

2021

smíšený
mód:

Design a
implementace
v českém
prostředí

*manuál
pro výzkumníky a
zadavatele*

PAPICAPICATICASICAWIA
PICAPICATICASICAWIPAP
ICAPICATICASICAWIPAPI
CAPICATICASICAWIPAPIC
APICATICASICAWIPAPICA
PICATICASICAWIPAPICAP
ICATICASICAWIPAICAPIC
ATICASICAWIPAPICAPICA
TICASICAWIPAPICAPIATI
CASICAWIPAPICAPICATIC
ASICAWIPAPICAPICATICA

Vývoj multimode sběru dat a zavádění tohoto typu dotazování v oblasti populačního, sociologického a marketingového výzkumu (TL02000152)

Smíšený mód: design a implementace v českém prostředí. Manuál pro výzkumníky a zadavatele.

Autorský tým:

Barbora Hubatková, Martin Lakomý (Katedra sociologie, Fakulta sociálních studií, Masarykova univerzita, Joštova 10, Brno)

Roman Skotnica, Lýdia Barbušová, Petr Horn, Filip Rozsívál (FOCUS – Marketing & Social research, Vrchlického sad 1894, Brno)

Publikace:

Brno, 2021

Poděkování:

Za užitečné komentáře k pracovní verzi textu děkujeme prof. Martinu Kreidlovi a dr. Petru Fučíkovi.

Dedikace:

Tento manuál vznikl jako výstup projektu TA ČR v rámci programu ÉTA, projekt „Vývoj multimode sběru dat a zavádění tohoto typu dotazování v oblasti populačního, sociologického a marketingového výzkumu“ (č. TL02000152).

Podmínky užití:

Užití tohoto díla se řídí mezinárodní licencí CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0>), která umožňuje neomezené využití, distribuci a kopírování díla pomocí jakéhokoliv média, za podmínky řádného uvedení původních autorů a zdroje.

Citační vzor:

Hubatková, Barbora, Martin Lakomý, Roman Skotnica a kol. 2021. *Smíšený mód: design a implementace v českém prostředí. Manuál pro výzkumníky a zadavatele*. [online] (https://is.muni.cz/auth/publication/1811637/Manual_smiseny.mod_final.pdf).

Obsah dokumentu

Hlavní zásady	4
Úvod – k čemu je nám dotazování?	5
Jakými způsoby dotazujeme	7
Výhody a problémy tradičních módů sběru dat	8
Výhody a problémy samovyplnění online	9
Smíšený mód a jeho varianty	9
Realizace smíšeného módu v České republice	12
Postupy při realizaci smíšeného módu v českém kontextu	14
Jak připravit dotazník	17
Zohlednění módů při tvorbě dotazníku	18
Dotazník – čistě prakticky	19
Několik slov k rekrutaci	22
Sběr dat smíšeným módem	24
Zaráz nebo za sebou?	24
Rozesílání dotazníku, urgencye, odměny	24
Módy a zařízení	26
Monitoring terénu v reálném čase a helpdesk	27
Analýza dat sebraných smíšeným módem	27
Zohlednění zařízení	29
Diseminace výsledků a dat	30
Závěrem	30
Seznam literatury	31

Hlavní zásady

- Dotazníková šetření jsou stále nejvhodnějším nástrojem k porozumění některým tématům, a reprezentativní výběr (v rámci národní populace nebo specifické skupiny lidí) je stále nejověřenějším postupem.
- Jednotlivé typy dotazování – osobní, telefonické, online – trpí řadou nedostatků, které se snažíme ošetřit jejich spojováním do tzv. smíšeného módu.
- V rámci metody smíšeného módu lze spojovat různé formy dotazování. Naším doporučením je jednoduché a stále dostupnější online dotazování doplněné například o dotazování osobní.
- Pokud je naším cílem náhodný výběr z celé populace, volíme mezi metodami náhodného navolávání čísel a osobní návštěvou domácností vybraných z registru adres.
- Pečlivá příprava dotazníku je klíčovou fází procesu, jejímž výsledkem by měl být mj. dotazník použitelný v různých módech dotazování a pro různé typy elektronických zařízení.
- Krom zvažování módu platí pro konstrukci dotazníku pravidlo jednoduchosti, srozumitelnosti a omezeného rozsahu – to by měly být vlastnosti celého dotazníku i jednotlivých otázek.
- Zvláštní pozornost věnujme citlivým a postojovým otázkám, u kterých je nejvyšší riziko zkreslení i potenciálních rozdílů mezi způsoby dotazování.
- Při rekrutaci respondentů ke spolupráci se snažíme využít všech možných motivačních prvků – získání důvěry, vyvolání zájmu, nabídnutí finanční odměny/ příspěvku na charitu, zdůraznění důležitosti výzkumu, případně nabídka dalšího způsobu dotazování.
- V celém procesu sběru dat je klíčová příjemná a profesionální komunikace, opakované kontaktování neodpovídajících čísel či domácností, připomínání se u dosud neukončených dotazníků a poděkování za účast.
- Součástí budování důvěry je využití výsledků a následné informování o jejich užitečnosti.
- Různé dotazovací módy zohledňujeme i při využití dat – kontrolujeme možné rozdíly v odpovědích, bereme do úvahy možná zkreslení způsobů dotazování.
- Vliv způsobu dotazování – a rozdílů mezi módy – bereme v potaz při jakékoliv práci s daty, vzhledem k rychle rostoucímu kombinování módů v různých typech šetření.

Úvod – k čemu je nám dotazování?

Populační cenzy, „big data“, online panely, sociální sítě – to vše působí dojmem, že klasická výběrová šetření již nemohou přinést mnoho skutečně nových informací, tedy poznatků, které by nebyly přístupné nějakou jinou cestou. Proč posílat tazatele do tisícovky vybraných domácností, sbírá-li Sčítání lidu údaje o všech českých domácnostech? Proč zdržovat respondenty dotazy na to koho volí, kam jezdí na víkend, co a kde nakupují, jaký je jejich rodinný stav nebo kolik mají celkem dětí, lze-li to vše v současné době buďto dohledat v populačních registrech či dovodit z polohy mobilního telefonu, historie online nákupů, sledovaných Twitterových účtů či z fotografií na Facebooku?

Idea klasického výběrového šetření se jeví zastaralá i vzhledem k proměnám životního stylu. Ve stále zaneprázdněnější každodennosti se na straně potenciálních respondentů poměrně obtížně hledá čas i chuť prosedět 20, 30 či 40 minut nad dotazníkem, a ještě k tomu v přítomnosti cizí osoby tazatele. Od tazatelů víme, že dotazování osobní (i telefonické) je kvůli neochotě oslovených osob stále obtížněji realizovatelné. To **zvyšuje náklady na dotazování a snižuje důvěru v nasbíraná data** i zájem o tazatelské pozice. Současně se mezi respondenty zvyšují obavy o vlastní soukromí a s tím narůstá i počet otázek, které mohou být vnímány jako citlivé, od „typicky problematické“ výše příjmu, po vybavení bytu. Stále schůdnější je proto pracovat s respondenty, kteří se dotazování účastnit aktivně chtějí, tedy např. se členy online panelů.

Nadto v posledních několika dekadách dochází k nebývalému **rozmachu různých průzkumů** (ať už jde o dotazníky spokojenosti s e-shopem či studentské průzkumy) i celkově zahrnující komunikace a možnosti vyjádřit názor, především přes sociální sítě. Lidé jsou navíc vystaveni nabídkám energií, investic či mobilních tarifů na hranici podvodného jednání, které dále snižují důvěru a pocit bezpečí při náhodném oslovení (osobně nebo přes telefon). Poměrně nízká je i důvěra v oficiální instituce. V tomto kontextu se vytrácí jak ochota na dotazováních participovat, tak v zásadě i celkový smysl dotazníkových šetření.

Přes všechny tyto změny dotazníková šetření i v dnešní době smysl mají. Z metodologického hlediska je například klíčové zachování kontinuity dlouhodobých šetření, která sledují změnu v čase. Online panely, ač pojmány jako reprezentativní, jsou pak typicky výsledkem „samovýběru“ zapojených respondentů, a jako takové **opomíjejí skupiny osob méně ochotných či méně schopných** se dotazování online formou zúčastnit.

Důvody jsou ale i čistě praktické. Množství tzv. velkých dat sice narůstá, tato data ale zdaleka nejsou samospásná, jejich zpracování je časově i technicky náročné, a navíc často obsahují jiný druh informace, než jaký výběrová šetření obvykle sbírají. Další typy údajů, například státem sbíraná data o

občanech (sňatky, rozvody, počty dětí apod.), jsou zase vysoce citlivé povahy, a proto i poměrně obtížně dostupné k analýze, nemluvě o propojení s jinými informačními zdroji. Do hry vstupuje též element dobrovolnosti – jsou-li data (například fotografie, komentáře, statusy na sociálních sítích) veřejně dostupná, souhlasí jejich autoři automaticky s jejich zpracováním a analýzou? Co nám vůbec taková data o lidech říkají? A jak přistupovat k lidem, kteří konkrétní sociální síť, potažmo internet nepoužívají? Diskvalifikujeme je z výzkumů? To jest, jak v této situaci zajistíme zobecnitelnost výsledků na populaci?

Abychom tedy zodpověděli otázku v nadpisu, **dotazníková šetření i v současné době slouží ke sběru specifických a tematických údajů** (například hodnot, postojů, chování) metodologicky zakotvenými, detailně popsanými, **replikovatelnými a otestovanými postupy**. Tyto postupy zaručují, že **sesbíraná data se známým, kvantifikovatelným způsobem vztahují zpět ke společnosti, v níž byla sebrána**. Současně dotazníková data umožňují i jednoduché propojení různých proměnných (například úroveň vzdělání s postojem k rozvodu) v jediném datové souboru – a tedy zodpovězení téměř libovolných výzkumných otázek. Mají také jasně daná pravidla dobrovolnosti a anonymizace, dávají zúčastněným vyšší kontrolu nad povahou zveřejněných informací a lze na ně aplikovat standardní i nově rozvíjené postupy statistické analýzy.

Nic z výše řečeného ovšem neznamená, že by oblast výběrových šetření neměla reagovat na změny „terénu“ a podléhat inovacím. Naopak. **Inovace** jsou pro zachování relevantnosti i samotného fungování dotazníkových šetření naprosto klíčové. Jednou z inovací je vedle využívání a propojování exitujících dat, rozvoje pasivního sběru dat nebo větší míry spolupráce a dostupnosti právě tzv. **smíšený mód**. Tomuto v současnosti rozsáhle rozvíjenému využívání kombinace módů ke kontaktování respondentů a sběru jejich odpovědí se věnujeme v tomto manuálu.

Tematicky manuál sestává ze dvou obecnějších částí. První část je teoretická a ve stručnosti shrnuje terminologii dotazníkových šetření a hlavní parametry smíšeného módu. Koncipována je jako komentovaná **rešerše současné odborné literatury uvedené na konci dokumentu**. Druhá část je pak praktická a shrnuje doporučený postup při přípravě a realizaci smíšeného módu ve verzi CAWI + CAPI a při analýze získaných dat. Zde přitom **vycházíme z výsledků testování navržené metodiky přímo v terénu**.

Tento manuál je určen pro **realizátory** (výzkumné agentury, akademická pracoviště), **zadavatele** (veřejné instituce, firmy, neziskové organizace, média, akademická pracoviště) i **uživatele** (všichni výše zmínění včetně dalších subjektů pracujících s daty) dotazníkových šetření. Pro potřeby těchto skupin čtenářů manuál představuje informace pro a) realizaci smíšeného módu, b) jeho zvážení při zadávání

sběru dat a c) reflektování vzniku dat při analýze a zpracování výsledků. Jednotlivé kapitoly na sebe navazují spíše volně a je tedy možné číst je selektivně dle aktuálních potřeb.

