

Metodika hodnocení dopadů klimatických kampaní

Jan Urban, Tomáš Chabada, Jan Skalík

Masarykova univerzita, Katedra environmentálních studií

Brno, 2017

T A
Č R

Tato metodika vznikla v rámci projektu TAČR „Nástroje pro hodnocení dopadu kampaní zaměřených na snížení individuálních emisí skleníkových plynů“, č. TD03000282.

Obsah

1. Úvod	3
2. Základní indikátory dopadu klimatických kampaní.....	11
3. Základní způsoby odhadu efektu kampaně a jejich omezení	25
4. Posuzování velikosti efektu kampaní	39
5. Náročnost a kvalita hodnocení dopadu kampaní	49
6. Subjekty provádějící hodnocení kauzálních dopadů kampaní.....	60
Přílohy	67

1. Úvod

1.1 Cíle metodiky

Cílem této metodiky je poskytnout referenční rámec pro hodnocení kauzálních dopadů klimatických kampaní, tedy kampaní usilujících o změnu individuálního chování zaměřeného na mitigaci či adaptaci vůči projevům změny klimatu (dále označovaného zkráceně jako mitigační a adaptační chování), anebo o změnu nějakých faktorů motivace (jako např. postojů, znalostí, nebo představ), které mohou k takové změně chování vést. Tato metodika představuje jednak postupy, které slouží k hodnocení kauzálních dopadů kampaní a dále pak kritéria, která umožňují hodnotit velikost kauzálních dopadů kampaní a správně je interpretovat.

Hodnocení kauzálních dopadů klimatických kampaní je možné provést pouze s využitím kvantitativních experimentálních a kvaziexperimentálních metod. Předkládaná metodika se odlišuje od stávající metodiky pro hodnocení environmentálního vzdělávání (Činčera, 2010, 2013a, 2013b), která se sice také zaměřuje na zkoumání kauzálních dopadů, ale se zaměřením na dopady environmentálního vzdělávání. Předkládaná metodika se také liší od jiných metodik, jejichž cílem je hodnocení jiných aspektů projektů než kauzálních dopadů, jako například hodnocení projektového cyklu (Ministerstvo pro místní rozvoj, 2016; Potulka & Špaček, 2013).

Hodnocení kauzálních dopadů klimatických kampaní je podobné hodnocení dopadů veřejných programů pomocí experimentálních a kvaziexperimentálních postupů (viz např. Chatterji, 2016; Orr, 1998). Hodnocení kauzálních dopadů se však liší od obecně známých a používaných postupů evaluací programů a projektů, které obvykle hodnotí jiné aspekty než kauzální dopady a kromě kvantitativních postupů využívají i jiné postupy hodnocení, které však nejsou pro hodnocení kauzálních dopadů relevantní (aktuální přehled těchto široce pojatých evaluačních metod uvádí např. Stufflebeam & Coryn, 2014).

Hodnocení kauzálních dopadů má v kontextu klimatických kampaní mimořádnou důležitost, protože potenciální přínos dobře fungujících kampaní je velký a velké mohou potenciálně být i finanční náklady spojené s realizací takových kampaní. Předkládaná metodika umožňuje hodnocení dopadů klimatických kampaní a z dlouhodobého hlediska by měla přispět k tomu, aby kampaně využívaly co nejúčinnější prostředky a dosahovaly co největší změny chování.

1.2. Pro jaké typy kampaní je metodika vhodná

Tato metodika představuje celé spektrum různých přístupů k hodnocení kauzálních dopadů klimatických kampaní. To ale neznamená, že všechny tyto přístupy jsou vhodné pro všechny typy kampaní. Metodika stanovuje několik základní principů, které by měly pomoci s výběrem vhodného

postupu hodnocení. Tyto principy jsou obsaženy v jednotlivých kapitolách metodiky; na tomto místě je stručně shrneme a odkážeme na příslušné kapitoly metodiky.

1.2.1 Pravidlo přiměřenosti

Základním pravidlem pro hodnocení kampaně je pravidlo přiměřenosti: finanční a odborná náročnost hodnocení kampaně by měla odpovídat nákladům kampaně a důvodům její realizace.

V metodice (Kapitola 5) představujeme tři úrovně hodnocení kauzálního efektu kampaní a uvádíme orientační informace o finanční a odborné náročnosti těchto přístupů. Obecně platí, že čím přesnější informace může určitý způsob hodnocení poskytnout o kauzálních dopadech kampaní, tím vyšší bude jeho finanční a odborná náročnost. Nyní tyto tři úrovně hodnocení krátce shrneme.

Orientační hodnocení je nejjednodušší a nejméně náročný (z finančního a odborného hlediska) způsob hodnocení dopadů určitého typu kampaní. Toto hodnocení není založeno na vlastním výzkumu, ale využívá přehledy nebo metaanalýzy existujících studií, které zkoumaly kauzální efekty stejných nebo typově podobných kampaní. Toto hodnocení dává odpověď na otázku, jestli stejný (nebo podobný) typ kampaně může vůbec vyvolat kauzální efekty, které jsou v kampani plánovány a jak velké tyto efekty mohou být. Takové hodnocení je orientační a v některých případech ho nebude možné provést (např. pokud neexistují studie, které by typově podobné kampaně zkoumaly). Nicméně i takové orientační hodnocení je užitečné v jakékoli fázi přípravy a realizaci kampaně (byť největší smysl má samozřejmě ve fázi přípravy kampaně), protože umožňuje zvolit optimální strategii pro realizaci kampaně a do budoucna vytváří podmínky pro to, aby byly skutečné dopady takové kampaně lépe prostudovány a prakticky využívány. Řádově se náklady u tohoto typu hodnocení budou pohybovat přibližně v intervalu 20-60 tis. Kč.

Základní hodnocení je vyšší úrovní hodnocení dopadů kampaní a je založeno na vlastním výzkumu, tedy sběru dat o dopadech prováděné kampaně. Základní hodnocení však využívá zjednodušené postupy (laboratorní experimenty, webové experimenty), které dávají odpověď na otázku, jestli konkrétní postupy využití v kampani mají, za ideálních podmínek, očekávané dopady a jak velké tyto dopady jsou. Tento typ hodnocení pravděpodobně bude muset realizovat odborník, který má zkušenosti s využitím kvantitativních výzkumných metod. Náklady tohoto typu hodnocení se budou pohybovat nad 100 tis. Kč.

Pokročilé hodnocení je nejvyšším stupněm hodnocení dopadů kampaní, a ačkoli může používat podobné postupy jako základní hodnocení, bude se od něj lišit tím, že se snaží odpovědět na otázku, jestli realizovaná kampaň skutečně měla zamýšlené kauzální dopady a jak velký dopad kampaně byl. Pokročilé hodnocení bude mít často podobu terénního experimentu, anebo bude realizováno jiným způsobem přímo na cílové skupině kampaně. Pokročilé hodnocení bude mít obvykle

vysokou statistickou a budou obecně průkaznější. Tím pádem budou výsledky takového hodnocení považovány za prokazatelnější a jistější. Náklady na tento typ hodnocení se budou pohybovat obvykle od 200 tis. Kč výše.

U malých kampaní, jejichž náklady na realizaci se pohybují na úrovni několika set tisíc korun, nemá obvykle smysl provádět náročné evaluace, jejichž náklady by dosahovaly podobné výše. U kampaní, jejichž náklady přesáhnou 400 tisíc Kč, je ale vhodné provést alespoň orientační evaluace, které stále mohou přinést důležité informace o dopadech typově podobných kampaní. Vyšší nároky (a tím i zvýšené výdaje) na evaluace malých kampaní ale mohou být oprávněné v případech, kdy by se jednalo o pilotní projekty, jejichž výsledky by rozhodovaly o realizaci větších a nákladnějších kampaní.

Naopak u velkých kampaní, jejichž náklady by se pohybovaly na úrovni vyšší než 4 miliony korun na jednu kampaň, je provedení pokročilých evaluací vhodné. Takové evaluace dávají přesnější informace o dopadech kampaně a jsou důležitým podkladem pro rozhodování o tom, jestli podobné kampaně realizovat v budoucnosti a za jakých podmínek.

Jako základní vodítko pro přiměřenost nákladů hodnocení kampaní doporučujeme rozmezí 5 % - 15 % nákladů samotné kampaně. Jinými slovy, pro kampaň, jejíž náklady se budou pohybovat na úrovni 400 000 Kč, by se očekávané náklady na evaluaci měly pohybovat mezi 20 a 60 tis. Kč. Takové prostředky umožňují maximálně provedení orientační evaluace, která je založena na rešerši dostupných studií a nebude zahrnovat vlastní výzkum. U kampaní s menším rozpočtem je při zachování pravidla přiměřenosti provedení hodnocení kauzálních dopadů obtížné. Naopak u kampaně, jejíž náklady přesáhnou 2 mil. Kč, se přiměřené náklady na evaluaci budou pohybovat nad 100 tis. Kč a měla by již být provedena minimálně základní nebo dokonce pokročilá evaluace dopadů kampaně.

1.2.2 Pravidlo prokazatelnosti kauzálních efektů

Tato metodika je založena na statistickém rámci pro hodnocení kauzálních dopadů a využití standardizovaných kvantitativních metod. Takový rámec je vhodný v případech, kdy může být evaluace postavena na pozorování minimálně desítek nebo spíše stovek jednotlivců a současně mohou být tato pozorování prováděna standardizovaným způsobem. Ve většině případů využívá hodnocení kauzálních dopadů postupy statistické indukce: hodnocení je prováděno na vzorku a výsledky hodnocení se zobecňují na populaci, z níž vzorek pochází. V těchto případech by měly být zajištěny podmínky pro použití statistické indukce (tj. například náhodný výběr vzorku z populace, anebo alespoň výběr, který dovede aproximovat vlastnosti náhodného výběru, např. kvótní výběr).

U menších vzorků recipientů kampaní nebude tato metodika vhodná z toho prostého důvodu, že metody, které využívá (statistickou indukci) budou mít příliš malou statistickou sílu a efekty kampaní nebudou moci prokázat, i kdyby skutečně existovaly. V případech, kdy nebude možné pozorování

standardizovat, bude docházet k podobnému problému: získaná pozorování budou zatížena velkou chybou měření anebo velkou systematickou chybou, což může mít za následek opět to, že evaluace nebude schopna efekt prokázat, i kdyby reálně existoval.

