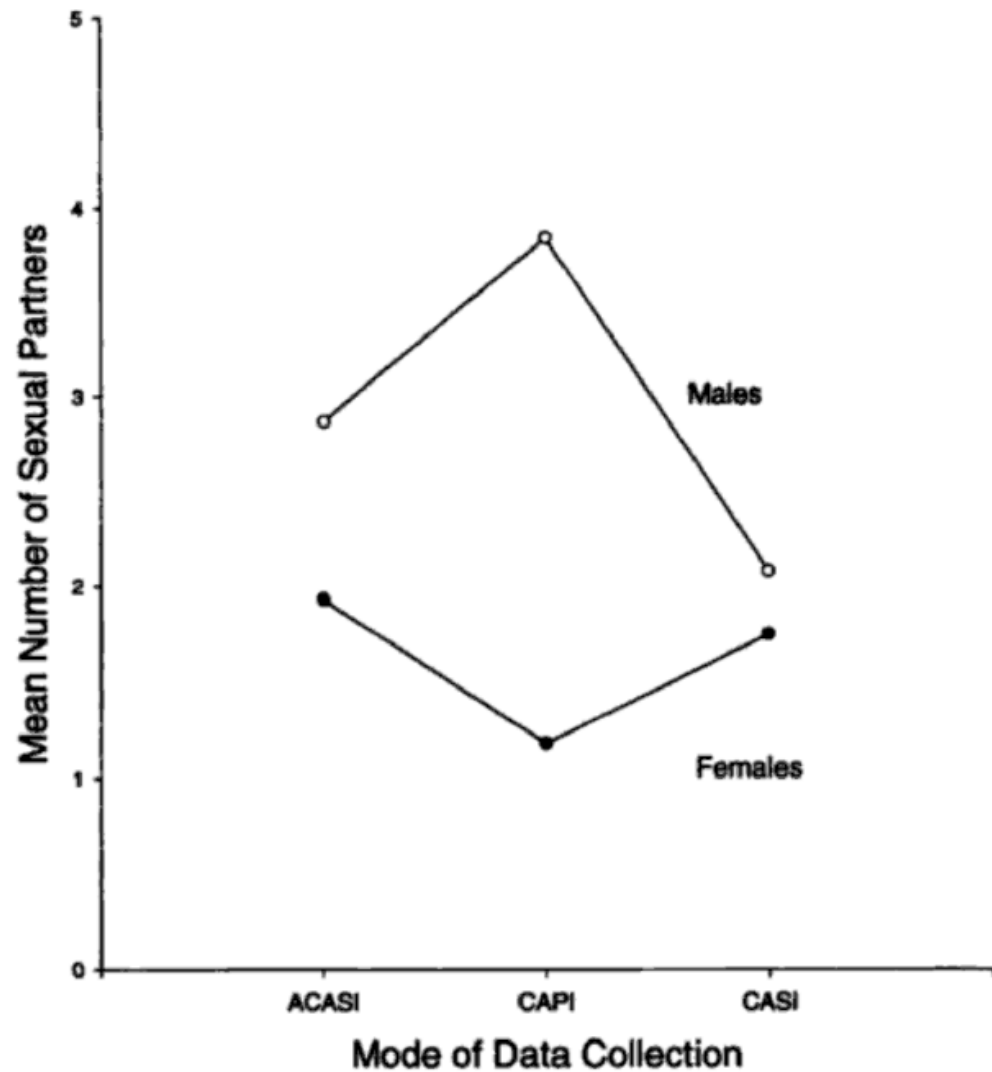
- Může být fajn mít instrukci typu „vyberte tu možnost, která vás vystihuje nejlépe“
 - Tím dáváme implicitně najevo, že to možná někomu nemusí sedět úplně nejlíp, ale ať se přiřadí k „best possible fit“ možnosti
 - Může pomoci i u hodnotících škál – pokud u kognitivního testování někomu bude chybět odpověď typu „jen velmi lehce souhlasím,“ kterou nechcete přidat, ale chcete jim usnadnit odpověď

JSME SCHOPNI ODPOVĚDĚT?

- Pozor na **mylné předpoklady**
 - „Jak důležité nebo nedůležité je pro vás vědět, jak je zabezpečená wifi síť, na kterou se připojete mimo domov?“
 - Zcela nedůležité - - - Velmi důležité
 - Předpokládá, že se připojuji na síť mimo domov
 - „Kde máte svůj telefon, když spíte?“
 - „Jak jste spokojeni se svým partnerským vztahem?“
- **Co s tím:**
 - Mít nejprve filtrovací položky („Máte vlastní mobilní telefon?“)
 - Ne vždy je to možné nebo chceme, pak: odpověď typu „nehodí se“, „nerelevantní“, „nemám XYZ“, „nedělám XY“

CHCEME ODPOVĚDĚT?