Jakými způsoby dotazujeme¹

Dotazníková data lze sbírat několika různými, dnes již ustálenými způsoby (tzv. módy). Ty se obecně dělí na (1) **dotazování s tazatelem** a (2) **samovyplnění**. Do první kategorie spadají metody skrývající se pod akronymy PAPI, CAPI a CATI, do druhé pak PAP, CASI a CAWI (přehled zkratk v Rámečku 1). PAPI a CAPI znamenají osobní dotazování tzv. tváří v tvář, tedy s účastí tazatele a liší se jen podobou dotazníku (papírový, elektronický); CATI je dotazování po telefonu. Při PAP, CASI a CAWI dotazování vyplňuje respondent samostatně dotazník papírový, elektronický off-line (např. na počítači tazatele) nebo přímo online dotazník na libovolném přístroji.

Rámeček 1. Slovníček pojmů: módy dotazování

PAPI (*pen-and-paper interviewing*) – osobní dotazování, dotazování za účasti tazatele s pomocí papírového dotazníku

PAP (*pen-and-paper*) – samovyplnění papírového dotazníku

CAPI (*computer-assisted personal interviewing*) – osobní dotazování, dotazování za účasti tazatele s pomocí dotazníku v počítači, typicky notebooku, příp. tabletu

CATI (*computer-assisted telephone interviewing*) – dotazování po telefonu s pomocí dotazníku v počítači

CASI (*computer-assisted self interviewing*) – samovyplnění elektronického off-line dotazníku na počítači, typicky samovyplnění na počítači tazatele

CAWI (*computer-assisted web interviewing*) – samovyplnění online dotazníku v prohlížeči nebo dotazníku staženého skrze aplikaci

Volba módu sběru dat vychází například z typu výzkumu, délky dotazníku, citlivosti otázek, finančních možností či dostupných kontaktních údajů. Osobní dotazování je vhodné pro delší a komplexnější dotazníky, realizované zejména jako součást akademického výzkumu. Telefonické dotazování slouží pro dotazníky krátké, ideálně cca do 10 minut. Samovyplnění se využívá také pro dotazníky kratší a

¹ Tato část až po kapitolu „Realizace smíšeného módu v České republice“ je silně založena na existujících zjištěních, jejichž seznam uvádíme na konci dokumentu. Od uvádění zdrojů přímo k informacím v textu jsme upustili kvůli a) nevědeckému formátu tohoto díla a b) následování cíle nejvyšší možné čtivosti textu.

jednodušší či dotazníky (nebo jejich části²) zjišťující soukromé či citlivé informace (např. o partnerství a sexualitě). Zdali tazatel při práci používá dotazník papírový či elektronický často záleží na praxi a vybavení dané agentury, ale trend jednoznačně směřuje k elektronickým variantám.

V současné době je stále častěji využíváno samovyplňování online (CAWI). V českém kontextu se prostřednictvím CAWI z velké části realizuje marketingový výzkum či výzkum veřejného mínění, přičemž dotazování jsou registrovaní členové online panelů³ (tj. nikoli respondenti náhodně vybraní). Základní pravidla tohoto typu dotazování shrnuje sdružení SIMAR ve svých [standardech](#) nebo v [obecném shrnutí](#). Zájem o využití CAWI roste také ve sféře akademického výzkumu, mj. jako důsledek omezení způsobených pandemií covidu-19. Využití CAWI je však v řadě případů problematické, neboť je s metodou spojený nenáhodný výběr respondentů, nízká návratnost a související nedostatky v kvalitě dat.

Výhody a problémy tradičních módů sběru dat

Ač dlouhodobě testované a používané, „tradiční“ metody sběru dotazníkových dat (PAPI, CAPI, CATI, CASI) mají své limity. **Dotazování tváří v tvář** je charakteristické vysokou návratností a dobrou kvalitou dat, ale je finančně i časově nákladné a odpovědi na některé otázky mohou být zkresleny přítomností tazatele (např. uvádění společensky přijatelných odpovědí). Při **telefonickém dotazování** mají zase respondenti větší tendenci souhlasit s položkami nebo udávat pozitivní odpovědi, typický je také tzv. „recency effect“, kdy respondenti v průměru častěji volí poslední z nabízených možností. Současně při telefonickém rozhovoru nelze pracovat s vizuální informací, respondenti nemohou například hodnotit obrázky či videa; problémem je také vizualizace slovně popsaných dlouhých škál, anebo udržení nabídnutých variant odpovědi v paměti. CATI není vhodné pro delší dotazníky, má nižší návratnost a tíhne k podreprzentaci osob bez telefonů, osob nezvedajících cizí čísla nebo lidí, kteří nejsou schopni nebo ochotni absolvovat dlouhé telefonáty. Telefonické dotazování je ale levnější než dotazování osobní, a netýkají se ho obavy z vpuštění cizího člověka do domácnosti. Módy postavené na **samovyplnění** jsou typicky levnější než módy tazatelské, jsou vhodné pro citlivé otázky, a respondenti je vcelku preferují (to se týká zejména papírových dotazníků). Zároveň ale respondentům nemůže nikdo asistovat, pokud nějaké otázce nerozumějí nebo si nejsou jistí, jak dotazník správně vyplnit. Ztrácíme také kontrolu nad tím, zdali dotazník skutečně vyplnil oslovený respondent. Samovyplnění na

² Při dotazování metodou CAPI může například tazatel jednu či více otázek nechat vyplnit respondentem samostatně nebo ke zcela samostatnému vyplnění předat určitou část dotazníku.

³ Online panelem rozumíme existující databázi opakovaně dotazovaných respondentů.

počítači (CASI) pak může narazit na nízkou počítačovou gramotnost respondenta (při osobním dotazování).⁴

Výhody a problémy samovyplnění online

Dotazování prostřednictvím internetu (CAWI) je na vzestupu přibližně od počátku nového milénia. Jeho přínos a rozmach tkví (ve srovnání s výše zmíněnými módy) v možnosti levnějšího, rychlejšího a více automatizovaného dotazování za použití nových technologií. CAWI je pružné z hlediska času i místa vyplnění a automatizace procesu, umožňuje zařazení komplexních filtrů, videí, obrázků a grafických prvků, pohodlnou analýzu odpovědí v průběhu sběru a redukci vlivu tazatele při dotazování citlivých otázek.

Přes uvedené výhody však online dotazování trpí zásadními nedostatky, které problematizují jeho využití jako samostatného módu sběru dat. Zaprvé, CAWI dotazování má, ve srovnání s ostatními módy, nejnižší návratnost. Zadruhé, populace uživatelů internetu není totožná s populací obecnou, takže se určitá část potenciálních respondentů nemá možnost dotazování zúčastnit. Zatřetí, CAWI se zdá být obzvláště náchylné k různým druhům tzv. satisficingu, tj. typu chování, kdy respondenti dotazníku nevěnují dostatečnou pozornost. Typickým příkladem je příliš rychlé vyplnění, nebo častá volba odpovědi „nevím“. Tyto problémy nabývají na důležitosti s větší délkou dotazníku nebo nižší motivací respondentů. Za čtvrté, online dotazníky lze vyplnit na řadě různých zařízení s různými parametry (velikost displeje, operační systém, internetový prohlížeč, způsob ovládání) od osobních počítačů po chytré telefony. Různým respondentům se dotazník tedy může zobrazit různě. Mobilní zařízení současně zvyšují pravděpodobnost, že respondenti se při vyplňování dotazníku budou věnovat jiným činnostem (tzv. *multitasking*), budou se nacházet v přítomnosti jiných osob nebo budou dotazník vyplňovat na místech, která pro to nejsou z metodologického hlediska vhodná. To vše se může nepříznivě podepsat na kvalitě sebraných dat.

Smíšený mód a jeho varianty

Účelem **tzv. smíšeného módu je využití silných stránek jednotlivých metod** kvantitativního dotazování při překonání jejich největších limitů. Tento přístup přitom cílí na rychlejší sběr dat s nižšími náklady, lepším pokrytím populace, vyšší návratností a vyšší výslednou kvalitou dat, než může nabídnout většina

⁴ Ke všem základním metodám vypracovalo sdružení SIMAR [standarty](#) kvalitního postupu, jejichž dodržování je u členských agentur kontrolováno.

módů samostatně. Relativní efektivita smíšeného módu ovšem záleží na tom, se kterou dotazovací metodou jej srovnáváme. V principu je „smíšený mód“ zastřešujícím metodologickým termínem, který odkazuje k poměrně bohaté škále postupů a možných kombinací. Jeho využití je tedy vždy nezbytné důkladně naplánovat – různé kombinace jsou totiž pro různé výzkumné záměry různě vhodné.

Pojem smíšený mód může odkazovat jak k samotné fázi sběru dat, tak i k fázi oslovení respondentů, případně k jejich kombinacím. Při sběru dat jej lze nasadit buďto **souběžně** či **sekvenčně** (následně). V souběžném designu probíhá sběr dat více módy současně (paralelně) a mód dotazování se odvíjí od volby respondenta, dostupných kontaktních údajů či technické vybavenosti respondentů. Sekvenční design obsahuje mód hlavní a mód následný, cílící například na podprezentované skupiny.

Šetření smíšeným módem může být dále realizováno tzv. „**v rámci**“ respondentů (všichni respondenti absolvují všechny módy, jde třeba o osobní dotazování doplněné o samovyplnění) nebo „**napříč**“ respondenty (část respondentů je dotazována jiným módem než část druhá). Smíšený mód může také vznikat **v čase**, typicky když longitudinální šetření postupem času změní způsob sběru dat. O smíšeném módu lze hovořit také v případech, kdy v rámci jednoho mezinárodního šetření jsou data téže vlny **v různých zemích sbírána různými módy** dle místních možností. **Ve druhé části tohoto manuálu se budeme primárně věnovat diskusi smíšeného módu „napříč“ respondenty v jednom časovém bodě a jedné zemi.**

Vybrané varianty kombinace módů jsou rozebrány v Tabulce 1. Předpokladem všech variant je **smysluplná srovnatelnost** získaných dat. Jednotlivé postupy sběru dat se ovšem liší ve třech hlavních ohledech, konkrétně (1) **přítomností či absencí osoby tazatele**, (2) **způsobem přijímání informace** (slyšená, čtená) a (3) **způsobem zaznamenání odpovědi** (ústně, písemně, na klávesnici a myši, na dotykové obrazovce). Lišit se tak mohou i výsledná data, čemuž lze do nějaké míry předcházet kombinováním módů příbuzných (tj. módů s/bez přítomnosti tazatele). Takový postup ale není vždy realizovatelný; například máme-li dotazník postavený na vizuální informaci (obrázky, videa, dlouhé škály, škály s piktogramy), nelze CAPI efektivně kombinovat s dotazováním telefonickým.

Alternativou k využití stejného dotazníku pro oba použité módy (tzv. *unimode design*), je přizpůsobení dotazníku logice každého módu, což by mělo zajistit lepší srovnatelnost dat. Některé metodologické příručky však tento typ designu nedoporučují. Obecně platí, **že v současnosti neexistuje jedna „nejlepší“ kombinace módů či přesný návod, jak smíšený mód designovat.** Vždy bychom ale měli maximalizovat vhodnost dotazníku pro jednotlivé módy a srovnatelnost výsledků napříč módy.