Typickými příklady kampaní, pro které není tato metodika hodnocení vhodná, jsou kampaně zaměřené na různé vlivové skupiny (poslance, podnikatele, skupiny *opinion leaderů*). Takové skupiny budou často velmi malé, anebo nebude technicky možné získat u těchto skupin standardizovaná pozorování (např. odpovědi na dotazník), při zajištění dostatečné reprezentativity. Je důležité poznamenat, že evaluaci kampaní zaměřených na malé skupiny respondentů tedy není možné provádět za pomoci kvantitativních metod popisovaných v této metodice. Tato metodika naopak bude vhodná v případech, kdy kampaně cílí na relativně velkou skupinu příjemců (obyvatelé ČR nebo nějaké oblasti, žáci nebo studenti, určitá socioprofesionální nebo věková skupina apod.).

1.3 Pro koho je metodika určena

Tato metodika je určena primárně jako metodické vodítko pro odborníky, kteří budou provádět hodnocení dopadů kampaní a pro oponenty, kteří mohou kvalitu takových evaluací posuzovat. Sekundárně je tato metodika určena i pro zadavatele hodnocení (což často budou donoři finančních prostředků) i pro realizátory klimatických kampaní. Nyní stručně shrneme, co každá z těchto skupin uživatelů může od metodiky očekávat.

1.3.1 Odborníci provádějící hodnocení dopadů kampaní

Odborníci provádějící hodnocení dopadů kampaní najdou v metodice rámcová pravidla pro provádění hodnocení, nástroje pro provádění evaluací (škály k měření konstruktů), minimální standardy hodnocení, praktické ukázky různých přístupů k hodnocení kampaní (v podobě demonstračních studií uvedených v informačních rámcích) a odkazy na další literaturu. Metodika je psána právě pro tuto skupinu uživatelů a je psána jazykem, který by měl být pro tuto skupinu srozumitelný a současně by měl být jednoznačný.

Metodika by neměla být chápána jako autoritativní dokument, který taxativně stanovuje přijatelné postupy hodnocení, ale spíše jako vodítko pro výběr a aplikaci obecně přijímaných metodologických postupů (např. různých typů experimentálních a kvaziexperimentálních výzkumných uspořádání), které jsou vhodné pro hodnocení kauzálních dopadů kampaní. Rádi bychom hodnotitele vybídli, aby sami dokumentovali prováděné evaluace (např. formou článků nebo zpráv) a sdíleli je s ostatními, nebo navrhovali lepší metodologická řešení, než jaká jsou popsána v této metodice. Takové poznatky budou užitečné při budoucích aktualizacích této metodiky a jako podpora práce ostatních evaluátorů.

1.3.2 Hodnotitelé evaluací

Hodnotitelé evaluací (odborní oponenti evaluačních zpráv) představují druhou skupinu uživatelů, pro něž byla tato metodika připravena. Hodnotitelé evaluací by v této metodice měli především nalézt minimální požadované standardy pro hodnocení kauzálních dopadů kampaní. Pomocí těchto kritérií budou moci posoudit, jestli evaluační studie dosahují potřebné kvality vzhledem k povaze samotných kampaní.

1.3.3 Zadavatelé hodnotících zpráv

Zadavatelé hodnotících zpráv, ať už to jsou donoři, anebo sami realizátoři kampaní, si mohou na základě této metodiky vytvořit rámcovou představu o tom, jaké jsou možné typy evaluací, na jaké otázky odpovídají a jaká je odborná a finanční náročnost je s nimi spojena. Na základě takových informací se mohou zadavatelé hodnotících zpráv rozhodnout, jak přesnou informaci o dopadech kampaní potřebují a tomu přizpůsobit požadavky na kvalitu evaluačních studií. Nároky na kvalitu evaluační studie se následně mohou promítnout do všech fází realizace kampaní od jejich plánování, před jejich návrh, průběh kampaní, až po závěrečné hodnocení výsledků kampaní.

Pro zadavatele hodnotících zpráv by měla být srozumitelná rámcová pravidla a doporučení popsaná v této metodice, ale je pravděpodobné, že některé technické podrobnosti evaluací jim mohou být méně srozumitelné.

1.3.4 Realizátoři kampaní

Realizátoři kampaní v této metodice naleznou rámcová pravidla pro hodnocení kampaní v závislosti na jejich rozsahu. Podstatné je přitom pravidlo přiměřenosti zmiňované výše a podrobněji v Kapitole 5. K jednotlivým velikostem kampaní jsou v metodice uvedeny příklady provedených hodnocení. Tyto informace by měly pomoci realizátorům při plánování kampaně a eventuálně i její evaluace. Realizátoři kampaní by také mohli být schopni na základě metodických doporučení obsažených v této metodice realizovat jednoduché orientační evaluace dopadů kampaně (formou rešerše předchozích evaluačních studií) a měli by být schopni formulovat rámcové zadání pro externí evaluaci v případě, kdy mají být použity náročnější evaluační postupy.

Tato metodika nepředpokládá, že by složitější evaluace prováděli sami realizátoři kampaní, i když i takový postup je možný. Z tohoto důvodu není tato metodika detailním návodem, jak přesně evaluaci kauzálních dopadů provádět, ale je spíše přehledem postupů, které by měly být při hodnocení kauzálních dopadů používány. Pro realizátory kampaní by proto měla být srozumitelná rámcová

pravidla a doporučení popsaná v této metodice, ale je pravděpodobné, že některé technické podrobnosti evaluací pro ně mohou být méně srozumitelné.

1.4. Části metodiky

Tato metodika je organizována do pěti kapitol. Každá kapitola se zabývá samostatným tématem souvisejícím s přípravou a realizací hodnotících studií. V závěru každé kapitoly je uvedeno shrnutí kapitoly a je zde navržena základní rozšiřující literatura a také kompletní seznam textů, které jsou v příslušné kapitole citovány. Kromě jednotlivých kapitol obsahuje metodika i přílohy, které obsahují popis nástrojů využitelných při evaluaci dopadů klimatických kampaní.

Po tomto úvodu následuje druhá kapitola, která popisuje základní indikátory dopadů klimatických kampaní. Pomocí těchto indikátorů je možné měřit to, co má klimatická kampaň dosahovat, ať už je to změna postojů, znalostí, nebo představ o změně klimatu, anebo změna chování souvisejícího s klimatickou změnou, tj. buď mitigačního nebo adaptačního chování. Součástí této kapitoly je i několik příloh, které obsahují používané nástroje pro hodnocení dopadů klimatických kampaní.

Třetí kapitola popisuje různé způsoby odhadu kauzálního efektu, které vycházejí z teoretického rámce potenciálních výsledků. Tento rámec je běžně využíván v aplikované statistice pro odhad kauzálních efektů. Tato kapitola dále popisuje využívání různých experimentálních a kvaziexperimentálních přístupů při zkoumání kauzálního dopadu a výhody a omezení těchto přístupů.

Čtvrtá kapitola pojednává o posuzování naměřeného kauzálního efektu a interpretaci jeho velikosti. Tato kapitola také upozorňuje na některá úskalí hodnocení velikosti efektu, zejména na fakt, že takové hodnocení je vždy nutné interpretovat v kontextu efektů naměřených v podobné oblasti, a také upozorňuje na rozdíly mezi statistickou a praktickou významností určitého efektu.

Pátá kapitola představuje tři úrovně evaluace: orientační, základní a pokročilou evaluaci. Tyto tři úrovně se liší jednak svými nároky (finančními, odbornými a jinými), ale také kvalitou hodnocení kauzálních dopadů. Obecně platí, že vyšší kvalita evaluace je spojena s vyšší náročností evaluace. Jednotlivé úrovně hodnocení jsou přiblíženy na příkladu demonstračních studií, které hodnotily dopady konkrétních kampaní. V této kapitole uvádíme také orientační náklady jednotlivých úrovní evaluace a navrhuje pravidlo přiměřenosti, které stanovuje, že náklady realizace evaluační studie by měly odpovídat 5 % - 15 % nákladů kampaně.

Konečně šestá kapitola se zabývá výběrem realizátorů hodnotících studií. Tato kapitola upozorňuje na fakt, že ve většině případů budou evaluace provádět externisté, jednotlivci nebo organizace, kteří se nepodíleli na realizaci kampaně. Ve většině případů budou evaluátoři buď specializovaní odborní hodnotitelé, anebo specialisté se širším profesním zaměřením, jako např.

kvantitativní výzkumníci. Interní evaluace kampaní, tedy evaluace prováděné samotnými realizátory kampaní, jsou pravděpodobné v případech, kdy rizika vyplývající ze střetu zájmů jsou relativně malá a kdy jsou realizátoři odborně schopní evaluace zpracovat (což bude často případ orientačních evaluací).

Přílohy, které jsou součástí metodiky, dokumentují škály využívané při měření postojů ke klimatické změně, měření znalostí mechanismu klimatické změny a měření mitigačního chování pomocí informací uváděných respondenty nebo pomocí kalkulačky individuální uhlíkové stopy.

Seznam literatury

- Činčera, J. (2010). Metodika evaluace programů environmentální výchovy. *Envigogika*, 5(3).
- Činčera, J. (2013a). Metodika pro hodnocení environmentální výchovy pro dospělé účastníky. *Envigogika*, 8(5). doi:10.14712/18023061.415
- Činčera, J. (2013b). Metodika pro hodnocení environmentální výchovy pro předškolní děti a mladší školní věk. *Envigogika*, 8(5). doi:10.14712/18023061.413
- Chatterji, M. (2016). Causal inferences on the effectiveness of complex social programs: Navigating assumptions, sources of complexity and evaluation design challenges. *Evaluation and Program Planning*, 59, 128–140. doi:10.1016/j.evalprogplan.2016.05.009
- Ministerstvo pro místní rozvoj. (2016). *Metodický pokyn pro evaluace v programovém období 2014 - 2020*. Praha: MMR. Získáno z http://www.dotaceeu.cz/getmedia/9cf7f984-4748-4695-901c-d9b10f12bcd6/MP-evaluace_v4.pdf?ext=.pdf
- Orr, L. L. (1998). *Social Experiments: Evaluating Public Programs With Experimental Methods* (1. vyd.). London, U.K.: Sage Publications.
- Potulka, O., & Špaček, M. (2013). *Postupy a metody kontrafaktuálních dopadových evaluací pro Operační program Zaměstnanost v období 2014 – 2020*. Získáno z https://www.mpsv.cz/files/clanky/17051/Metodika_CIE_MPSV_131015.pdf
- Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, and applications* (Second edition). San Francisco: Jossey-Bass & Pfeiffer Imprints, Wiley.