- **Citlivé nebo jinak ohrožující otázky**
 - Potenciálně mohou vést k újmě (pověst, vztahy..), pokud by došlo k porušení důvěrnosti dat nebo respondent z nějakého důvodu nechce odpověď sdělit (výzkumné agentuře, tazateli,..)
 - Otázky se sociálně žádoucí odpovědí – pro otázky, kde je ustálená (nebo jen vnímaná) sociální norma – otázky se vlastně ptají, jestli danou normu neporušujeme
 - Potenciální poškození vlastního sebeobrazu, kdyby měl respondent přiznat nějaké chování, které dělá a přitom by neměl, nebo nedělá a přitom by měl (stud, vina,..)
 - Trestní činnost, chování sociálně sankcionované nebo odmítané (kouření během těhotenství)
 - Ale i prosociální, pomáhající činnost, občanská participace (nechození k volbám, nepřispívání na charity..) nebo „správné“ a „špatné“ chování (pravidelná fyzická aktivita, prokrastinace v práci, čtení knih)
 - Otázky zasahující do (vnímané) výlučně soukromé oblasti
 - Sexualita a sexuální chování, výdělek, náboženství, ale třeba i různé tělesné projevy
- Častější chybějící odpovědi anebo uvedení nepravdivé odpovědi – nadhodnocení žádoucího a podhodnocení nežádoucího chování
 - U některých jiné tendence pro různé skupiny respondentů – gender a počet sexuálních partnerů nebo frekvence



CAPI – s tazatelem
ACASI, CASI – self-interviewing (A = s
audiem, přednahráné otázky, respondent pak
kliká odpověď)

CO S TÍM

- Největší problém u dotazování s tazatelem → udělat **self-completion sekce**
- Na některé citlivé otázky na chování pomáhá **ptát se v delším časovém rámci** (např. za celý život)
 - Protože obvykle jsou relativně málo časté
 - A umožňují respondentovi „tvářit se“ že je dané chování už dávno minulostí
 - Kriminální, delikventní typy chování apod.
- Někdy lze **přidat jiné, méně citlivé otázky** – kontext ostatních méně citlivých otázek vede k vnímání naší „target“ otázky jako méně citlivé, než kdybychom se ptali jen na ni
 - Jak často jste během posledního roku...
 - Pil energy drinky
 - Pil alkohol
 - Kouřil cigarety
 - Kouřil marihuanu

CO S TÍM

- **Normalizovat chování**
 - V současné době pandemie se někteří rodiče dětem mohou věnovat více, jiní méně nebo vůbec. Jak moc jste se ve svém volném čase (tj. mimo práci) mohli věnovat tomuto dítěti během posledních měsíců Vy?
 - V mnoha zemích je účast na komunálních volbách nízká. Byl/a jste volit v posledních komunálních volbách v roce 2018?
 - I velmi klidní lidé se někdy na své děti rozzlobí.
- **Ujistit, že je ok některé věci nezvládat, neumět, nedělat**
 - Nikdo nedokáže pracovat na 100% každý den. Jak často se vám stane, že během pracovní doby prokrastinujete...
 - Možná byste u některých otázek byl/a rád/a, kdybyste věci dělal/a jinak, ale momentálně vám to nejde. Nás teď zajímá právě to, jak často je skutečně za poslední měsíce děláte.
- **Potvrdit, že každý má právo chovat se, jak sám chce**
 - Každý rodič může přistupovat k výchově dětí jinak. Prosím, označte, co platí pro Vás a dítě.
 - Každý reaguje na stresové situace jinak. Jak...

CO S TÍM

- **Použít „hezčí“ slova pro popis chování**
 - Jak často si stahujete filmy s pomocí torrentů? (namísto „jak často nelegálně sdílíte...“)
 - Jak často vám po pití alkoholu v posledním měsíci bylo výrazně fyzicky nevolno?
- **Připomenout anonymitu**
- **Podpořit pravdivé odpovídání** a uznat, že víte, že to nemusí být snadné
- **Poděkovat** za jejich odpovědi
 - I když se vám může na tyto otázky odpovídat obtížně, prosíme, udělejte to. Jen díky upřímným odpovědím budeme schopni zjistit, co se v rodinách skutečně děje. Vašich odpovědí si opravdu vážíme.

LITERATURA

- Gideon, L. (Ed.). (2012). *Handbook of survey methodology for the social sciences*. New York: Springer.
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (Vol. 561). John Wiley & Sons.
- Tourangeau, R., Ripps, L. J., Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.