Tabulka 1. Použití smíšeného módu pro sběr dat napříč respondenty – vybrané varianty, bez rozlišení “hlavního” a “vedlejšího” módu

	CAPI/PAPI	CATI	PAP (samovyplnění)
CAWI	<p>PLUS: Vhodné pro komplexní dotazníky i pro dotazníky s vizuálními prvky. Vhodné tam, kde je pro část respondentů nevhodné dotazování online nebo nechce interagovat s tazatelem. Vhodné pro citlivé otázky, pokud je CAPI doplněno o CASI. Vhodné pro dlouhodobější šetření, kde probíhá rekrutace a získání emailu návštěvou tazatele.</p> <p>MINUS: Nutnost získat emailovou adresu respondenta. Odkaz na dotazník může „zapadnout“ do spamu. Online dotazník je jednodušší vyplnit s nízkou mírou důslednosti, případně i nedokončit. Každý mód je charakterizován jiným typem přijímání informace (čtená/ slyšená).</p> <p>ZKRESLENÍ: recency effect (CAPI), primacy effect (CAWI), zvýšené riziko satisficingu (CAWI)</p>	<p>PLUS: Vhodné pro kratší dotazníky bez ohledu na komplexnost; vhodné, pokud část respondentů nemá čas na telefonické dotazování během běžných pracovních hodin.</p> <p>MINUS: Nepříliš vhodné pro dotazování citlivých otázek. Problém použití dlouhých škál a dalších vizuálních prvků. Každý mód je charakterizován jiným typem přijímání informace (čtená/ slyšená).</p> <p>ZKRESLENÍ: recency effect (CATI), tendence k souhlasu (CATI), primacy effect (CAWI) + zvýšené riziko satisficingu (CAWI)</p>	<p>PLUS: Vhodné tam, kde část respondentů nedisponuje přístupem k internetu, příp. nemá k online dotazníku důvěru, a pokud dotazník neobsahuje vizuální efekty (např. videa) a složité filtry. Stejný typ přijímané informace (čtená). Stejná flexibilita, co se týče časových možností respondenta.</p> <p>MINUS: Nutnost získat emailovou adresu respondenta. Odkaz na dotazník může „zapadnout“ do spamu. Data z papírových dotazníků musejí být (ručně, počítačově) převedena do digitální podoby, což je časově náročné a může dojít k chybám.</p> <p>ZKRESLENÍ: primacy effect (obojí), zvýšené riziko satisficingu (CAWI/obojí)</p>
CAPI /PAPI		<p>PLUS: Vhodné pro kratší dotazníky bez dlouhých škál a jiných vizuálních prvků. V případě delšího dotazníku může být příp. uskutečněno několik hovorů. Vhodné, pokud část respondentů nesouhlasí s návštěvou tazatele. Neliší se způsobem přijímání informace.</p> <p>MINUS: Nevhodné pro dotazníky obsahující citlivé otázky – CAPI postupy pro dotazování citlivých otázek (karty, CASI) nelze při telefonickém dotazování využít. Nelze použít obrázky nebo videa.</p> <p>ZKRESLENÍ: recency effect a social desirability bias (oba módy), tendence dávat pozitivní nebo souhlasné odpovědi při CATI</p>	<p>PLUS: Vhodné pro méně komplexní dotazníky bez velkého množství filtrů. Vhodné tam, kde část respondentů nemůže nebo nechce interagovat s tazatelem. Papírový dotazník lze zaslat poštou či zanechat ve schránce při návštěvě tazatele spolu s informačním letákem a instrukcemi k odeslání.</p> <p>MINUS: Potenciálně nižší motivace k účasti při samovyplnění. Data z papírových dotazníků musejí být (ručně, počítačově) převedena do digitální podoby, což je časově náročné a může dojít k chybám. Každý mód je charakterizován jiným typem přijímání informace (čtená/ slyšená).</p> <p>ZKRESLENÍ: recency effect (CAPI), primacy effect (PAP)</p>

V průběhu času se častým elementem designu smíšeného módu stává dotazování online⁵, typicky ve spojení s osobním dotazováním, jež se takto z doposud primární metody přesouvá k metodě dosběru v sub-populacích méně orientovaných na ICT. Tato tendence tzv. push-to-web převažuje v řadě velkých mezinárodních longitudinálních šetření, jak průřezových, tak panelových. Vzhledem k tomu, že ČR je dlouholetým členem mnoha z těchto výzkumných infrastruktur⁶, postupné zavádění smíšeného módu se odehrává i v domácím kontextu. Poznatky ze zahraničí bude přitom nebudou přímo aplikovatelné, neboť úspěšnost různých technik sběru dat, jejich prvků či kombinací je v **různých zemích různá**.

Realizace smíšeného módu v České republice⁷

Jak jsme výše uvedli, úspěšná realizace smíšeného módu je do určité míry **ovlivněna národním kontextem**. Země se liší tradicí sběru dat, chováním respondentů, technickou vybaveností, právním rámcem či dostupnými oporami výběru. V důsledku toho jsou různé módy a jejich kombinace různě efektivní. I z tohoto důvodu je obtížné – alespoň v současné době – formulovat univerzální doporučení či nejlepší strategii implementace smíšeného módu.

V tuzemském kontextu se volba módu často odvíjí od typu výzkumu: výzkum akademický (s delšími a složitějšími dotazníky) preferuje osobní dotazování (CAPI nebo PAPI), výzkum marketingový častěji preferuje dotazování po telefonu či dotazování online panelů. Smíšený mód se v současné praxi využívá, typicky jde buďto o kombinaci CAPI a PAPI (tj. tazatel používá dotazník v počítači nebo dotazník papírový) nebo, v menší míře, o smíšený mód „v rámci“ respondentů (tj. všichni respondenti vyplňují různé typy dotazníků). Jiné kombinace jsou poměrně vzácné, a chybí i systematická metodologická diskuse.⁸

Obecně lze říci, že i česká populace – stejně jako tomu je jinde v západním světě – trpí tzv. **přezkoumáním**⁹, i když by dle nás byl výstižnější termín přesycení. Problém přitom nepramení z

⁵ představuje totiž jakousi „přemosťovací“ fázi mezi tradičními postupy sběru dat a úplným přechodem do online prostředí očekávaným v blízké budoucnosti

⁶ například European Values Study, European Social Survey, Generations and Gender Programme, International Social Survey Programme

⁷ Následující části textu jsou založené především na našich zjištěních, ale do velké míry i na existujících poznatcích. Ani zde neodkazujeme na zdroje přímo v textu, a to ze stejných důvodů nevědeckého formátu díla a b) následování cíle nejvyšší možné čtivosti textu.

⁸ Razantní zlom v dosavadní praxi nastává především s příchodem pandemie covid-19, která řadu tradičních sběrů přesunula do online nebo telefonického dotazování a do možností osobního dotazování pravděpodobně přenesla částečně nevratné změny (interakce, důvěra, tazatelské sítě).

⁹ Termín „přezkoumáním“ odkazuje k jisté únavě potenciálních respondentů způsobené příliš velkým množstvím výzkumů a průzkumů různého druhu, jimž jsou každodenně vystaveni. Tato „únava“ pak ústí v neochotu na šetřeních participovat, často bez ohledu na jejich typ, téma a zaměření.

nárůstu počtu profesionálních dotazníkových šetření, ale z nevyžádaných telefonátů a návštěv různých prodejců, z přesycení newslettery, dotazníky zákaznické spokojenosti, dotazníky studentskými a podobnými typy výzkumů. Kvůli této zátěži jsou respondenti k legitimním výzkumům spíše skeptičtí. Na vzestupu jsou obavy o soukromí a tradiční metody sběru dat, například dotazování v domácnostech, přestávají být efektivní. Tento trend urychlený pandemií covid-19 se projevuje stále klesající návratností CAPI i CATI.

Současně ale dotazování obecné populace **nelze jednoduše „přesunout online“**. Pokrytí internetem zatím není úplné a výrazně se liší dle socio-demografických charakteristik domácností a jednotlivců – stejně jako používání počítačů a chytrých telefonů (viz Tabulka 2, Graf 1). Krom toho je nejefektivnějším způsobem rekrutace respondentů stále návštěva tazatele v domácnosti.

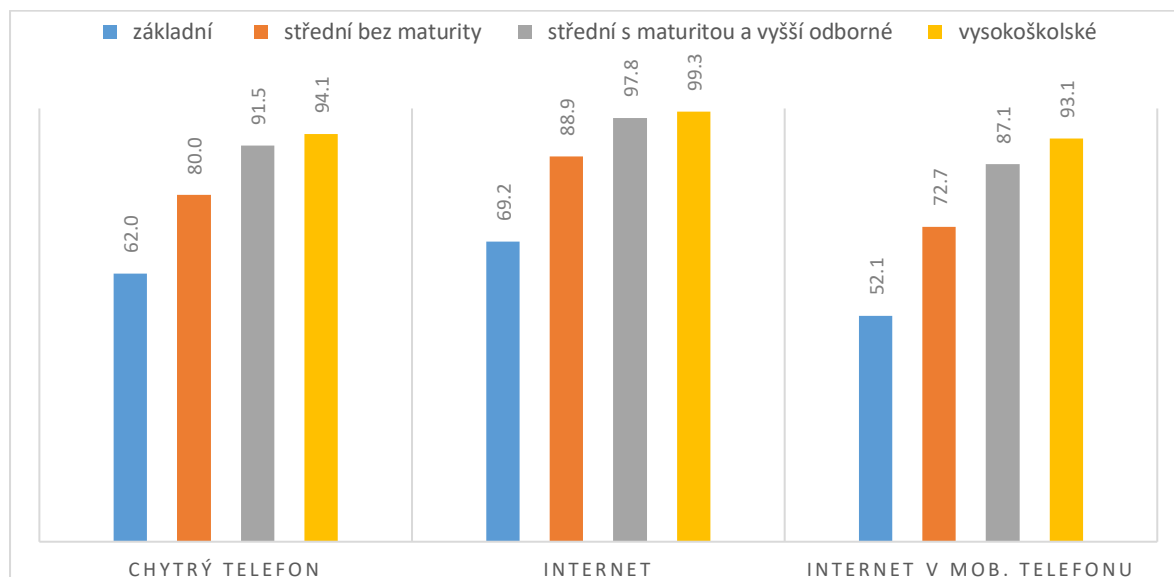
Tabulka 2. Využívání ICT v ČR – domácnosti a jednotlivci, 2020

	%
Domácnosti s počítačem celkem	79
<i>Domácnosti osob ve věku 65+ s počítačem</i>	40
<i>Domácnosti s nejnižšími příjmy s počítačem (1. příjmový kvartil)</i>	44
Domácnosti s internetem celkem	82
<i>Domácnosti osob ve věku 65+ s internetem</i>	41
<i>Domácnosti s nejnižšími příjmy s internetem (1. příjmový kvartil)</i>	50
Domácnosti s pevnou linkou celkem*	14
<i>Domácnosti osob ve věku 70+ s pevnou linkou*</i>	25
Domácnosti s mobilním telefonem celkem	99,5
z toho s chytrým telefonem	76
<i>Domácnosti osob ve věku 65+ s mob. telefonem</i>	98
z toho s chytrým telefonem	27
<i>Domácnosti s nejnižšími příjmy s mob. telefonem (1. příjmový kvartil)</i>	98
z toho s chytrým telefonem	43
Osoby ve věku 16+ používající mobilní telefon	99
z toho používající chytrý telefon	73
<i>Osoby ve věkové skupině 65+ používající mob. telefon</i>	98
z toho používající chytrý telefon	23
Osoby ve věku 16+ používající internet	81
<i>Osoby ve věkové skupině 65+ používající internet</i>	40
Osoby ve věku 16+ používající internet v mob. telefonu	68
<i>Osoby ve věkové skupině 65+ používající internet v mob. telefonu</i>	16

Zdroj: ČSÚ, 2021, Informační společnost v číslech - 2021, <https://www.czso.cz/csu/czso/informacni-spolecnost-v-cislech-2021>

*za rok 2016. ČSÚ, 2019, Informační společnost v číslech – 2019, <https://www.czso.cz/csu/czso/informacni-spolecnost-v-cislech-2018>; poslední dostupná data

Graf 1. Využívání chytrého telefonu, internetu a internetu v mobilním telefonu dle nejvyššího dosaženého vzdělání v roce 2020, osoby ve věku 25-64 let



Zdroj: vlastní zpracování dat ČSÚ, 2021, Informační společnost v číslech – 2021, <https://www.czso.cz/csu/czso/informacni-spolecnost-v-cislech-2021> (tabulky C1, C2, C3)

U populace dostupné přes různé komunikační kanály je klíčová ochota spolupracovat, kterou bychom měli podporovat všemi prostředky a) u všech oslovených, a ještě speciálně b) u pravidelně podreprezentovaných skupin. V rámci poskytování vnější motivace příliš nefungují loterie nebo dárky (dárek jen v rámci CAPI), efektivnější jsou peněžní poukazy, a především odměna v hotovosti. Odměna v hotovosti nefunguje jen při vyplacení respondentovi, ale i jako možnost věnovat odměnu na dobročinné účely (především v designu, kde není možné hotovost jednoduše předat). Ochota také závisí na tématu dotazování. Při přípravě šetření je tedy užitečné myslet i na to, jak je rámovat a představit respondentům.