2. Základní indikátory dopadu klimatických kampaní

Klimatické kampaně mohou mít různé cíle. Mohou se zaměřovat na zvýšení *znalostí* mechanismu klimatické změny a jejích dopadů, anebo na zvýšení znalostí opatření, které pomáhají čelit klimatické změně. Klimatické kampaně se také mohou zaměřovat na změnu *přesvědčení* a *postojů* vztahujících se ke klimatické změně. V neposlední řadě se klimatické kampaně mohou také zaměřovat přímo na změnu *chování*, které ovlivňuje individuální emise skleníkových plynů nebo jinak souvisí s klimatickými mitigacemi a adaptacemi (např. podpora klimatických politik). Pomocí evaluací klimatických kampaní je možné hodnotit, jestli kampaně dosahují těchto stanovených cílů, tedy jestli mají měřitelný kauzální efekt na znalosti, přesvědčení, postoje, nebo chování.

Jevy, které se klimatické kampaně snaží ovlivnit, jsou v odborné literatuře označovány jako konstrukty. Tyto konstrukty není často možné přímo pozorovat, a proto bývají někdy označovány také jako latentní proměnné. I většina typů chování, na něž mohou klimatické kampaně cílit, je obtížně přímo pozorovatelná; často se totiž jedná o abstraktní třídy chování, které se mohou projevovat různými konkrétními způsoby. Například tendence k individuálnímu mitigačnímu chování se může projevovat současně jako tendence k omezování využívání osobního automobilu, tendence k úsporám energie v domácnosti, anebo jako tendence pořizovat si úsporné spotřebiče. Pokud chceme takovou tendenci odhalit, nemůžeme vycházet pouze z pozorování jednoho konkrétního chování (Kaiser, Byrka, & Hartig, 2010). Ačkoli latentní proměnné není možné přímo pozorovat, je možné je měřit prostřednictvím jejich pozorovatelných indikátorů (Hendl & Remr, 2017).

Protože se evaluační studie klimatických kampaní musejí opírat o teorii měření a musejí používat sady indikátorů pro měření latentních konstruktů, představíme ve zbytku této kapitoly teorii měření a uvedeme příklady indikátorů používaných pro měření některých vybraných konstruktů vhodných pro hodnocení klimatických kampaní. Podrobně jsou tyto indikátory popsány v příloze k této kapitole. Pro hodnocení dopadů realizovaných klimatických programů je možné použít i jiné indikátory, za předpokladu, že mají dostatečnou validitu a reliabilitu.

2.1 Měření latentních konstruktů

Měření představuje podstatu empirického výzkumu. Základní myšlenkou jakékoli teorie měření je představa, že je možné kvantitativně popsat intenzitu určitého jevu.¹ Např. postoj může mít různé

¹ Na tomto místě bychom měli upozornit na fakt, že v sociálních vědách existují různé teorie měření. Tzv. klasická teorie měření vychází z představy, že měření je proces přiřazování číselných hodnot různým úrovním určitého jevu na základě arbitrárního pravidla (Stevens, 1946). Typické je pro klasickou teorii měření

úrovně, které je možné vyjádřit číselně. Pokud tedy měříme úroveň postoje určitého člověka, znamená to, že jsme pomocí jednoho čísla schopni vyjádřit intenzitu tohoto postoje (viz např. (Fishbein & Ajzen, 2010; Kaiser et al., 2010).

Chceme-li nějaký jev změřit, musíme vědět, jak je možné zjistit různé úrovně tohoto jevu (jeho intenzity). Protože latentní konstrukty nemůžeme pozorovat přímo, ale pouze prostřednictvím jejich indikátorů, je důležitým krokem měření tzv. operacionalizace, tedy převod abstraktních konstruktů do sady měřitelných indikátorů (Mareš, Rabušic, & Soukup, 2015). Operacionalizací rozumíme postup, který vede k vytvoření empirické sady indikátorů umožňujících měřit rozdíly v intenzitě určitého jevu. Indikátory jsou přímo zjistitelné údaje vypovídající o intenzitě určitého jevu, zatímco konstrukty jsou teoretické entity (Hendl & Remr, 2017). Při měření latentních proměnných jsou typicky používány škály, což jsou stupnice, jejichž hodnoty jednoznačně odpovídají úrovním určité vlastnosti. Škály mají různé formáty a bývají určeny svým měřítkem a způsobem uspořádání svých hodnot.²

Pokud jsou v rámci evaluace měřeny dopady kampaně na latentní konstrukty, je vhodné využívat již existující operacionalizace daného konstruktu (např. existující škály pro měření postojů, znalostí apod.). Pokud žádný takový nástroj neexistuje, je nutné vytvořit nový nástroj měření, což zpravidla zahrnuje časově náročný proces návrhu indikátorů a jejich opakovaného testování a kalibraci (Hendl & Remr, 2017; Wilson, 2005).

2.1.1. Kvalita měření

přesvědčení, že číselné hodnoty vyjadřující intenzitu daného jevu neodpovídají nějakým objektivním rozdílům v intenzitě tohoto jevu, ale jsou dány na základě pravidla měření (Hendl & Remr, 2017). Oproti tomu některé jiné teorie měření předpokládají možnost tzv. objektivního měření (Bond & Fox, 2007). Pro tento typ měření je naopak typický předpoklad, že naměřené hodnoty odpovídají skutečně existujícím rozdílům v intenzitě měřeného jevu a jsou nezávislé na situaci měření (tj. např. na vzorku participantů, na nichž je měření prováděno nebo instrumentu, s nímž je měření prováděno).

² Zde je dobré alespoň stručně zmínit fakt, že existují dva typy indikátorů latentních proměnných. Zatímco reflektivní indikátory jsou indikátory reflektující nějaký reálně existující jev, formativní indikátory jsou indikátory zachycující různé jevy, které společně vyvolávají nějaký efekt. Příkladem reflektivních indikátorů mohou být indikátory používané k měření postoje k životnímu prostředí (Kaiser, Byrka, & Hartig, 2010). U reflexivních indikátorů se předpokládá, že úroveň latentního postoje ovlivňuje to, jak lidé na jednotlivé položky reagují. Naopak příkladem formativního indikátoru jsou např. některé indexy kvality života. Tyto indexy zachycují různé aspekty života člověka, které společně přispívají ke kvalitě života (např. Better life index; Durand, 2015). Rozdíl mezi oběma typy indikátorů se týká směru kauzality: zatímco reflexivní indikátory jsou ovlivňovány měřenou latentní proměnnou, formativní indikátory naopak ovlivňují tuto latentní proměnnou.

Použitý způsob měření by měl být relevantní, objektivní, validní a reliabilní. Relevance spočívá ve vhodnosti použité procedury měření ve vztahu k problému. Objektivita měřícího nástroje znamená, že výsledek není závislý na tom, kdo měřící nástroj použije. Validita je shoda měřeného jevu s vlastností, kterou chceme měřit. Měření je validní, pokud měří to, co by mělo teoreticky měřit. Reliabilita popisuje spolehlivost výsledků měření, tedy to, do jaké míry odpovídají rozdíly v naměřených hodnotách skutečným rozdílům v intenzitě měřených jevů (Mareš et al., 2015).

Různé způsoby zjišťování reliability shrnuje Tabulka 2.1. Každý z těchto postupů zjišťuje jiný aspekt reliability. V ideálním případě by bylo vhodné použít různé postupy a ověřovat různé aspekty reliability, avšak kvůli finančním, časovým a jiným omezením to často není možné. Navíc některé nástroje měření byly již dříve testovány a jejich reliabilita je známá, a proto je není nutné v dalších aplikacích již tak rigorózně testovat.

Tabulka 2.1: Přehled různých způsobů zjišťování reliability.

Aspekty reliability.	Charakteristika.	Způsob měření.	Způsob vyhodnocení.
Test-retest	Shoda opakovaných testů.	Každý participant je měřen opakovaně stejným nástrojem.	Reliabilita je úměrná korelaci mezi výsledky jednotlivých testů provedených v různých časových okamžicích.
Dělení testu	Shoda mezi dvěma různými částmi téhož nástroje.	Každý participant je měřen pouze jednou, a to každou z obou částí nástroje.	Reliabilita je úměrná korelaci mezi výsledky získanými pomocí jednotlivých částí nástroje.
Reliabilita mezi hodnotiteli	Míra shody mezi hodnocením různých soudců.	Různí soudci hodnotí nějaký objekt (často např. obsahovou validitu nějaké škály).	Reliabilita je úměrná míře shody mezi soudci. Míra shody je měřena pomocí speciálních statistik, z nichž některé

			vycházejí z korelačních koeficientů.
Interní konzistence	Velikost asociace mezi jednotlivými položkami měřícími stejný konstrukt.	Konstrukt je měřen pomocí baterie obsahující více položek. Každý účastník odpovídá na baterii pouze jednou.	Interní konzistence je ideálně mírou toho, jaká by byla korelace mezi různými testy měřícími týž konstrukt. Tato míra je ve skutečnosti dolní hranicí korelace mezi testy.
Paralelní formy	Míra shody mezi různými nástroji měřícími týž konstrukt.	Jsou porovnávány výsledky měření daného konstruktu pomocí nástroje, jehož reliabilitu chceme ověřit a pomocí dalších nástrojů, jejichž reliabilita je známá, anebo očekávaná. Každý nástroj se použije jednou. Každý participant je měřen všemi různými nástroji.	Reliabilita je měřena jako asociace mezi výsledky testů získanými různými nástroji.