Postupy při realizaci smíšeného módu v českém kontextu

V následující části manuálu se již budeme plně věnovat praktickým náležitostem implementace smíšeného módu v českém prostředí. Konkrétně představíme **nejperspektivnější kombinaci dotazování online (CAWI) a osobního dotazování (CAPI), a to napříč respondenty**. Doporučení nejsou sice univerzální, vycházejí ale ze zahraniční literatury a z testování tohoto designu v českém prostředí (více Rámeček 2). Jsou tedy (s určitou mírou reflexe) dobře aplikovatelné pro různé typy šetření.

Rámeček 2. Východiska – design a testování smíšeného módu v rámci projektu TAČR „Vývoj multimode sběru dat a zavádění tohoto typu dotazování v oblasti populačního, sociologického a marketingového výzkumu“

V tomto manuálu vycházíme ze závěrů projektu „Vývoj multimode sběru dat a zavádění tohoto typu dotazování v oblasti populačního, sociologického a marketingového výzkumu“ realizovaného mezi lety 2019 a 2022 pracovníky Katedry sociologie na Fakultě sociálních studií Masarykovy univerzity a agenturou FOCUS – Marketing & Social research. Jak je z názvu patrné, projekt se zaměřil právě na **návrh a testování nejperspektivnější varianty smíšeného módu v českém prostředí**. Testována byla kombinace online a osobního dotazování (CAWI a CAPI) „napříč“ respondenty. Účelem projektu dále bylo inovovat software FocusIS tak, aby jeho pomocí bylo data smíšeným módem možné kvalitně a efektivně sbírat.

Projekt měl sociologickou a marketingovou větev a smíšený design byl testován na dotaznících typických pro obě tyto oblasti. Marketingový dotazník se zaměřil na životní styl, nakupování, hospodaření či „radosti a neřesti“. Sociologický dotazník se zaměřil na tematické oblasti spojené s partnerstvím, politickými preferencemi či kvalitou života a obsahoval i řadu citlivých otázek. Oba **dotazníky byly záměrně navrženy dlouhé (40-50 min)**, aby bylo možné otestovat ochotu respondentů takové dotazníky (typické zejména pro sociologická šetření) vyplnit online. Respondentům náležela odměna za vyplnění, které se mohli vzdát ve prospěch daru na charitu. Online dotazník bylo možné průběžně uložit a vrátit se k němu později.

Vzorek CAWI respondentů byl sestaven náhodným výběrem a doplňkový vzorek pro CAPI kvótním výběrem z obecné populace ve věku 18+. Projekt nepracoval se členy online panelů.

Design smíšeného módu v rámci tohoto projektu lze ve zkratce popsat takto:

- (1) rekrutace metodou náhodného navolávání čísel
- (2) zaslání odkazu, sběr dat metodou CAWI
- (3) dosběr respondentů prostřednictvím CAPI se zařazením CASI pro citlivé otázky

Testování smíšeného módu se zaměřilo na tyto hlavní úrovně:

- (1) Implementace – testovali jsme, zdali a jak úspěšně lze v současnosti data dvěma zvolenými módy, zejména pak online módem, sesbírat. Testování ukazuje, že **po internetu lze skutečně sesbírat větší část dotazníků** a CAPI lze tedy nasadit „jen“ jako doplňkovou metodu pro sběr podreprezentovaných skupin. Technicky je tedy navržený design bez větších problémů realizovatelný.
- (2) Efektivita – náklady jsou vyšší než při použití online panelu, s ambicí dosáhnout lepšího pokrytí populace. Zároveň jsou náklady nižší než při klasickém osobním dotazování a rekrutaci, a to za cenu rizika nižší kvality vzorku. Srovnání s CATI není relevantní, tento mód se nabízí k jiným typům dotazování (kratší, bez citlivých témat).
- (3) Návratnost – **míra návratnosti je spíše nízká**, i když ji nelze vyčíslit ani v rámci CATI rekrutace (nespolehlivé rozlišení neodpovídajících a nepoužívaných/neexistujících čísel) ani v rámci kvótního osobního dosběru. Skupiny podreprezentované v CAWI lze kvótně doplnit v rámci CAPI, anebo je částečně doplnit osobním náhodným výběrem.

(4) Efekt módu – dále jsme testovali, zdali a do jaké míry jsou data sesbíraná CAWI a CAPI módem srovnatelná, jaká je jejich kvalita a jak v českém kontextu vypadá efekt módu. Zaměřili jsme se mj. na srovnání otázek na příjem, partnerské a sexuální chování, finanční problémy a exekuce, společensky problematické chování (hazard, marihuana, alkohol). Výsledky indikují, že **efekt módu v českých datech jednoznačně existuje, jeho síla a závažnost se ale liší dle povahy otázek**. Nalézáme ho zejména u otázek citlivých (navzdory využití CASI v osobním dotazování) a u baterií, přičemž kvalita dat z baterií se jeví vyšší v online dotazování. Efekt módu je ale velmi složité odlišit od vlivu výběru a odlišného složení vzorku. Smíšený design CAWI+CAPI, jakožto i ekvivalenci jím získaných dat, bude nezbytné dále testovat v rámci budoucího metodologického výzkumu.

Dotazování online je v našem designu primárním módem. Účelem je sebrat co nejvyšší podíl dotazníků právě prostřednictvím CAWI. To by standardně mělo být jednak rychlejší, jednak úspornější. **Rychlejší** proto, že respondent může dotazník vyplnit de facto okamžitě po obdržení odkazu. Rychlost i časová náročnost tedy mohou být výrazně nižší než u jiných módů. Současně lze bez větších problémů využít delší, komplexnější dotazník s vizuálními prvky. Navíc není rychlost sběru dat omezena počtem tazatelů. **Úspornější** proto, že není nutné hradit výdaje a odměny tazatelů. Úspornost je ale třeba brát s jistou rezervou – finanční prostředky je nezbytné vkládat do rekrutace, ať už telefonické, osobní či poštovní. Dále je důležité respondenty motivovat odměnou za vyplnění. Vzhledem k aktuálnímu stavu terénu mohou být tyto vstupní náklady vysoké, měly by ale zaručit dobrou návratnost i kvalitu získaných dat.

Osobní dotazování (CAPI) pak slouží k dosběru skupin systematicky podreprezentovaných v rámci hlavního módu. V tomto ohledu je možný osobní dosběr určitých kvótních znaků nebo náhodný osobní výběr (s předpokladem vyššího zastoupení mladších skupin v online dotazování a starších skupin v osobním dotazování). V ideálním případě respondenty pro oba módy vybíráme tímž způsobem (např. pravděpodobnostním či kvótním výběrem). Pokud to není možné, je nutné brát v úvahu, že efekt módu (o němž bude řeč níže) nebude při analýze dat možné spolehlivě odlišit od efektu metody výběru.

Ještě poznamenejme, že **CAPI je v našem designu preferováno před PAPI**, protože pro obě skupiny respondentů může být použit totožný dotazník (např. náročné filtry, dlouhé škály, obrázky a videa apod.). Využití CAPI či PAPI často záleží na praxi agentury; s agenturou nebo institucí realizující sběr dat je tedy nutné se domluvit na přesné podobě osobního dotazování.

Jak připravit dotazník

Příprava dotazníku je – vedle metody výběru vzorku – klíčovým bodem každého dotazníkového šetření. Bez **kvalitního výzkumného nástroje** totiž nemůžeme získat kvalitní, validní a reliabilní data. A pokud nemaximalizujeme šanci na získání takových dat, nemá realizace šetření prakticky žádný smysl a ve výsledku jen přispívá k přezkoumání populace, tj. ke zhoršujícím se podmínkám pro sběr dat. Kvalitní data by současně neměla být cílem sama o sobě – primární motivací všech dotazníkových šetření by mělo být zodpovězení jasně formulovaných výzkumných otázek, ať už za účelem akademické práce, návrhu veřejných politik, nebo splnění zadání klienta. Obecně by tedy měla příprava dotazníku začít až po přesném vymezení cílů sběru.

I jsou-li naše záměry jasně definovány, by přípravě vlastního sběru měla předcházet **rešerše existujících, dostupných datových zdrojů**. Kýžená data již mohla být sebrána v rámci tuzemských či mezinárodních šetření, a to pravděpodobně ve větším rozsahu a kvalitě, než jsou běžní zadavatelé schopni dosáhnout. Příkladem informačně bohatých a spolehlivých datových zdrojů jsou mj. Evropský výzkum hodnot (European Values Study), Evropský sociální výzkum (European Social Survey), Survey of Health, Ageing and Retirement in Europe či měsíční šetření CVVM *Naše společnost*. Mnoho datových zdrojů je po jednoduché registraci dostupných například v Českém sociálněvědním datovém archivu (ČSDA). I pokud takovéto datové soubory neobsahují všechny potřebné informace, jsou užitečné jako zdroj prověřených otázek pro vlastní šetření. Chceme-li se ptát na politické preference, pocit osamělosti, postoje k manželství či na spokojenost se životem, je vždy lepší použít ustálené formulace nežli vymýšlet otázky nové. Jinými slovy, otázky nemusejí být (a často by ani být neměly) originální, nýbrž spolehlivé (jasné, srozumitelné, nezavádějící, otestované).

Tím se ovšem již dostáváme k pravidlům přípravy dotazníku, jejichž diskuse přesahuje zaměření tohoto manuálu. Pro více informací čtenáře odkazujeme na příslušnou metodologickou literaturu. Ve stručnosti však shrňme, že dotazník by měl být co nejkratší, srozumitelný všem skupinám v cílové populaci, a měl by být pro všechny i relevantní, tematicky i obsahově by měl být pro respondenty zajímavý a neměl by obsahovat otázky, které by respondenti mohli subjektivně vnímat jako urážlivé. Při určování pořadí otázek se snažíme o logickou provázanost, od otázek jednoduchých a obecných ke složitějším, se sociodemografickými otázkami na konci (pokud je potřebujeme vědět). Samotné otázky by měly být jednoduché, jasné, s plným pokrytím možných odpovědí včetně „nevím“ či „nechci odpovídat“, důkladně zkontrolované několika páry očí. Tzv. otevřené otázky je dobré používat minimálně.

Zohlednění módů při tvorbě dotazníku

Ke všem obecným pravidlům spojeným s přípravou instrumentu pro dotazníkové šetření přibývá v našem případě ještě **nutnost zohlednit specifika obou použitých módů**, tedy CAWI a CAPI. Pravidlo přípravy dotazníku použitelného ve více módech platí vždy, když existuje možnost přesunu dotazování do jiného módu. Jak shrnuje již Tabulka 1, zatímco CAPI je mód postavený na interakci mezi tazatelem a respondentem (tj. tazatel čte otázky, respondent odpovídá), CAWI mód je postavený na samovyplnění (respondent otázky čte i zodpovídá). Módy se tedy liší způsobem přijímání informace (sluchem vs. vizuálně), který je v literatuře spojován s poněkud odlišnými **kognitivními procesy**¹⁰. Příkladem mohou být dlouhé škály: slovní popis „Odpovězte pomocí škály od 0 do 10, kde 0 znamená *vůbec* a 10 *zcela*“ je něco jiného, než vidíme-li tuto škálu před sebou¹¹ jako

Vůbec 0 1 2 3 4 5 6 7 8 9 10 Zcela

Navíc, při samovyplnění si respondenti mohou otázku i odpovědi přečíst kolikrát potřebují, zatímco v dotazovací situaci může stud nebo nedostatek času opakování omezit na minimum. Ani sebestřednější formulace pak nezaručí, že obě skupiny respondentů zodpovídají de facto tutéž otázku. Platí totiž, že při samovyplnění může respondent dotazníku věnovat tolik času, kolik potřebuje, případně vyplňování i docela přerušit a vrátit se k němu později (umožňuje-li to dotazník), zatímco při vyplnění s tazatelem existuje úzký časový rámec, na který se interakce musí omezit (např. 60 minut). Na druhou stranu takto omezená a jasně strukturovaná dotazovací situace zaručuje, že dotazník je úspěšně vyplněn; u samovyplnění hrozí, že respondent nakonec dotazník kvůli odkladům nedokončí. Příliš rychlé vyplnění v obou módech (a příliš pomalé vyplnění především v CAWI) nám přitom indikuje potenciální problém s kvalitou dat.

Příbuzným problémem, který se týká především CAWI respondentů, je **zobrazení dotazníku na různých zařízeních** a v různých internetových prohlížečích. To opět znamená, že různí respondenti dotazník vidí a vnímají různě a různě jej mohou i zodpovídat. Například naše škála od 0 do 10 se pohodlně vejde na monitor běžného stolního či přenosného počítače i tabletu, na displeji mobilního telefonu (zejména při orientaci na výšku) může ale dojít k jejímu „useknutí“, zalomení nebo neúměrnému zmenšení, např.:

**Vůbec 0 1 2 3 4 5
6 7 8 9 10 Zcela**

¹⁰ Zároveň z každého typu dotazování systematicky vypadávají respondenti se zrakovými nebo sluchovými omezeními, což může být pro specifická témata důležité zkrácení.

¹¹ Tento problém ale není zas až tak signifikantní. Osobní dotazování u dlouhých škál či náročnějších výčtů standardně využívá papírové karty nebo otočení monitoru k respondentovi. Jde tedy zejména o to při přípravě CAPI dotazování na tyto vizuální pomůcky nezapomenout.

Při použití mobilu tak respondenti mohou častěji vybírat odpověď, která je zrovna viditelná, nikoli tu, která nejlépe odpovídá jejich názoru či situaci. Kvůli nepřehlednosti dotazníku také mohou dotazování zcela přerušit. Na druhou stranu, respondenti **často volí zařízení, se kterým sami nejčastěji pracují**. Parametry použitých zařízení, které bychom mohli vnímat jako problematické, tedy nutně nemusejí na kvalitu dat působit – právě proto, že respondenti jsou na své zařízení zvyklí. Rozdílům mezi zařízeními se ještě budeme věnovat v podkapitolách o sběru a analýze dat.

Důležité ke zvážení jsou **citlivé otázky**. Zatímco samovyplnění je pro takové otázky vhodné, osobní dotazování může být problematické. Zodpovídá-li respondent citlivou otázku položenou tazatelem (např. „užíval/a jste někdy drogy“, „kolik jste v životě měl/a sexuálních partnerů“, „pokusil/a jste se někdy o sebevraždu“), může mít tendenci odpovídat dle společenských konvencí, nikoli pravdivě, čímž v datech dochází ke zkreslení. V krajním případě je pak respondent takovými otázkami zaskočen natolik, že dotazování ukončí. To je ovšem eventualita, které se zcela nevyhneme ani při samovyplnění. Před zařazením potenciálně citlivé otázky tedy zvažte, je-li skutečně potřeba, a pokud ano, přidejte k ní krátký úvod či vysvětlení; např. „Různí lidé mívají různé pohledy na svět/ životní zkušenosti/ názory atp.“. Vhodné je také využití CASI módu.

Módy se liší i prací se speciálními odpověďmi typu „nevím“ či „nechci odpovídat“, které bychom měli v online dotazování (kde jsou zpravidla všechny otázky povinné) nabízet tam, kde jejich zařazení dává smysl. Tazatel tyto varianty většinou aktivně nenabízí. Ve smíšeném módu lze zvážit i jednotný přístup a nabízet speciální odpovědi v obou variantách (upřímné „nevím“ je cennější informací než v podstatě vynucená věcná odpověď). V našich sběrech jsme speciální odpovědi nabízeli jen v CAWI (v CAPI nastávaly spontánně) a jejich množství se významně nelišilo, takže to prozatím přijímáme jako funkční praxi (při opatřeních jako administrace citlivých otázek přes CASI a nabídku CAWI variant). V CAWI můžeme tyto varianty také vizuálně odlišit od platných odpovědí, abychom simulovali jejich implicitní přítomnost při dotazování CAPI. To lze udělat odsazením, umístěním ke kraji stránky, umožněním přechodu na další stranu bez zakliknutí odpovědi nebo zvýrazněním možnosti „nevím“ při pokusu otázku přeskočit.

Dotazník – čistě prakticky

Při konstrukci dotazníku pro sběr dat kombinací CAWI a CAPI doporučujeme následující:

Dotazník konstruuje jeden, v nejvyšší možné shodě pro CAWI (s pokyny pro respondenta, jsou-li nutné) a CAPI (s pokyny pro tazatele). Jinými slovy, dotazník je nutné přizpůsobit oběma módům tak,

aby v nich byl realizovatelný týmž způsobem (tzv. *unimode design*). V této kombinaci módů se možnosti CAWI přizpůsobují limitacím CAPI:

Formulace otázek by měla být natolik jasná, aby respondenti otázku snadno pochopili jak psanou, tak slyšenou formou. **Možnosti odpovědí** by také měly být jasné, nepřiliš obsáhlé. U složitějších otázek a dlouhých škál, baterií či výčtů v CAPI navrhujeme využít CASI mód (otočení monitoru k respondentovi) a/nebo papírové karty. Dochází tak je sjednocení způsobu přijímané informace alespoň tam, kde by potenciálně mohly vznikat problémy. Tím by mělo dojít ke zmenšení efektu módu.

Citlivé otázky je třeba ošetřit v CAPI módu. Stejně jako v předchozím případě navrhujeme strategické využití **CASI módu**, kdy respondent příslušné otázky vyplňuje samostatně na počítači tazatele, s co nejvyšší mírou soukromí (více viz Rámeček 3). Mezi typicky citlivé se řadí otázky na příjem, zdravotní stav, tzv. patologické chování či sexuální život. Různé skupiny respondentů ale mohou citlivost konkrétních otázek vnímat různě. Míru citlivosti lze zjistit v rámci důkladné pilotáže.

Délka dotazníku by měla být v rámci potřeb výzkumu co nejmenší.¹² Je třeba myslet na to, že ačkoli CAWI respondenti typicky mohou dotazník vyplnit i na několikrát, CAPI respondenti tuto možnost obvykle nemají. CAWI respondenti budou současně méně motivovaní, aby dlouhý dotazník dokončili. Proto je především v CAWI důležité zapojení motivačních prvků (peněžní odměna vs dar charitě za dokončený dotazník, důležitost tématu, využitelnost výsledků, nepostradatelnost respondenta, hladká a zdvořilá komunikace, důvěra ve výzkum a realizátora).

Rámeček 3. CASI mód

Computer-assisted self-interviewing (CASI) **definujeme jako samovyplnění části naprogramovaného dotazníku na počítači/tabletu tazatele** jako součást osobního dotazování. Vyplnění je v největší možné míře **soukromé**, tj. otázku respondent čte i zodpovídá sám, s monitorem otočeným k sobě, a odpovědi následně sám ukládá tak, aby tazatel neznal odpověď. CASI mód je nutné v dotazníku uvést pokynem “otočte monitor k respondentovi” a doplnit i krátkým vysvětlením, co má respondent udělat. Ani CASI nemůže plně simulovat samovyplnění v soukromí, bez tazatele. Už samotná přítomnost cizí osoby má totiž vždy potenciál odpovídání zkruslit. Mód vyplnění je tedy třeba zohlednit při analýze.

¹² 2000 slov včetně numerických znaků v odpovědích rovná se cca 15 minut hovoru, CAWI může být mírně rychlejší. Často se dozvíme, že i pro online dotazování se hodí krátké dotazníky do 10 minut, například v sociálních vědách je ale dotazování často koncipováno na dolní hranici 30 minut. Naše vlastní testování ukazuje, že hlavním problémem je spíše rekrutace respondentů nežli dlouhý dotazník: pokud už respondenti dotazník rozkliknou, vyplní jej ve většině případů celý bez ohledu na délku. Současně mohou využít možnosti dotazník vyplnit po částech.

Vizuální podoba dotazníku se musí odvíjet od skutečnosti, že CAPI respondenti dotazník buďto neuvidí vůbec, nebo uvidí pouze některé jeho části. Vzhledem k tomu, že i vizuál (např. barvy, animované přechody mezi stránkami, ilustrační obrázky apod.) může odpovídání ovlivnit, doporučujeme jeho střízlivé využití. Použít samozřejmě lze obrázky, videa nebo piktogramy (např. škály se smajlíky), musejí je ale vidět všichni respondenti. V pokynech pro tazatele u vybraných otázek tedy nezapomeňme na „otočte monitor k respondentovi“. Tento pokyn také můžeme použít u dlouhých škál nebo baterií; případně lze tazatele vybavit papírovými kartami.

Kvůli poměrně široké škále **různých zařízení** (stolní počítače, notebooky, chytré telefony, tablety), na které lze CAWI dotazník vyplnit, je třeba zvážit počet otázek na jedné stránce dotazníku či fungování dlouhých škál na malých displejích. Například v delších bateriích můžeme rozložit každou položku na zvláštní stránku; takové opatření mírně zpomaluje dotazování, ale nevedlo v našich dotazování k odpadávání respondentů a značně zvýšilo kvalitu dat. Dlouhé škály je lepší omezit nebo nahradit výběrem číselné odpovědi v rolovátku (to ale zase přináší více středních a extrémních hodnot).

Respondentům v pokynech můžeme vhodné zařízení **doporučit**, neměli bychom ale zařízení vyžadovat či přímo blokovat „nehodná“ zařízení. Může tím totiž dojít k posílení vlivu selekce (různé zastoupení zařízení v různých skupinách osob), posílení chyby měření (pokud bude část respondentů pracovat se zařízením, na které není zvyklá) či rovnou odmítnutí účasti v šetření. Vhodnějším řešením je nastavením dotazníku umožnit změnu zařízení v průběhu vyplňování (tj. dotazník lze průběžně uložit, zavřít a znovu otevřít na jiném zařízení). Doporučit také lze nejvhodnější orientaci malých mobilních zařízení (na výšku, na šířku).