(Upraveno podle: Hendl & Remr, 2017)

Stejně jako existují různé způsoby ověřování reliability, existují i různé způsoby ověřování validity. Validita je nezávislá na reliabilitě: i nástroj s vysokou reliabilitou může být málo validní. Jinými slovy, takový nástroj může měřit přesně, ale přesto může měřit něco jiného, než co má měřit. Takovou situaci můžeme popsat i tak, že získaná měření jsou zatížena systematickou chybou, která způsobuje, že nástroj systematicky měří něco jiného, než co by měl měřit.³

³ Oproti tomu nízká reliabilita znamená, že měření jsou zatížena větší náhodnou chybou, která se projevuje jako chyba měření.

Validizace je dlouhodobým a zpravidla postupným procesem ověřování validity nástroje, při němž se postupně zvyšuje naše důvěra ve výsledky měření (Hendl & Remr, 2017). Základní způsoby validizace měřicího nástroje shrnuje Tabulka 2.2.

Tabulka 2.2: Charakteristiky vybraných typů validity měřicího nástroje a možnosti přezkoušení.

Typ validity.	Charakteristika.	Způsob měření.	Způsob vyhodnocení.
Obsahová validita	Obsahová validita odpovídá tomu, do jaké míry nástroj měření reprezentuje všechny aspekty měřeného konstrukt.	Panel soudců (expertů v dané oblasti) určuje, jestli indikátory obsažené v daném nástroji zachycují dostatečně konstrukt, který má tento nástroj měřit.	Na základě vyjádření expertů, resp. shody v jejich hodnocení, určujeme, zda náš nástroj adekvátně zachycuje měřený konstrukt.
Kriteriální validita - souběžná validita	Míra shody nástroje měření s měřeními provedenými pomocí jiných validních nástrojů měření téhož konstrukt ve stejném čase.	Ve stejném čase měříme konstrukt dalšími validními měřicími nástroji.	Posouzení korelace nástroje měření s dalšími validními nástroji.
Kriteriální validita - prediktivní validita	Schopnost nástroje měření předpovídat výsledky budoucích měření stejného konstrukt.	Nejdříve je provedeno měření pomocí nástroje. V pozdější době se shromažďují výsledky dalších měření	Posouzení korelace nástroje měření s výsledky měření v pozdější době.
Konstruktová validita - konvergentní validita	Konstruktová validita označuje, do jaké míry	Různé postupy s ohledem na využívané	Sleduje se těsnost vztahu mezi mírou

	daný nástroj měří příslušný teoreticky definovaný konstrukt. Konvergentní validita označuje míru, do které dvě měření konstruktů (které mají spolu teoreticky souviset) spolu fakticky souvisí.	statistické procedury.	konstrukt, který testujeme a mírami jiných konstruktů, které by s ním měli teoreticky souviset. Tento vztah by měl být dostatečně těsný.
Konstruktová validita - diskriminační validita	Konstruktová validita označuje, do jaké míry daný nástroj měří příslušný teoreticky definovaný konstrukt. Diskriminační validita označuje, do jaké míry daný nástroj neměří (diskriminuje) ostatní konstrukty.	Různé postupy s ohledem na využívané statistické procedury	Sleduje se těsnost vztahu mezi mírou konstrukt, který testujeme a mírami jiných konstruktů, které by s ním teoreticky neměli souviset. Tento vztah by měl být těsný.

(Upraveno podle: Hendl & Remr, 2017)

2.2 Hlavní indikátory relevantní pro klimatické kampaně

Indikátory relevantní pro měření dopadů klimatických kampaní můžeme rozdělit do dvou širších skupin. První skupinu tvoří indikátory, které měří individuální chování relevantní z hlediska klimatické změny (např. Individuální mitigační a adaptační chování, záměr takového chování, anebo podporu a vnímání přijatelnosti politik týkajících se klimatické změny). Druhou skupinu pak tvoří indikátory měřící konstrukty, které buď ovlivňují, anebo souvisí s konstrukty spadajícími do první skupiny. Mezi konstrukty spadající do druhé skupiny bychom zařadili například znalosti, přesvědčení a postoje související se změnou klimatu. Řada kampaní přitom cílí na změnu klimatického chování právě prostřednictvím změny těchto konstruktů (např. vzdělávací kampaně často usilují o změnu znalostí a postojů). U takových kampaní má logicky smysl sledovat, jestli jsou úspěšné při ovlivňování těchto

konstruktů spadajících do druhé skupiny, a to i v případě, že se jim nepodaří ovlivnit samotné chování. Nyní si tyto dvě skupiny indikátorů přiblížíme.

2.2.1 Behaviorální indikátory: mitigační chování, přijatelnost a podpora politik

Behaviorální indikátory měří nejen chování jednotlivců v soukromé oblasti (například dopravní chování, spotřebu energií, nakupování, či nakládání s odpady), ale také chování ve veřejné oblasti (např. podpis peticí, podporu zájmových relevantních organizací, či účast na kampaních a protestech), nebo přijatelnost klimatických politik a jejich podporu (Dunlap and Brulle, 2015; Hornsey, Harris, Bain, & Fielding, 2016a).

Metaanalýza dosavadních výzkumů ukázala, že se zvyšujícím se přesvědčením o existenci klimatické změny roste u jednotlivců jak záměr k mitigačnímu a adaptačnímu chování, tak i pravděpodobnost, že se takové chování uskuteční (Hornsey et al., 2016a). Ukazuje se však, že přesvědčení mají větší souvislost se záměrem chování, než se skutečným chováním. Kampaně, které jsou úspěšné při změně záměru chování, proto nemusí nutně dosáhnout změny skutečného chování (Hornsey et al., 2016a). Individuální chování se totiž uskutečňuje v kontextu sociálních příležitostí a omezení zahrnujících např. dostupnost a cenu nízkouhlíkových alternativ, sociální normy, hodnoty a znalosti. Tyto faktory způsobují, že některé záměry chování nejsou ve skutečnosti realizovány (Dunlap and Brulle, 2015).

Obecně se ukazuje, že přesvědčení o klimatické změně pozitivně souvisí také s ochotou upřednostňovat zájmy životního prostředí před ekonomickými zájmy (Hornsey et al., 2016a). Přesvědčení o existenci klimatické změny má pozitivní vliv také na podporu konkrétních veřejných politik (např. podporu alternativních zdrojů energie, podporu uhlíkové daně, podporu obchodování s emisními povolenkami, či podporu vytváření zelených politik v rámci organizací) (Hornsey et al., 2016a). Celkově se ukazuje, že vztah mezi přesvědčením o klimatické změně a podporou klimatických politik je tím silnější, čím obecnější jsou tyto politiky (Hornsey et al., 2016a).

2.2.2 Indikátory faktorů, které mají vliv na chování

Některé klimatické kampaně se nezaměřují na změnu samotného chování, ale usilují o změnu určitých faktorů chování, jako jsou znalosti mechanismu klimatické změny a jejích dopadů nebo přesvědčení o existenci klimatické změny a jejích dopadech. Výsledky dosavadních výzkumů ukázaly, že s rostoucími znalostmi (ať už zjištěnými na základě sebehodnocení nebo objektivními testy) se zvyšuje přesvědčení o existenci klimatické změny (Hornsey et al., 2016a). Například zvýšení znalosti o mechanismu klimatické změny má pozitivní vliv na přesvědčení o existenci klimatické změny a na přesvědčení o

antropogenních příčinách klimatické změny (Aitken, Chapman, & McClure, 2011; Ranney & Clark, 2016b).

Mezi základní přesvědčení související s klimatickou změnou se řadí přesvědčení o existenci klimatické změny. Může se jednat o obecné přesvědčení o tom, jestli klimatická změna probíhá, nebo přesvědčení o antropogenních příčinách klimatické změny. Oba typy přesvědčení mají pozitivní vliv na záměr mitigačního chování a skutečné mitigační chování (Hornsey et al., 2016a).

Jiným důležitým typem přesvědčení, které má vliv na chování, je přesvědčení o míře vědeckého konsensu o různých aspektech klimatické změny a o jejích příčinách. Významná role vědeckého konsensu plyne z toho, že jednotlivci mají velmi omezené možnosti poznat příčiny, mechanismus a dopady klimatické změny na základě vlastní zkušenosti, a tudíž je vědecké poznání tohoto problému prakticky jediným zdrojem informací o tomto jevu. Představy lidí o povaze vědeckého poznání klimatické změny proto významně ovlivňují, co si o tomto problému myslí a jak se chovají. Přesvědčení o míře vědeckého konsensu má pozitivní vliv například na přesvědčení o antropogenních příčinách klimatické změny, vnímání závažnosti a reálnosti klimatické změny, na podporu klimatických politik, a také na záměr mitigačního chování (Bolsen, Leeper, & Shapiro, 2013; Ding, Maibach, Zhao, Roser-Renouf, & Leiserowitz, 2011; Linden, Leiserowitz, Feinberg, & Maibach, 2015; McCright, Dunlap, & Xiao, 2013, 2014).

Význam představ o vědeckém konsensu plyne také z faktu, že veřejnost vnímá míru vědeckého konsensu o různých aspektech klimatické změny zkresleně. Zatímco výrazná většina vědců (97% nebo více; viz Carlton, Perry-Hill, Huber, & Prokopy, 2015; Cook et al., 2013, 2016) se shoduje na antropogenních příčinách klimatické, široká veřejnost míru této shody významně podhodnocuje (Cook et al., 2016). Představy o vědeckém konsensu jsou tedy potenciálně zajímavým cílem klimatických kampaní.

Přesvědčení o lokálních a globálních dopadech klimatické změny spolu souvisí a oba typy mají pozitivní vliv na podporu klimatických politik (Dietz, Dan, & Shwom, 2007; Ding et al., 2011). Pozitivní vliv na podporu klimatických politik má také přesvědčení o časové blízkosti dopadů klimatické změny (Ding et al., 2011; McCright et al., 2013, 2014). Také míra znepokojení klimatickou změnou a vnímání rizik a hrozeb klimatické změny mají pozitivní vliv na podporu mitigačních politik, záměr mitigačního chování a samotné mitigační chování (Aitken et al., 2011; Linden et al., 2015; McCright et al., 2013, 2014; Spence, Poortinga, & Pidgeon, 2012).