Kombinace CAWI a CAPI předpokládá, že dotazník bude **naprogramován ve specializovaném softwaru**. Sbírá-li pro nás data agentura, pak je třeba dopředu konzultovat, jaké možnosti jimi využívaný software nabízí. Obecně doporučujeme naprogramovat dotazník tak, aby jej bylo možné průběžně uložit a vrátit se k němu později. To by mělo zohlednit délku dotazníku, časové možnosti CAWI respondentů a umožnit případnou změnu zařízení. Funkce ale může přijít vhod i tazatelům v terénu. Hodit se mohou i možnosti průběžného nahlížení do sbíraných dat a ukládání obou módů do společné databáze.

Dále bychom doporučili:

- velmi stručnou a přehlednou první stranu dotazníku
- detailnější pokyny až na druhou stranu nebo do průvodní zprávy
- umožnit vracení se dotazníkem
- vynucovat odpovědi (CAWI), ale s nabídkou speciálních odpovědí

- ukazatel průběhu u dlouhých dotazníků nezobrazovat, u krátkých dotazníků bez složitých filtrů můžou motivovat

Použití specializovaného softwaru by také mělo umožnit sběr tzv. **paradat**, tedy doplňkových údajů o dotazovací situaci, které mohou být využity při kontrole kvality dat. Typickým údajem je délka vyplnění dotazníku; v závislosti na typu softwaru by ale mělo být možné posbírat i údaje jako čas strávený na stránce dotazníku, změny odpovědí, listování dotazníkem nazpět, typ použitého zařízení, operační systém, geografická poloha respondenta¹³ apod. Možnosti softwaru je opět nutné konzultovat se subjektem sbírajícím data.

Několik slov k rekrutaci

Dobře připravený dotazník je pouze jednou částí úspěšného sběru dat. Dalším klíčovým elementem je výběr a rekrutace respondentů. Především o druhém tématu se vedou intenzivní metodologické diskuse. Zatímco v akademické sféře převládají zastánci tradičních metodologických postupů (tj. různé typy náhodného výběru), zástupci agentur, které pro vědce data často sbírají, doporučují i výběry nenáhodné, např. výběr kvótní. Upozorňují totiž na to, že v českém kontextu obecně – na rozdíl od zemí, které využívají populační registry – není náhodný výběr v současné době smysluplně realizovatelný. Na jednu stranu jsou respondenti nejen méně ochotní k účasti, ale i hůře zastizitelní na náhodně vybraných adresách. Na straně druhé se snižují i počty tazatelů a jejich pokrytí prostoru; a to kvůli zhoršujícímu se stavu terénu, intenzivnější kontrole ze strany agentur i dlouhodobému výpadku činnosti v období pandemie. I v případě smíšeného módu zůstávají tyto problémy relevantní a představují jednu z hlavních **bariér** úspěšné osobní rekrutace i sběru dat.

V našem projektu jsme se rozhodli nepoužít metodu náhodného výběru adres nebo metodu náhodné procházky (obojí přes tazatele), ale rekrutovat respondenty **náhodným navoláváním čísel** (tzv. RDD – Random Digit Dialing), což je další z ověřených způsobů rekrutace náhodného vzorku. Na každé číslo doporučujeme volat alespoň třikrát v různé denní doby, aby byl vzorek co nejvyváženější. Do procesu byla zařazena mobilní telefonní čísla a čísla na pevné linky v poměru 90:10. Ačkoli v současnosti již většina obyvatel ČR vlastní mobilní telefon, zatímco pevné linky jsou na ústupu a koncentrují se například ve firmách, považujeme zařazení pevných linek za metodologicky korektní kvůli lepšímu pokrytí podreprezentované starší části populace (podrobněji viz Tabulka 2).

¹³ Poloha a jiné podobně citlivé údaje by měly být sbírány se souhlasem respondenta.

I telefonická rekrutace náhodného vzorku naráží na několik úskalí. Proces generuje velké množství neexistujících čísel. I pokud ale číslo existuje, je problémem nezvednutí, a navíc není vždy možné rozlišit čísla nedovolaná od neexistujících. I tam, kde se operátor dovolá, je častým výsledkem odmítnutí, které je zhruba stejně časté jako při rekrutaci tazateli na fyzických adresách. Následně ještě použitelný podíl respondentů účast přislíbí, ale jen necelá polovina z nich dotazník skutečně vyplní. Celková míra návratnosti tedy nejde přesně spočítat, ale je velmi nízká. Některé parametry respondentů (pohlaví, kraj) sice odpovídají populaci, ale zastoupení z hlediska věku a (především) vzdělání jsou vychýlené. Náklady na lidské zdroje (operátoři call centra agentury) jsou nižší než u tazatelů, ale tuto výhodu snižují vysoké odměny respondentů.

Poznamenejme ještě, že některé navolané osoby budou na telefonát reagovat negativně a řada bude chtít vědět, odkud má volající jejich soukromé číslo. Je tedy potřeba připravit **krátké vysvětlení** logiky náhodného vytáčení čísel a dovolaného ujistit, že číslo nebylo převzato z žádné databáze a do žádné databáze ani nebude uloženo. Ačkoli takto nelze zaručit souhlas s účastí, je takovýto postup přinejmenším korektní.

V rámci rekrutačního hovoru je třeba co nejrychleji získat **důvěru** dovolaného, především zmíněním věrohodného zadavatele výzkumu, renomé agentury (která by měla zajistit snadné dohledání informací na webu) zároveň s finanční odměnou. Dle naší analýzy jsou nejčastějšími důvody odmítnutí účasti obecný nezájem, nedostatek času v danou chvíli, hovor v práci, případně nedostupnost internetu (při rekrutaci pro CAWI). Pro takové situace je dobré mít nachystané flexibilní plánování času volání, i když každý z těchto důvodů je pro část volajících pouze výmluvou. V ideálním případě by respondenti bez přístupu k internetu měli být přesunuti do osobního módu (tzv. souběžný design). Naše testování ovšem ukazuje, že a) jen velmi málo těchto osob souhlasí s návštěvou tazatele a b) systém telefonické rekrutace pro CAWI s následným vysláním tazatele za konkrétním respondentem je velmi nákladný. Poznamenejme ale, že pro další práci s daty je nezbytné, aby byly respondenti **pro oba módy rekrutování týmž způsobem**, ať už náhodně či nenáhodně (např. kvótně). V opačném případě nebude možné analyticky odlišit případný vliv módu sběru dat od vlivu sestavení vzorku.

Obecně lze shrnout, že telefonní rekrutace náhodného vzorku není zcela ideální, ale při nižším rozpočtu a nižším nároku na návratnost (ve srovnání s osobní rekrutací) celkem funguje. Návštěva tazatele například ve spojení s vysokou finanční odměnou zůstává zlatým standardem pro rekrutaci náhodného vzorku, ale její udržitelnost může nadále klesat jak z hlediska finančních nákladů a ochoty vpouštět tazatele do domácnosti, tak z hlediska klesající hustoty tazatelských sítí. Online nebo telefonní panely jsou alternativami cílícími na kvótní výběr, u nichž kvalita dat závisí především na povaze a tématu šetření.

Sběr dat smíšeným módem

Zaráz nebo za sebou?

Jak jsme již shrnuli v první části tohoto manuálu, sběr dat smíšeným módem může probíhat dvěma způsoby, tj. **paralelně** – souběžně (kdy sběr oběma módy probíhá současně), nebo **sekvenčně** – postupně (kdy nejprve proběhne sběr jedním, poté druhým módem). Výběr konkrétního postupu závisí z části na znalosti zkoumané populace. Umíme-li dopředu odhadnout, které skupiny respondentů se budou pravděpodobně účastnit konkrétním módem, není důvod, proč zvolené módy nenasadit souběžně. Například při volbě náhodného (nebo kvótního výběru) pro oba módy lze spustit oba módy současně. Naopak nejsme-li si úspěšností rekrutace pro daný mód zcela jisti, můžeme s nasazením druhého módu počkat, až budeme mít k dispozici základní charakteristicky vzorku a v módu druhém cílit na specifické sub-populace. Při tomto postupu ale padá ambice na náhodný výběr.

Z našeho dosavadního zkoumání lze formulovat doporučení pro sběry dat v obecné populaci a při rekrutaci pro CAWI po telefonu. Ze sesbíraného vzorku, ve srovnání s obecnou populací ČR, systematicky vyjadávají skupiny osob s nižším vzděláním (vzdělání základní a střední bez maturity) a osoby nad 65 let (při nadreprezentování jediné věkové skupiny 45-54 let). K získání většího zastoupení skupin s nižším vzděláním lze používat cílené motivace, jako více pokusů či upomínek, navýšení odměny v upomínce, odmítání CAWI respondentů s vyšším vzděláním. Alternativou je osobní rekrutace nebo výběr z online panelu. Určitou nevyváženost náhodného výběru lze řešit i váhami.

Rozesílání dotazníku, urgency, odměny

V případě CAWI lze respondentovi dotazník odeslat v co nejkratší době po souhlasu s účastí na šetření (a získání emailové adresy) a tento postup **doporučujeme**. Respondent má takto dotazník téměř ihned k dispozici a může jej začít i rovnou vyplňovat. Odpadá možná prodleva mezi rekrutací a samotným dotazováním, což podporuje zrychlení celého procesu a vyšší návratnost. K tomuto postupu doporučujeme softwarové řešení (např. viz Rámeček 4), neboť ruční rozesílání je náročnější a pomalejší. Pokud software k dispozici nemáme, je nezbytné systematizovat proces tak, aby byly prodlevy malé a aby při rozesílání dotazníků nedocházelo k chybám, duplicitám apod. Čím větší je náš vzorek, tím důležitější tato koordinace je. V principu by měla stačit i jednoduchá tabulka, je ale naprosto klíčové, aby údaje v ní byly aktualizovány v reálném čase (a zároveň byly plně zabezpečeny).

Rámeček 4. Software FocusIS

V rámci projektu jsme nevyvíjeli jen optimální design smíšeného módu, ale i softwarové řešení této metody. Agentura Focus jako spoluúčastník projektu dlouhodobě používá a vyvíjí software FocusIS, který byl přizpůsoben současným nárokům na online dotazování, a především **dotazování ve všech hlavních kombinacích módů**. FocusIS tak získal sadu funkcí v oblasti základního provozu, formátu otázek, automatizace kontaktování a distribuce dotazníku, kompatibility módů a další.

Klíčové pro metodu smíšeného módu jsou především možnosti:

- a) sdílení kvót mezi módy
- b) správa kontaktování mezi módy
- c) možnost pokračovat ve vyplňování na jiném zařízení
- d) automatické rozesílání nově narekrutovaných + připomínek + poděkování
- e) automatické/ruční plánování času hovoru
- f) přizpůsobení různým typům zařízení
- g) zabezpečená archivace nových typů dat

Nikoliv nezbytné, ale metodologicky přínosné jsou možnosti automatického sběru paradat (čas sběru, mód, typ zařízení a jeho vlastností, volitelně i GPS polohy) nebo provádění tzv. split-ballot experimentu (náhodné přiřazování podmínek dotazování nebo variant otázek/odpovědí respondentům).

Na českém trhu a v českých agenturách nebo mezinárodních pobočkách českých agentur jsou samozřejmě dostupná další vyhovující softwarová řešení. Vždy je ovšem vhodné si před začátkem spolupráce ověřit, zda jsou naše výzkumné představy u agentury plně technicky proveditelné.

Je ovšem zřejmé, že ne každý CAWI respondent dotazník vyplní ihned po obdržení. Důvody mohou být různé, od momentálního nedostatku času po strategii „souhlasit a nevyplnit“, kde je souhlas s účastí de facto jen únikem z rekrutační situace. Jak dlouhá prodleva je či není přijatelná závisí na konkrétním designu výzkumu a případných časových omezeních terénních prací. Obecně lze ale říci, že pravděpodobnost, že respondent dotazník vyplní, se v čase snižuje. Neboť každý souhlas s účastí „se počítá“, v designu výzkumu by neměly chybět **urgence** (upomínky). Jejich časování a podoba jsou opět záležitostmi designu, určitě je ale dobré, aby přinejmenším jedna upomínka proběhla jiným módem, než jakým probíhá sběr dat. V našem projektu se osvědčila emailová připomínka po pěti dnech od souhlasu s účastí a telefonická připomínka (zahrnující kontrolu emailové adresy) po osmi dnech od souhlasu s účastí; další připomínky ani prodlužování intervalu neměly věcný přínos.

Ochotu k účasti lze podpořit nasazením **odměn za vyplnění**. Nejvhodnější pro smíšený design je **odměna finanční** (spíše než dárky, poukázky, loterie), ve výši dle možností našeho rozpočtu a náročnosti šetření (typicky v řádu stokorun). V CAPI lze sice jednoduše předat všechny typy odměn, hlavním motivačním prvkem je osoba tazatele, a kromě finančního obnosu funguje i dárek. V CAWI či CATI je ale obtížné nově narekrutovaným respondentům odměnu odevzdat, stejně jako je přesvědčit o poctivosti nabídky.

Zařadit lze také možnost darovat svou odměnu **na charitativní účely**. To jednak může respondenty motivovat k účasti; za druhé může tato možnost sloužit jako smysluplná „úniková varianta“ pro ty, kteří nechtějí vyplnit osobní údaje nutné pro vyplacení odměny. Pokud dar na charitu jako možnost zařadíme, je o jeho uskutečnění respondenty vhodné nějakým způsobem informovat, např. zveřejněním informace na webových stránkách agentury či naší instituce, organizace. Výběr charitativní organizace je citlivá záležitost, v rámci které je na základě dílčích poznatků (např. z Focus groups) vhodné vyhnout se organizacím napojeným na církve, orientovaným na otázky ochrany přírody, pomoci uprchlíkům apod. Variantou je pracovat s obecným „darem charitě“ a zvolit organizaci po skončení sběru nebo v komplexnější variantě dávat na výběr z organizací.

Ve čtyřech našich šetřeních příslib finanční odměny sice CATI rekrutaci napomáhal, ale spíše, než její absolutní výše fungoval jakýkoliv příslib. Kolem 60 % respondentů se ale nakonec odměny vzdalo ve prospěch dobročinných účelů. I když výše přislíbené odměny neměla dopad na míru návratnosti, vyšší odměna ovlivnila ochotu vyplnit osobní údaje k jejímu vyplacení ve srovnání s možností vzdát se odměny ve prospěch dobrovolné organizace.

Na závěr poznamenejme, že ani sebelepší konfigurace odměn a upomínek nikdy není plně úspěšná. Je ale třeba mít na zřeteli, že účast na dotazování je dobrovolná, a že tedy má **každý respondent právo dotazník jednoduše nevyplnit**, bez ohledu na předchozí souhlas.

Módy a zařízení

Ačkoliv o tomto tématu byla již řeč výše, pro úplnost shrňme hlavní body: Pokud neudáme jinak, budou CAWI respondenti online dotazník vyplňovat **na různých zařízeních s různými parametry** (různé operační systémy, osobní nastavení, velikosti displeje, typy klávesnic apod.). Data z našeho projektu indikují, že chytrý telefon k vyplnění využívá zhruba pětina online respondentů, stolní počítač zhruba třetina respondentů, notebook necelá polovina respondentů a tablet jen přibližně 5 % online respondentů.

Vhodné zařízení můžeme respondentům v pokynech **doporučit**, neměli bychom je ale nařizovat, abychom se vyhnuli chybám měření či přímo ztrátám pozorování. Pokud jsme různá zařízení zohlednili již ve fázi přípravy a programování dotazníku, nemělo by docházet k problémům. Jako prevenci odpadávání respondentů kvůli špatné volbě zařízení doporučujeme naprogramovat dotazník tak, aby bylo možné jej **přerušit a otevřít jej znovu na jiném zařízení**. O této možnosti můžeme respondenty informovat v průvodním dopise nebo na úvodní straně samotného dotazníku.

CAPi mód by měl být ve vztahu k zařízení bezproblémový tam, kde data sbírá profesionální agentura. Ta totiž své tazatele buď vybavuje stejným či podobným typem přenosného počítače, nebo tazatele proškolí a poskytne jim specializovaný software. Pokud data sbíráme sami – a nemáme-li možnost vybavení pro tazatele nakoupit, pak bychom si měli alespoň udělat přehled o zařízeních, která budou tazatelé používat a případně i přidat „zařízení“ jako proměnnou do CAPI verze dotazníku.

Monitoring terénu v reálném čase a helpdesk

Základní **přehled postupu terénních prací** probíhající v reálném čase by měl patřit ke standardům každého sběru dat bez ohledu na mód, obzvláště v případě módu smíšeného je ale naprosto klíčový. Může totiž upozornit na případné problémy a umožnit jejich včasné řešení. Při sekvenčním designu pak slouží k určení toho, které skupiny se budou dosbírávat druhým módem.

Standardem by také měl být komunikační kanál k řešení dotazů či problémů respondentů, ať už telefonní číslo nebo speciálně založený email. Respondentů, kteří se k nám budou obracet pro radu či se svými komentáři pravděpodobně sice nebude mnoho, naši pomoc ale rozhodně ocení. Kontakt na zadavatele (především výzkumníka, státní orgán, ne firmu) zvyšuje důvěryhodnost šetření, ale pro praktické otázky je lepší uvést kontakt na subjekt zajišťující sběr dat.

Analýza dat sebraných smíšeným módem

Vzhledem k rozdílům mezi módy (viz Rámeček 1) je nutné zohlednit skutečnost, že různé postupy sběru dat mohou generovat odlišná data (tzv. **efekt módu**). Respondenti se stejnými charakteristikami totiž mohou v různých módech odpovídat odlišně, a to právě vlivem použitého módu, například proto, že jedni otázku slyší, zatímco druzí ji čtou, nebo kvůli různé míře soukromí. Například vdané vysokoškolačky ve věku 30-34 let mohou při otázce na spokojenost se životem na naší škále v průměru volit hodnu 8,0 odpovídají-li v CAWI módu a hodnotu 8,7 odpovídají-li v CAPI módu. Zahraniční literatura dokumentuje dobrou ekvivalenci příbuzných módů a drobné systematické rozdíly mezi módy s a bez tazatele (které jsou kombinovány v rámci našeho designu). Jinými slovy, mód je nezbytné nějakým způsobem zohlednit v analýze, abychom předešli zkresleným výsledkům. Na tyto možné rozdíly a potřebu jejich kontroly by měla u smíšených šetření upozorňovat **dokumentace k datům**.

Standardním východiskem pro další analytický postup je jednoduchá **explorační analýza**, v jejímž rámci se díváme, zdali a jak se odpovídání na jednotlivé otázky liší podle módu a zdali jsou rozdíly mezi módy věcně a statisticky významné. To lze učinit prostřednictvím dvourozměrných tabulek a srovnání průměrů vč. obvyklých statistických testů (t-test, chí-kvadrát; takové srovnání funguje při relativně

podobném složení vzorků, jinak rozdíly musíme kontrolovat, viz dále). Vhodné ke srovnání jsou otázky citlivé a postojové. Předmětem srovnání může být i podíl neplatných odpovědí a chybějících hodnot, nejen prosté podobnosti/rozdíly v odpovědích platných. My jsme například srovnávali chybějící odpovědi (varianta „nechci odpovědět“) mezi módy pro otázku „Jaký je celkový měsíční příjem celé Vaší domácnosti?“ a „Měl(a) jste někdy pohlavní styk?“. První otázku odmítlo zodpovědět 12 % CAWI a 14 % CAPI respondentů, druhou 9 % CAWI a 10 % CAPI respondentů. I při kontrole odlišného složení vzorků je tento indikátor plně srovnatelný (obě otázky tazatelé nabízeli prostřednictvím CASI).

Zjistíme-li v rámci explorační studie, že odpovědi se podle módu statisticky významně (či věcně) liší, je vhodné přikročit k **analýze dat komplexnějšími a sofistikovanějšími statistickými metodami**, které lépe ukáží, zdali lze data smysluplně používat jako celek (více viz Rámeček 5).

Poznamenejme ale ještě, že rozdíly v odpovídání, zejména rozdíly zjištěné na deskriptivní úrovni, **nemusejí být nutně výsledkem efektu módu**. Mohou totiž vyplývat ze **specifik složení vzorků**, typicky demografických charakteristik (tam, kde dosbíváme podreprezentované skupiny) či selekce (tam, kde respondenti sami volí preferovaný mód). Součástí explorační studie by tak měla být i **kontrola** efektu módu pro základní sadu charakteristik respondentů (pohlaví, věk, vzdělání, velikost obce, kraj apod.).

Rámeček 5. Některé statistické postupy pro testování ekvivalence dat

Ačkoli základní deskriptivní statistiky vybraných proměnných (podíly, průměry, chybějící odpovědi apod.) poskytují užitečný základní vhled do datového souboru, dostatečně přesně nezodpovídají otázku ohledně ekvivalence dat. Jinými slovy, umíme sice najít rozdíly, nevíme ale, jsou-li způsobeny samotným módem, selekcí, nebo nějakou jinou systematickou chybou.

Způsobů, jak pokročit za elementární deskripci, je v literatuře diskutována řada, a v blízké době jich bude pravděpodobně přibývat. Jedním z nich může být například **regrese**: v rámci testování navrhne model logistické či lineární regrese s proměnnými, jejichž jednotlivé efekty jsou z dřívějších analýz známy. Model pak odhadneme na datech za každý mód zvlášť i za oba souhrnně a výsledky porovnáme. V úvahu přichází také srovnání se s jinými datovými zdroji či skutečným stavem, například u volebního chování. Vychází-li modely výrazně jinak, data pravděpodobně nelze za ekvivalentní považovat.

K určení několika stupňů ekvivalence (někdy terminologicky také „invariance“) slouží tzv. **multi-group konfirmační faktorová analýza (MG CFA)**. V jejím rámci odhadujeme sérii modelů, z nichž každý následný model má striktnější specifikaci než model předchozí. Srovnáním modelů pak

můžeme potvrdit či vyvrátit ekvivalenci metrickou a ekvivalenci skalární. Spolehlivá data by přitom měla být ekvivalentní na obou těchto úrovních.

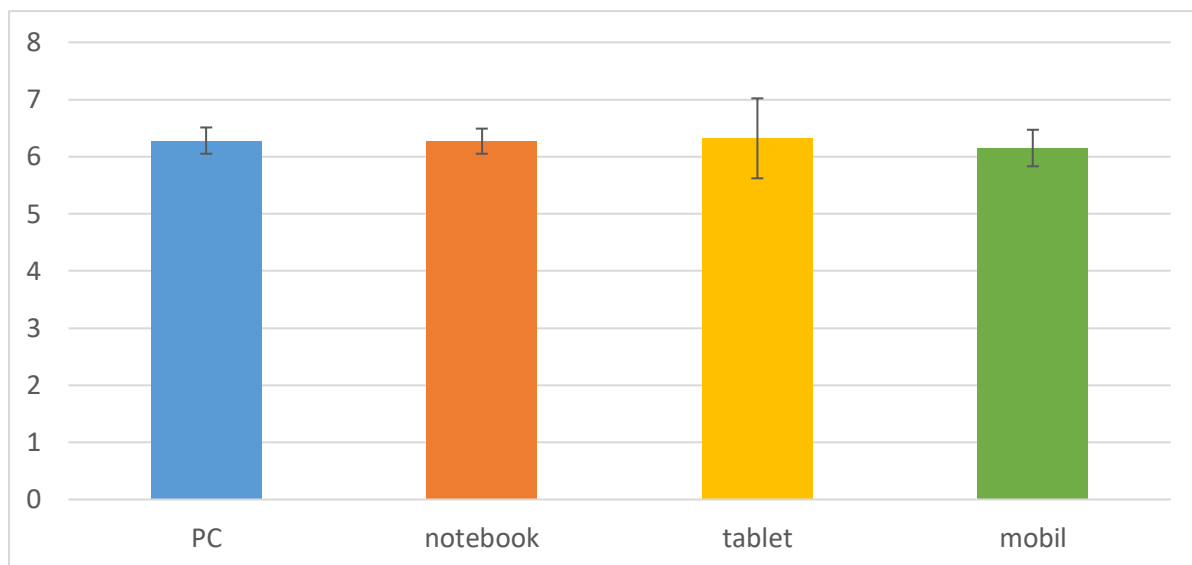
Nepodaří-li se nám doložit ekvivalenci dat sesbíraných různými módy, je třeba provést validizaci prostřednictvím logických kontrol, jiných zdrojů dat, a analyzovat data s módem jako kontrolní proměnnou (pokud data z obou módů považujeme za relevantní, tedy ne zcela zavádějící).