2.3. Nástroje měření

Nyní představíme některé nástroje (škály vhodné např. do dotazníků), které se uplatní v evaluačních studiích zaměřených na hodnocení dopadů klimatických kampaní. Pomocí těchto nástrojů je možným

měřit změnu znalostí, přesvědčení, postojů a mitigačního chování jednotlivců v důsledku klimatických kampaní.

Nástroje pro měření znalostí klimatické změny (viz Příloha 1) využívají buď objektivní testy znalostí (Ranney & Clark, 2016b), nebo subjektivní hodnocení vlastních znalostí (Boyes, Skamp, & Stanisstreet, 2009). I když objektivní znalosti a subjektivní percepce vlastních znalostí spolu pozitivně korelují, objektivní znalosti mají obvykle těsnější vztah s mitigačním chováním a přesvědčeními o klimatické změně (Hornsey et al., 2016a). Kampaně navíc přirozeně usilují o zvýšení úrovně objektivních znalostí, proto jsou škály objektivních znalostí vhodným a důležitým nástrojem při evaluaci dopadů klimatických kampaní. V Příloze 1.2.1 uvádíme příklad kalibrované škály znalostí mechanismu klimatické změny obsahující 32 položek.

Existuje několik nástrojů pro měření vnímání vědeckého konsensu (viz Příloha 1.3). Jeden z těchto nástrojů je postaven na zjišťování vnímané shody vědců o antropogenních příčinách klimatické změny u široké veřejnosti (Linden, Leiserowitz, Feinberg, & Maibach, 2014). Jiný nástroj je postaven na zjišťování souhlasu široké veřejnosti s výroky popisujícími míru shody v pohledu vědecké komunity na změnu klimatu z nástrojů (Ding et al., 2011; Leiserowitz et al., 2014).

K měření postojů týkajících se klimatické změny je možné využít například vícepoložkovou škálu měřící skepticismus k různým aspektům klimatické změny (viz Příloha 2.1; viz také Brügger, Morton, & Dessai, 2016). K měření přesvědčení o existenci klimatické změny (viz Příloha 3.1) se používají jednopoložkové škály, která hodnotí představy o existenci klimatické změny (Leiserowitz et al., 2014; Poortinga, Spence, Whitmarsh, Capstick, & Pidgeon, 2011) nebo sílu přesvědčení o realitě klimatické změny (Linden et al., 2014).

Přesvědčení o antropogenních příčinách klimatické změny (viz Příloha 3.2) je možné měřit pomocí jednopoložkové škály, která zjišťuje, do jaké míry lidé připisují příčinu klimatické změny přirozeným změnám prostředí, nebo naopak působení lidí (Bolsen et al., 2013; Linden et al., 2014), či otázkou zjišťující vnímaný podíl lidské činnosti na změně klimatu (Aitken et al., 2011). Přesvědčení o načasování dopadů klimatické změny (viz Příloha 3.3) je možné zjišťovat pomocí jednopoložkové škály, na níž participanti určují, do jaké míry si myslí, že se dopady globálního oteplování projevují již nyní, anebo k nim naopak nikdy nedojde (McCright et al., 2013, 2014). Podobná škála se užívá také k měření představ o tom, kdy začnou tyto dopady působit na lidi (Ding et al., 2011).

Přesvědčení o dopadech klimatické změny (viz Příloha 3.4) je možné měřit pomocí škál, které zjišťují sílu přesvědčení o budoucích negativních dopadech klimatické změny na lidskou populaci a přírodu. První z těchto škál zjišťuje představy o tom, do jaké míry změna klimatu poškodí různé entity, jako například přírodu, lidstvo, nebo participanta osobně (Ashworth, Jeanneret, Gardner, & Shaw, 2011; Ding et al., 2011; Leiserowitz et al., 2014; Zhao, Leiserowitz, Maibach, & Roser-Renouf, 2011).

Druhá škála zjišťuje vnímanou pravděpodobnosti toho, že se konkrétní následky klimatické změny projeví na lokální nebo globální úrovni (Dietz et al., 2007). Další možností je měřit výhradně přesvědčení o dopadech klimatické změny na vlastní osobu (Hornsey et al., 2016a). Znepokojení klimatickou změnou (viz Příloha 4) je možné zjišťovat pomocí jednopoložkové čtyřnádové nebo jedenáctibodové škály škály, která zjišťuje míru obav z klimatické změny (Linden et al., 2014; McCright et al., 2013, 2014).

Záměr mitigačního chování (viz Příloha 5) je možné na obecné úrovni měřit zjišťováním ochoty podniknout osobní kroky ke snížení vlastních emisí skleníkových plynů (Bolsen et al., 2013). Jinou možností je využití škály zachycující záměr realizovat konkrétní mitigační chování (Brügger et al., 2016), která může být případně doplněna také o škálu zjišťující záměr realizovat konkrétní adaptační chování (Brügger et al., 2016).

Při měření podpory politik (viz Příloha 6) je potřebné, aby posuzované politiky reflektovaly současný rámec politik v oblasti změn klimatu a byly zasazené do vhodného kontextu (např. politiky na úrovni EU, nebo ČR). Z tohoto důvodu jsme vytvořili vlastní sadu 12 otázek optimalizovanou pro Českou republiku, která vychází z předešlých škál využívaných v zahraničí (Ding et al., 2011; Leiserowitz et al., 2014). Kromě toho přílohou metodiky uvádíme také škálu obsahující obecně formulované mitigační politiky (Brügger et al., 2016) a škálu zaměřenou na podporu adaptačních politik (Brügger et al., 2016).

Individuální mitigační chování (viz Příloha 7) je možné měřit například pomocí volně přístupného softwaru, Kalkulačky uhlíkové stopy (dostupná na internetové adrese www.uhlikovastopa.cz). Tento nástroj odhaduje na základě informací o chování přímé emise skleníkových plynů v oblasti bydlení, dopravy, spotřeby potravin a další spotřeby. Tato kalkulačka může být snadno integrována do jiných měřících nástrojů (např. dotazníků) a oproti jiným nástrojům má výhodu transparentního algoritmu (popis metodiky kalkulačky je součástí její dokumentace) a dokumentované validity a reliability.

2.4. Shrnutí kapitoly

V této kapitole jsme představili teorii měření latentních konstruktů, jakými jsou například postoje, znalosti nebo třídy mitigačního chování, a která proto nalezne uplatnění při hodnocení dopadů klimatických kampaní. Aby bylo možné tyto latentní konstrukty měřit, je nutné využít proces operacionalizace, který vede k návrhu měřitelných indikátorů. Dále jsme představili hlavní aspekty kvality měření, mezi které patří jeho reliability a validita. Tato kapitola také představila hlavní indikátory využitelné při hodnocení dopadů klimatických kampaní, jako jsou indikátory postojů, představ a znalostí spojených s různými aspekty klimatické změny a indikátory individuálního chování,

kteřé souvisí s klimatickou změnou. Na závěr kapitoly jsme uvedli příklady konkrétních škál používaných pro měření těchto konstruktů.

Rozšiřující četba

Coolican, H. (2014). *Research methods and statistics in psychology* (6. vyd.). London: Psychology Press.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory* (1. vyd.). Mahwah, United States: Psychology Press.

Hendl, J., & Remr, J. (2017). *Metody výzkumu a evaluace* (1. vyd.). Praha: Portál.

Seznam literatury

Aitken, C., Chapman, R., & McClure, J. (2011). Climate change, powerlessness and the commons dilemma: Assessing New Zealanders' preparedness to act. *Global Environmental Change*, 21(2), 752–760. doi:10.1016/j.gloenvcha.2011.01.002

Ashworth, P., Jeanneret, T., Gardner, J., & Shaw, H. (2011). *Communication and climate change: What the Australian public thinks*. Canberra: Csiro publishing.

Bolsen, T., Leeper, T. J., & Shapiro, M. A. (2013). Doing what others do: Norms, science, and collective action on global warming. *American Politics Research*, 42(1), 65–89. doi:10.1177/1532673X13484173

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: fundamental measurement in the human sciences* (2. vyd.). Mahwah, N.J: Lawrence Erlbaum Associates.

Boyes, E., Skamp, K., & Stanisstreet, M. (2009). Australian secondary students' views about global warming: Beliefs about actions, and willingness to act. *Research in Science Education*, 39(5), 661–680. doi:10.1007/s11165-008-9098-5

Brügger, A., Morton, T. A., & Dessai, S. (2016). "Proximising" climate change reconsidered: A construal level theory perspective. *Journal of Environmental Psychology*, 46, 125–142. doi:10.1016/j.jenvp.2016.04.004

Carlton, J. S., Perry-Hill, R., Huber, M., & Prokopy, L. S. (2015). The climate change consensus extends beyond climate scientists. *Environmental Research Letters*, 10(9), 094025. doi:10.1088/1748-9326/10/9/094025

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., ... Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, 8(2), 024024. doi:10.1088/1748-9326/8/2/024024

Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R. L., Verheggen, B., Maibach, E. W., ... Rice, K. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, 11(4), 048002. doi:10.1088/1748-9326/11/4/048002

- Dietz, T., Dan, A., & Shwom, R. (2007). Support for climate change policy: Social psychological and social structural influences. *Rural Sociology*, 72(2), 185–214.
doi:10.1526/003601107781170026
- Ding, D., Maibach, E. W., Zhao, X., Roser-Renouf, C., & Leiserowitz, A. (2011). Support for climate policy and societal action are linked to perceptions about scientific agreement. *Nature Climate Change*, 1(9), 462–466. doi:10.1038/NCLIMATE1295
- Dunlap, R. E. & Brulle, R. J. 2015. *Climate change and society: sociological perspectives*. New York: Oxford University Press.
- Durand, M. (2015). The OECD Better Life Initiative: How's Life? and the Measurement of Well-Being. *Review of Income and Wealth*, 61(1), 4–17. doi:10.1111/roiw.12156
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: the reasoned action approach* (Roč. 1). New York: Psychology Press.
- Hendl, J., & Remr, J. (2017). *Metody výzkumu a evaluace* (1. vyd.). Praha: Portál.
- Hornsey, M. J., Harris, E. A., Bain, P. G., & Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change*, 6(6), 622–626.
doi:10.1038/NCLIMATE2943
- Kaiser, F. G., Byrka, K., & Hartig, T. (2010). Reviving Campbell's paradigm for attitude research. *Personality and Social Psychology Review*, 14(4), 351–367. doi:10.1177/1088868310366452
- Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinberg, G., Rosenthal, S., & Marlon, J. (2014). *Climate change in the American mind: Americans' global warming beliefs and attitudes in November, 2013*. New Haven, CT: Yale Project on Climate Change.
- Linden, S. L. van der, Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2014). How to communicate the scientific consensus on climate change: plain facts, pie charts or metaphors? *Climatic Change*, 126(1–2), 255–262. doi:10.1007/s10584-014-1190-4
- Linden, S. L. van der, Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2015). The Scientific Consensus on Climate Change as a Gateway Belief: Experimental Evidence. *PLOS ONE*, 10(2). doi:10.1371/journal.pone.0118489
- Mareš, P., Rabušic, L., & Soukup, P. (2015). *Analýza sociálněvědních dat (nejen) v SPSS* (1. vyd.). Brno: Masarykova univerzita.
- McCright, A. M., Dunlap, R. E., & Xiao, C. (2013). Perceived scientific agreement and support for government action on climate change in the USA. *Climatic Change*, 119(2), 511–518.
doi:10.1007/s10584-013-0704-9
- McCright, A. M., Dunlap, R. E., & Xiao, C. (2014). Increasing influence of party identification on perceived scientific agreement and support for government action on climate change in the

- United States, 2006–12. *Weather, Climate, and Society*, 6(2), 194–201. doi:10.1175/WCAS-D-13-00058.1
- Poortinga, W., Spence, A., Whitmarsh, L., Capstick, S., & Pidgeon, N. F. (2011). Uncertain climate: An investigation into public scepticism about anthropogenic climate change. *Global environmental change*, 21(3), 1015–1024. doi:10.1016/j.gloenvcha.2011.03.001
- Ranney, M. A., & Clark, D. (2016). Climate change conceptual change: Scientific information can transform attitudes. *Topics in Cognitive Science*, 8(1), 49–75. doi:10.1111/tops.12187
- Spence, A., Poortinga, W., & Pidgeon, N. (2012). The psychological distance of climate change. *Risk Analysis*, 32(6), 957–972. doi:10.1111/j.1539-6924.2011.01695.x
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 677–680.
- Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Zhao, X., Leiserowitz, A. A., Maibach, E. W., & Roser-Renouf, C. (2011). Attention to science/environment news positively predicts and attention to political news negatively predicts global warming risk perceptions and policy support. *Journal of Communication*, 61(4), 713–731. doi:10.1111/j.1460-2466.2011.01563.x

3. Základní způsoby odhadu efektu kampaně a jejich omezení

Tato kapitola představuje teoretický rámec pro analýzu kauzálních efektů, známý jako Neyman–Rubinův model potenciálních výsledků. Dále tato kapitola představuje základní výzkumná uspořádání vhodná pro analýzu kauzálních efektů (tj. experimentální a kvaziexperimentální uspořádání) a zmiňuje jejich výhody, nevýhody a možnosti jejich využití.

3.1 Problém odhadu efektů

Klíčovým pojmem při tvorbě a hodnocení kampaní je pojem efekt kampaně. Ačkoli je tento pojem běžně užíván odbornou i širokou veřejností přicházející do styku s kampaněmi, jedná se o pojem, který je často užíván chybně. Pod pojmem efekt kampaně chápeme v této metodice kauzální efekt, který kampaň má nebo může mít na určitou skupinu lidí. Patrně nejrozšířenějším konceptuálním rámcem pro analýzu takového kauzálního efektu je tzv. Rubinův nebo také Neyman–Rubinův model potenciálních výsledků (Holland, 1986; D. Rubin, 1974; D. B. Rubin, 2005; pro přehled viz Sekhon, 2007).

Rubinův kauzální model je postaven na pojmu potenciálních výsledků, který je velmi užitečný pro pochopení toho, co je to kauzální efekt. Kauzální efekt je v tomto modelu definován jako rozdíl mezi alternativními potenciálními stavy, které mohou nastat v důsledku toho, že změníme úroveň kauzálního faktoru (například poskytneme určitým lidem informace o klimatické změně) a tato změna vyvolá kauzální efekt (tito lidé začnou snižovat své individuální emise skleníkových plynů). Kauzální efekt je tedy rozdíl mezi tím, když daný kauzální faktor působí (lidé dostanou informace) a když nepůsobí (lidé nedostanou informace), resp. rozdíl mezi následky těchto dvou stavů.

Ačkoli je tato definice jednoduchá, ukazuje na důležité problémy kauzálního usuzování. Jedním z těchto problémů je tzv. fundamentální problém kauzálního usuzování (Holland, 1986; D. Rubin, 1974), který spočívá v tom, že nemůžeme konkrétního jedince (nebo skupinu individuí) pozorovat současně ve dvou různých výsledných stavech. Pokud se vrátíme k výše zmíněnému příkladu kauzálního vlivu informace na chování, je evidentní, že nemůžeme téhož člověka současně pozorovat ve stavu, kdy určitou informaci má a kdy ji nemá. Protože jedinec (nebo skupina jedinců) nemůže současně být a nebýt vystaven působení kauzálního faktoru, jedná se o fundamentální problém, který je principiálně neřešitelný. Statistické a jiné přístupy se snaží tento problém přeformulovat tak, aby byl řešitelný, a to zavedením dodatečných předpokladů. Musíme tedy mít na paměti, že všechna tato řešení jsou platná jen tehdy, pokud tyto předpoklady (z nichž některé jsou netestovatelné) platí. Jedním takovým řešením například je, že určitého člověka můžeme pozorovat před tím, než informaci získá a potom, kdy tuto informaci získal. Rozdíl mezi oběma stavy je ale mírou kauzálního efektu pouze

za určitých, poměrně silných, předpokladů (např. mezi oběma pozorováními nedochází k působení dalších faktorů). Různé přístupy k odhadu kauzálního efektu, které představíme v této kapitole, se liší zejména v tom, jak silné předpoklady musejí při odhadu kauzálního efektu přijmout a tedy tím, jak závislé jsou na platnosti těchto předpokladů.

Jednou z důležitých podmínek, která musí platit, abychom mohli vůbec hovořit o kauzálním efektu je tzv. podmínka vystavitelnosti (Rubin, 1974). Podmínka vystavitelnosti stanovuje, že o kauzálním efektu můžeme hovořit pouze u takových faktorů, které se u daného individua mohou, alespoň teoreticky, měnit nezávisle na tomto individuu. Podmínka vystavitelnosti v podstatě znamená, že kauzálním faktorem může být pouze takový faktor, u něž si dovedeme představit hypotetický experiment, který by měnil úroveň tohoto faktoru nezávisle na individuu. Ne každý faktor takovou podmínku, a tudíž ne všechny statistické efekty je možné interpretovat jako kauzální efekty. Příkladem faktoru, který je ze statistického hlediska důležitým prediktorem, ale není kauzálním faktorem, je pohlaví. Pohlaví má statistické souvislosti s psychosociálními projevy obecně (Zell, Krizan, & Teeter, 2015) a konkrétně i v oblasti ochrany klimatu. Avšak pohlaví je obecným rysem, který definuje danou osobu a není možné ho změnit nezávisle na individuu. Pokud bychom chtěli interpretovat statistické efekty pohlaví jako kauzální efekty, museli bychom odpovědět na otázku, co konkrétně způsobuje pozorované statistické efekty. Ideálně by to měl být nějaký faktor (např. způsob výchovy, geneticky dané odpovědi organismu apod.), který – kdybychom ho změnili v hypotetickém experimentu – by vedl k pozorovaným statistickým efektům pohlaví. Kauzálním faktorem potom není pohlaví člověka samo o sobě, ale právě tento faktor, který splňuje podmínku vystavitelnosti. Důležité je uvědomit si, že zatímco technicky je možné odhadnout statistické efekty nezávislých proměnných, které nesplňují podmínku vystavitelnosti, z logicko-teoretického hlediska nemohou být takové efekty považovány za kauzální efekty.

3.2 Typy efektů

Pokud hovoříme o kauzálním efektu, máme tím nejčastěji na mysli průměrný efekt faktoru (*average treatment effect, ATE*), tj. očekávanou hodnotu rozdílu závislé proměnné v situaci působení kauzálního faktoru a v situaci, kdy tento faktor nepůsobí. Formálně je možné vyjádřit tento efekt jako očekávanou hodnotu rozdílu dvou úrovní závislé proměnné,

$$ATE = E(Y_{1i} - Y_{0i}) \quad (3.1)$$

kde y_0 je hodnota závislé proměnné před působením kauzálního faktoru a y_1 je hodnota závislé proměnné po působení kauzálního faktoru a i indexuje individua.

Jinou často používanou mírou je průměrný efekt u lidí vystavených působení kauzálního faktoru (*average treatment effect on treated, ATET*), který je vyjádřen jako očekávaná hodnota rozdílu závislé proměnné po a před působením kauzálního faktoru, ovšem pouze u skupiny jedinců vystavených působení kauzálního faktoru. Formálně je možné vyjádřit tento efekt jako podmíněnou očekávanou hodnotu,

$$ATET = E[(Y_{1i} - Y_{0i}) | D_i = 1] \quad (3.2)$$

kde y_0 je hodnota závislé proměnné před působením kauzálního faktoru a y_1 je hodnota závislé proměnné po působení kauzálního faktoru a i indexuje jednotky skupiny (individua) a D je indikátor skupiny vystavené působení kauzálního efektu (tj. skupina, v níž je $D = 1$).