Zohlednění zařízení

Vše shora řečené platí i pro **zařízení** použitá pro vyplnění online módu. Ačkoli řada zahraničních studií neukazuje mezi zařízeními s ohledem na kvalitu a srovnatelnost výsledných dat velké rozdíly, zařízení se typicky liší délkou vyplnění, přičemž vyplnění trvá nejdéle na chytrých telefonech. Nemusí to být jen kvůli kvalitě připojení (tj. prodlevami mezi přechodem na další stránku dotazníku), ale i kvůli delšímu času strávenému na stránce. To může znamenat, že se mobilní respondenti při vyplňování věnují ještě jiným činnostem. Pomalejší vyplňování ale může být způsobeno i podobou dotazníku na malém displeji. Zahraniční zjištění také ukazují, že respondenti odpovídající pomocí mobilního telefonu se svými základními charakteristikami liší od uživatelů ostatních zařízení. Je ovšem nezbytné poznamenat, že tyto rozdíly se v čase pravděpodobně snižují s tím, jak stoupá počet uživatelů chytrých telefonů.

V našich šetřeních jsme výraznější rozdíly mezi jednotlivými zařízeními nezaznamenali. Jako příklad můžeme uvést průměrný počet nejčastěji volené kategorie odpovědi v 14položkové baterii technostresu. Jak shrnuje Graf 2, pro jednotlivá zařízení v rámci CAWI se tato hodnota neliší.

Graf 2. Průměrný počet nejčastěji volené kategorie jako indikátor kvality dat v 14položkové baterii.



Diseminace výsledků a dat

Pokud jsme mód dostatečně zohlednili při analýze, není nezbytně nutné design šetření diskutovat tehdy, kdy prezentujeme výsledky stručnou či popularizační formou. Je naopak vhodné jej zmínit v odborných textech, které by měly též obsahovat detailnější popis logiky konstrukce vzorku, efektu módu, postupu při analýze a případných problémů či přetrvávajících zkreslení. Přesný popis sběru dat a vhodné práce s nimi by měl být také součástí **dokumentace k datům**, pokud data dáváme k dispozici širší vědecké komunitě.

Jako součást **pěče o terén** by mělo být dobrou praxí respondenty nějakým způsobem informovat o výsledcích šetření, kterého se účastnili. Lze tak učinit adresně (rozesílání brožurek či odkazů na výsledky) nebo i nezacíleně, prostřednictvím zveřejnění výsledků širšímu publiku na webových stránkách či sociálních sítích zadavatele nebo agentury, anebo jako novinový článek v online či tištěných médiích. Mysleme přitom na to, že v nastupující informační společnosti bychom (i kvůli budoucímu přesvědčování respondentů k účasti) měli usilovat o co největší využití našich dat, zajištění jejich jednoduché dostupnosti a popularizaci zjištění.

Závěrem

Pokud se v dnešní době nespokojíme s dostupnými daty/informacemi, a rozhodneme se pustit do dobrodružství vlastního sběru dat, je dobré o něm přemýšlet jako procesu propojujícím několik skupin

lidí. V první řadě je důležité proces od začátku plánovat za konzultace s odborníky (z agentury nebo sítě agentur SIMAR, akademické instituce). Dále nezapomínáme na to, abychom se co nejlépe dostali k naší cílové populaci a data dokázali efektivně posbírat. Nakonec jsou to především respondenti, kteří nám věnují čas a informace; k těm také přistupujeme s úctou, a kromě uvedených motivačních prvků vyjádříme vděk i po skončení, nechme prostor si na dotazník „postěžovat“, nabídneme zaslání základních výsledků. Celý proces je často náročný, ale přesto tímto způsobem můžeme získat nenahraditelné, většinou jiným způsobem nedostupné informace.

Seznam literatury

ANDREWS, Dorine, Blair NONNECKE a Jennifer PREECE. 2003. „Electronic Survey Methodology: A Case Study in Reaching Hard-To-Involve Internet Users.“ *International Journal of Human-Computer Interaction* 16(2): 185–210.

BETHLEHEM, Jelke a Silvia BIFFIGNANDI. 2011. *Handbook of web surveys* (Vol. 567). Hoboken, NJ: John Wiley & Sons.

BOSNJAK, Michael, Robert BAUER a Kai W. WEYANDT. 2018. „Mixed Devices in Online Surveys: Prevalence, Determinants, and Consequences.“ Pp. 53-65 in Axel THEOBALD (ed.). *Mobile Research*. Wiesbaden: Springer Gabler.

BUCHTIK, Martin. 2012. „Smíšené módy sběru dat v kvantitativním sociálněvědním výzkumu.“ *Data a výzkum-SDA Info* 6(2): 129–149.

CALLEGARO, Mario, Katja Lozar MANFREDA a Vasja VEHOVAR. 2015. *Web survey methodology*. London: Sage.

CONRAD Frederick G., Mick P. COUPER, Roger TOURANGEAU a Chan ZHANG. 2017. „Reducing Speeding in Web Surveys by Providing Immediate Feedback.“ *Survey Research Methods* 11(1): 45–61.

COOK, Colleen, Fred HEATH a Russel L. THOMPSON. 2000. „A Meta-Analysis of Response Rates in Web-Or Internet-Based Surveys.“ *Educational and Psychological Measurement* 60(6): 821–836.

COUPER, Mick P. 2000. „Web Surveys – A Review of Issues and Approaches.“ *Public Opinion Quarterly* 64(4): 464–494.

COUPER, Mick P. 2017. „New Developments in Survey Data Collection.“ *Annual Review of Sociology* 43: 121–145.

ČSÚ, 2019, Informační společnost v číslech – 2019, <https://www.czso.cz/csu/czso/informacni-spolecnost-v-cislech-2018>; poslední dostupná data

ČSÚ, 2021, Informační společnost v číslech – 2021, <https://www.czso.cz/csu/czso/informacni-spolecnost-v-cislech-2021>

DE LEEUW, Edith. D. a Joop J. HOX. 2011. „Internet Surveys As Part Of A Mixed-Mode Design.“ Pp. 45–76 in Marcel DAS et al. (eds.). *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*. New York: Routledge/ Taylor & Francis Group.

DE LEEUW, Edith D., Z. Tuba SUZER-GURTEKIN a Joop J. HOX. 2018. „The Design and Implementation of Mixed-mode Surveys.“ Pp. 387–408 in Timothy P. JOHNSON, Beth-Ellen PENNELL, Ineke A. L. STOOP a Brita DORER (eds.). *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC)*. New York: John Wiley & Sons.

DE LEEUW, Edith D. a Vera TOEPOEL. 2018. „Mixed-Mode and Mixed-Device Surveys“. Pp. 51-61 in David L. VANNETTE a Jon A. KROSNICK (eds). *The Palgrave Handbook of Survey Research*. Cham: Palgrave Macmillan.

DILLMAN, Don A., Jolene D. SMYTH a Leah M. CHRISTIAN. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Hoboken, NJ: John Wiley & Sons.

EVANS, Joel. R. a Anil MATHUR. 2005. „The Value of Online Surveys.“ *Internet research* 15(2): 195–219.

FAN, Weimiao a Zheng YAN. 2010. „Factors Affecting Response Rates of The Web Survey: A Systematic Review.“ *Computers in Human Behavior* 16(2): 132–139.

HOX, Joop. Edith D. DE LEEUW, E. D. a Thoma KLAUSCH. 2017. „Mixed Mode Research: Issues in Design and Analysis.“ Pp. 511–530 in Paul BIEMER et al. (eds.). *Total Survey Error in Practice: Improving Quality in the Era of Big Data*. New York: Wiley.

HUBATKOVÁ, Barbora a Martin LAKOMÝ. 2021. „Smišený mód a online dotazování: mezinárodní praxe a tuzemská připravenost na implementaci.“ *Sociální studia/Social Studies* 18(1): 13-33.

KREJČÍ, Jindřich. 2009. *Kvalita sociálněvědních výběrových šetření v České republice*. Praha: Sociologické nakladatelství (SLON).

LUGTIG, Peter a Vera TOEPOEL. 2016. „The Use of Pcs, Smartphones, and Tablets in a Probability-Based Panel Survey: Effects on Survey Measurement Error.“ *Social Science Computer Review* 34(1): 78–94.

LYNN, Peter. 2020. „Evaluating Push-To-Web Methodology for Mixed-Mode Surveys Using Address-Based Samples.“ *Survey Research Methods* 14(1): 19–30.

MARTIN, Peter. 2011. „A Good Mix? Mixed Mode Data Collection and Cross-National Surveys.“ *Ask. Research & Methods* 20: 5–26.

PILECKÁ, Jarmila. 2016. *Smišené módy sběru dat: PAPI a CAPI srovnání*. Diplomová práce. Praha: Univerzita Karlova.

RYBÍN, Filip. 2021. *Sampling Fusion: Ekvivalence měření tří módů sběru dat*. Diplomová práce. Praha: Univerzita Karlova.

SIMAR. 2019. *Výroční zpráva o stavu výzkumného trhu*. Cit. 17. dubna 2020 (https://simar.cz/assets/media/files/assets/uploads/Vyrocní_zpráva_o_stavu_vyzkumneho_trhu_2018.pdf).

SMYTH, Jolene D. a Jennie E. PEARSON. 2011. „Internet Survey Methods: A Review of Strengths, Weaknesses, and Innovations.“ Pp. 11–44 in Marcel DAS et al. (eds.). *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*. New York: Routledge/ Taylor & Francis Group.

STRUMINSKAYA, Bella, Kai WEYANDT a Michael BOSNJAK. 2015. „The Effects of Questionnaire Completion Using Mobile Devices on Data Quality. Evidence from a Probability-Based General Population Panel.“ *Methods, Data, Analyses* 9(2): 261–292.

TOEPOEL, Vera a Peter LUGTIG. 2015. “Introduction: Online Surveys are Mixed-Device Surveys. Issues Associated with the Use of Different (Mobile) Devices in Web Surveys.” *Methods, Data, Analyses* 9(2): 155–162.

TOURANGEAU, Roger. 2017. „Mixing Modes: Tradeoffs among Coverage, Nonresponse, and Measurement Error.“ Pp. 115–132 in Paul BIEMER et al. (eds.). *Total Survey Error in Practice: Improving Quality in the Era of Big Data*. New York: Wiley.

VILLAR, Ana a Rory FITZGERALD. 2017. „Using Mixed Modes in Survey Data Research: Results from Six Experiments.“ Pp. 273–310 in Michael J. BREEN (ed.). *Values and Identities in Europe: Evidence from the European Social Survey*. London: Routledge.