Rozdíly mezi těmito dvěma způsoby měření kauzálního efektu jsou následující. *ATE* odpovídá na otázku, jaký je efekt kampaně v rámci skupin, které jsou kampani aktuálně vystaveny i těch, které by jí mohly být vystaveny v budoucnu. Naopak *ATET* měří jen efekt pro skupinu, která je aktuálně vystavena působení kauzálního faktoru; na základě *ATET* je tudíž obtížné zobecňovat na další skupiny. Pokud jsou působení kauzálního faktoru vystaveny všechny skupiny, platí, že $ATE = ATET$. Usuzovat z *ATET* na *ATE* není v mnoha případech možné a může nastat jak situace kdy $ATET < ATE$ (např. kampaň je zaměřena na část populace, která tuto kampaň odmítá), ale může také nastat situace, kdy $ATET > ATE$ (kampaň je zaměřena na část populace, která kampaň velmi dobře přijímá). Odhadnout *ATET* je v mnoha případech jednodušší než odhadnout *ATE*, protože při odhadu *ATE* je nutné zobecňovat i na skupinu, která působení kauzálního faktoru nebyla doposud vystavena. To je možné pouze s využitím statistické indukce, tzn. statistických postupů, které umožňují zobecnit ze skupiny aktuálně vystavené působení kampaně na skupinu, která ještě jejímu působení vystavena nebyla.

Hovoříme-li o kauzálních efektech, je také třeba odlišovat kauzální efekty pozorované ve vzorku účastníků studie (*SATE, sample average treatment effect* a *STATET, sample average treatment effect on treated*) a populaci (*PATE, population average treatment effect* a *PATET, population average treatment effect on treated*). Toto rozlišení nás upozorňuje na fakt, že o populačních odhadech můžeme hovořit pouze tehdy, pokud můžeme na základě statistické indukce zobecnit odhady efektů na celou populaci (*PATE*) nebo její část vystavenou intervenci (*PATET*). Přísně vzato, takové zobecnění je možné pouze tehdy, pokud jsou účastníci studie z populace vybírání náhodně (tj. tvoří náhodný vzorek populace). Toto kritérium náhodného výběru ale nesmíme zaměňovat s procesem náhodného rozdělení účastníků experimentu do experimentálních skupin (randomizace), který je typický pro randomizované experimenty, o nich budeme hovořit v této kapitole později.

Měření průměrného kauzálního efektu je spojeno s představou homogenního efektu na všechny jednotky (v našem případě zpravidla jednotlivce, ale také např. skupiny) vystavené působení kauzálního faktoru. V takovém případě přináší údaj o průměrném efektu informaci o vlivu kauzálního faktoru na studované jednotky. Ve skutečnosti není samozřejmě efekt kauzálního faktoru na studované jednotky nikdy homogenní. Mohou proto nastat situace, kdy je tato heterogenita natolik velká, že údaj o průměrné velikosti kauzálního efektu přestává být informativní a naopak zajímavými se stávají informace o rozdílech v kauzálních efektech na různé jednotky. Taková situace může nastat, například když nějaké vlastnosti studovaných jednotlivců ovlivňují velikost kauzálního efektu (např. proto, že jedinci mají různou motivaci číst informace prezentované v kampani, anebo jim věnují různou pozornost).⁴ Jak vyplývá z povahy fundamentálního problému kauzálního usuzování, není možné stanovit velikost kauzálního efektu na úrovni individua, ale pouze skupin. Kvůli tomuto omezení je možné studovat heterogenitu kauzálního efektu pouze na úrovni skupin. Takto měřený kauzální efekt bývá označován jako průměrný lokální kauzální efekt (*local average treatment effect, LATE*). Protože hlavním cílem této metodiky je poskytnout vodítko pro hodnocení efektů kampaní a nikoli pro jejich explorativní zkoumání, nebudeme se v této metodice lokálními efekty zabývat a případné zájemce odkážeme na specializovanou literaturu (Angrist & Pischke, 2009; Imbens & Angrist, 1994).

3.3 Experimentální a kvazi-experimentální studie

Kauzální efekty je možné odhadnout s využitím experimentálních nebo kvazi-experimentálních studií (např. Angrist & Pischke, 2009; Cox, 1992; Cox & Reid, 2000; Shadish, Cook, & Campbell, 2001). Experimentální studie se vyznačují tím, že mohou měnit úroveň kauzálního faktoru a tak systematicky studovat jeho vliv na závislou proměnnou. Oproti tomu kvazi-experimentální studie neusilují přímo o změnu úrovně kauzálního faktoru, ale místo toho se zaměřují na zkoumání podobných jednotek, které jsou vystaveny různým úrovním kauzálního faktorů. Takto kvazi-experimentální studie napodobují uspořádání randomizovaných experimentálních studií, jak si dále v této kapitole vysvětlíme.

Experimentální studie přijímají při hodnocení kauzálních efektů relativně slabé předpoklady, a tudíž jsou jejich výsledky relativně méně otevřené kritice v důsledku porušení těchto předpokladů. Oproti tomu kvaziexperimentální studie musejí přijímat daleko silnější předpoklady, které jsou často netestovatelné (např. předpoklad o podmíněné nezávislosti, viz Holland, 1986), a proto jsou jejich výsledky více otevřené případné kritice. Často však není možné zajistit v evaluačních studiích randomizaci, tedy náhodné rozdělení jednotlivců do skupin, které jsou v různé míře vystaveny kampani, což je klíčový prvek randomizovaných experimentálních studií. Náhodnému přiřazování

⁴ Z technického hlediska se jedná o problém endogenity, kdy nezávislá proměnná (např. vzdělávací program) a závislá proměnná (např. ochota k mitigačním aktivitám) jsou obě ovlivněny třetí neznámou proměnnou (např. postojem ke klimatické změně).

individuí brání často etické ohledy (např. není možné nutit participanty, aby se účastnili dobrovolně a bez náhrady určitých aktivit, anebo aby zásadním způsobem měnili své chování), anebo praktické důvody (např. vliv environmentálních postojů jako kauzálního faktoru je obtížné studovat experimentálně, protože environmentální postoje jsou relativně stabilní a mění se spíše v řádu desetiletí, viz Otto & Kaiser, 2014). V takových situacích jsou kvaziexperimentální studie nejlepší volbou pro zkoumání kauzálního efektu, ovšem s řadou omezení, které tyto přístupy mají (v důsledku silných předpokladů, které musejí, na rozdíl od experimentálních studií, přijmout).

3.3.1 Experimentální přístupy

Experimentální studie mohou mít různé typy výzkumných uspořádání (nazývaných také experimentální plány nebo designy; anglicky *experimental designs*). Experimentální studie využívají dva základní typy uspořádání: vnitrosubjektové (*within-subjects design*) a mezisubjektové (*between-subjects design*), popř. jejich kombinace. Vnitrosubjektové uspořádání spočívá na opakovém pozorování téhož individua, přičemž jednotlivá pozorování poskytují kontrafaktuální evidenci⁵ pro odhad kauzálního efektu. Oproti tomu mezisubjektové uspořádání spočívá na porovnávání skupin tvořených různými individui. Kontrafaktuální evidence je v takovém případě zajištěna náhodným rozdělením individuí do skupin, které vede k tomu, že míry centrální tendence (průměr skupin) nejsou ovlivněny vnějšími faktory a mohou tak posloužit jako kontrafaktuální evidence v procesu testování kauzálních hypotéz.

3.3.1.1 Experimentální vnitrosubjektové uspořádání

Společným rysem experimentálních studií s vnitrosubjektovým uspořádáním (základní informace o tomto uspořádání uvádí např. Myers & Hansen, 2006) je to, že využívají opakované měření na jednom individuu pro odlišení kauzálního efektu intervence od vlivu dalších faktorů, které souvisejí s vlastnostmi individuí (např. jejich postoji) a které mohou potenciálně ovlivnit závislou proměnnou. Opakované měření na týchž subjektech (minimálně jednou před – měření označované jako pretest – a minimálně jednou po působení kauzálního faktoru – měření označované jako posttest) vlastně umožňuje pozorovat totéž individuum ve stavu, kdy na něj nepůsobí kauzální faktor a ve stavu, kdy na něj tento faktor působí. Tímto způsobem poskytují vnitrosubjektové experimenty kontrafaktuální evidenci o kauzálním efektu určitého faktoru.

⁵ Kontrafaktuální evidencí se v tomto textu, stejně jako v odborné literatuře zaměřené na kauzální inferenci, myslí informace o tom, jaké důsledky by mělo vystavení (nebo naopak nevystavení) určité jednotky (často jednotlivce) působení kauzálního faktoru, oproti stávajícímu stavu, kdy tomuto působení vystaven není (anebo naopak je).

Nejjednodušším příkladem vnitrosubjektového experimentálního uspořádání jsou experimenty s jednou skupinou, které měří závislou proměnnou před a po intervenci. Průměr rozdílů úrovně závislé proměnné přes všechna individua účastníci se experimentu je nejjednodušší mírou kauzálního efektu (jedná se pochopitelně o míru ATET).

Takové vnitrosubjektové experimenty s jednou skupinou trpí jedním závažným nedostatkem, a to, že zanedbávají efekt času, resp. efekt všech událostí, které mohou nastat v období mezi pretestem a posttestem. Určitým způsobem, jak se s tímto problémem vyrovnat, je zahrnout explicitně takové dodatečné faktory (označované jako kovariáty, protože kovariují se závislou proměnnou, ale nejsou v experimentu systematicky měněny) jako dodatečné vysvětlující proměnné. Protože ale v řadě situací není možné všechny kovariáty zahrnout (buď je neznáme, anebo je jich příliš mnoho), může toto řešení vést ke zkresleným odhadům efektu. Jinou možností je použití vnitrosubjektového uspořádání s kontrolní skupinou.

Vnitrosubjektové experimenty s kontrolní skupinou využívají dodatečnou kontrolní skupinu pro kontrolování faktorů, které mohou ovlivnit závislou proměnnou nad rámec kauzálního efektu faktoru, který nás v experimentu zajímá především. Nejjednodušším příkladem takového uspořádání jsou vnitrosubjektové experimenty s kontrolní a experimentální skupinou. V obou skupinách je proveden pretest a posttest na stejných jedincích, ale pouze v experimentální skupině je mezi pretestem a posttestem provedena intervence (v kontrolní skupině intervence provedena není). V případě, že je rozdělení jednotlivců do kontrolní a experimentální skupiny provedeno náhodně, poskytuje takový experiment poměrně přesný odhad kauzálního efektu. Vhodnou mírou kauzálního efektu představuje průměrný rozdíl mezi rozdíly individuálních skóre v pretestu a posttestu v experimentální oproti kontrolní skupině (tento způsob odhadu bývá označován jako rozdíl v rozdílech, anglicky *difference in differences* a jedná se o míru kauzálního efektu ATE). Více informací o odhadu kauzálních efektů v experimentech s opakovaným měřením nalezne čtenář ve specializované literatuře (Bonate, 2000).

Velkou výhodou experimentů s vnitrosubjektovým uspořádáním je to, že umožňují velmi dobře kontrolovat odlišnosti a zvláštnosti jednotlivců. Je to dáno tím, že toto uspořádání je založeno na porovnávání rozdílů mezi pretestovým a posttestovým měřením na témže jednotlivci. Odhadovaný kauzální efekt proto není ovlivněn těmi charakteristikami jednotlivců, které zůstávají mezi pretestem a posttestem stabilní, i když se jedná o charakteristiky, které zakládají významné rozdíly mezi jednotlivci. Tato vlastnost vnitrosubjektového uspořádání experimentů má za následek to, že za jinak stejných podmínek, mají vnitrosubjektové experimenty asi 4-6x vyšší statistickou sílu než mezisubjektové experimenty (Bellemare & Bissonnette, 2014). Prakticky to znamená, že studie využívající vnitrosubjektové uspořádání mohou při zkoumání kauzálních efektů využívat

několikanásobně menší vzorky respondentů při zachování stejné síly statistického testu oproti mezisubjektovým experimentálním studiím nebo kvaziexperimentálním studiím.

Opakované měření na týchž jednotlivcích, které je velkou výhodou vnitrosubjektových experimentů, je zároveň jejich slabinou. Samotný fakt, že jedinec je opakovaně vystaven podobným nebo stejným testovým situacím, totiž může ovlivnit výsledky těchto měření. Opakovaná měření, která jsou navíc doprovázena experimentální manipulací, totiž často vedou účastníky experimentálních studií k tomu, že se snaží uhádnout hypotézy výzkumu, což následně vede k ovlivnění jejich reakcí (ať už vědomě, např. ve snaze „pomoci“ výzkumníkovi, anebo podvědomě). Takové efekty jsou označovány jako efekty experimentální poptávky (anglicky *demand effects*). Jinými efekty, které mohou zkreslit výsledky opakovaných měření, jsou efekty pořadí (jedná se obecně o všechny typy nežádoucích efektů, které vznikají díky tomu, že jedinci se seznamují s procesem měření, což ovlivňuje jejich výsledek). Tyto potenciální problémy vnitrosubjektových experimentů jsou natolik vážné, že vedou řadu autorů k tomu, že doporučují zcela upustit od vnitrosubjektových experimentů (viz Charness, Gneezy, & Kuhn, 2012 pro shrnutí této diskuse).

3.3.1.2 Experimentální mezisubjektové uspořádání

Alternativu k vnitrosubjektovému uspořádání představuje mezisubjektové experimentální uspořádání (základní informace o tomto uspořádání uvádí např. Myers & Hansen, 2006). Toto uspořádání nabízí jiné řešení fundamentálního problému kauzálního usuzování než vnitrosubjektové uspořádání. Zatímco vnitrosubjektové uspořádání je postaveno na porovnávání různých výsledných stavů závislé proměnné v závislosti na experimentálních podmínkách u stejných individuí, mezisubjektové uspořádání porovnává výsledné stavy u různých skupin jedinců, přičemž každá z těchto skupin je vystavena jiným úrovním kauzálního faktoru (tj. jiným experimentálním podmínkám).

Je důležité si uvědomit, že porovnání různých výsledných stavů u různých individuí není samo o sobě mírou kauzálního efektu ATE nebo ATET, protože tato individua se budou vždy lišit v určitých podstatných rysech, které mohou tyto výsledné stavy ovlivnit. Pokud například vystavíme určitého člověka působení informační kampaně a porovnáme jeho chování s chováním jiného člověka, který této kampani vystaven nebyl, pak rozdíl výsledných stavů těchto dvou lidí (tj. např. rozdíly v jejich chování) nejsou mírou kauzálního efektu poskytnuté informace na chování. Je tomu tak proto, že tito dva jedinci se pravděpodobně chovali rozdílně již předtím, než byl jeden z nich vystaven působení informace a navíc se mohli lišit i v psychosociálních charakteristikách (např. znalostech nebo postojích), které ovlivňují efekt poskytnutých informací na chování.

Řešením tohoto problému mezisubjektového experimentálního uspořádání je náhodné rozdělení jedinců do skupin (randomizace) a následné porovnávání centrálních tendencí ve skupinách.

Jak vyplývá z teorie velkých čísel, náhodné rozdělení jedinců vede k tomu, že s rostoucí velikostí vzorků konvergují průměrné hodnoty ve skupinách k populačnímu průměru. Jinými slovy, náhodné rozdělení jedinců do skupin nám umožňuje eliminovat působení dalších faktorů, které měly u každého jednotlivce odlišnou úroveň ještě před tím, než vůbec začal působit námi sledovaný kauzální faktor. Je však třeba si uvědomit, že randomizace nevede k eliminaci rozdílů mezi jednotlivci: jednotlivci přiřazení náhodně do skupin jsou stále stejně odlišní, pouze průměrné hodnoty všech charakteristik individuů (např. inteligence, postoje apod.), by měly být na úrovni skupin podobné (podle teorie velkých čísel budou konvergovat ke stejné populační hodnotě a při nekonečně velkých skupinách budou randomizované skupiny shodné). Randomizace takto umožňuje naplnit předpoklad nepodmíněné zanedbatelnosti (*unconditional ignorability*), což je předpoklad, že vystavení experimentálním podmínkám je nezávislé na úrovni závislé proměnné. Jinými slovy: to, do jaké experimentální skupiny bude jedinec zařazen a jestli bude vystaven působení kauzálního faktoru, nezávisí na tom, jaká je úroveň závislé proměnné (u daného člověka) před experimentem. V takovém případě je možné zanedbat vliv třetích proměnných, které nejsou součástí experimentu jako nezávislé proměnné (tj. všechny charakteristiky jednotlivců, které mohou odlišovat jednotlivce a které mohou ovlivňovat výsledné stavy nad rámec kauzálního faktoru).

Na tomto místě je dobré zdůraznit zcela jedinečnou roli randomizace a s ní spojenou nepodmíněnou zanedbatelnost při odhadu kauzálních efektů. V oblasti sociálně vědního výzkumu se obvykle setkáváme s tím, že jakákoli proměnná je ovlivněna téměř nekonečnou řadou jiných proměnných. Nepodmíněná zanedbatelnost umožňuje "odstínit" vliv třetích proměnných (které nejsou předmětem výzkumu, ale mohou potenciálně ovlivnit výsledný stav). Nepodmíněná zanedbatelnost je dosažitelná pouze v randomizovaných experimentech, ale ne například v kvaziexperimentálních nebo korelačních studiích. Tyto přístupy často usilují o podmíněnou zanedbatelnost (*conditional ignorability*), což je stav, kdy explicitní zahrnutí třetích faktorů do statistických modelů umožňuje odhadnout nezkrácený kauzální efekt. Problémem takového předpokladu ovšem je to, že spoléhá na zahrnutí všech faktorů, které souvisejí s kauzálním faktorem a výsledným stavem (což je většinou nerealistický předpoklad) a že je netestovatelný.

Z teorie velkých čísel vyplývá, že míra jakou randomizace „odstíní“ vliv třetích faktorů, závisí na velikosti vzorků, které jsou použité pro zkoumání efektu a na různorodosti populace. Různorodostí zde myslíme, jak jsou jedinci odlišní z hlediska všech proměnných, které potenciálně ovlivňují výsledný stav závislé proměnné. Pokud je populace hodně různorodá (což může být třeba případ názorů na klimatickou změnu v české společnosti), anebo pokud jsou použité vzorky relativně malé, bude odhad kauzálního efektu relativně nepřesný (tj. interval spolehlivosti pro odhad efektu bude relativně široký). Odhad kauzálního efektu v mezisubjektovém uspořádání je možné zpřesnit zahrnutím dodatečných

vysvětlujících proměnných, které zachycují rozdíly mezi jednotlivci před působením kauzálního faktoru (např. sociodemografické proměnné nebo postoje a znalosti). Dokonce je možné zahrnout jako dodatečnou vysvětlující proměnnou i úroveň závislé proměnné před působením kauzálního faktoru, čímž se toto uspořádání přibližuje vnitrosubjektovému uspořádání (vice informací o takovém uspořádání s individuálními vnitrosubjektovými prediktory dává např. Bonate, 2000).

Nepřesnost odhadu kauzálního efektu v mezisubjektovém uspořádání v důsledku různosti jednotlivců může vyřešit také párování nebo blokování. Obecně se jedná o takové postupy, kdy jsou jedinci buď spárováni na základě podobnosti z hlediska klíčových proměnných, které mohou mít potenciálně vliv na závislou proměnnou, anebo jsou na základě takové podobnosti rozděleni do tzv. bloků, ještě před působením kauzálního faktoru. Využívá se pro to například metoda *propensity score matching*, při níž se identifikují proměnné, které jsou nejvíce asociované s příslušností ke kontrolní či experimentální skupině a následně se v rámci analýzy srovnají páry jedinců z obou skupin, kteří jsou si v hodnotách těchto proměnných nejpodobnější. Následně je provedeno náhodné rozdělení jednotlivců do skupin, při němž jsou jedinci rozdělováni náhodně do skupin v rámci každého páru nebo bloku (fakticky se tedy jedná o stratifikovaný náhodný výběr). Teprve potom je jedna (nebo více skupin) vystavena působení kauzálního faktoru.

3.3.1.3 Typy mezisubjektového uspořádání

Nejjednodušším typem mezisubjektového experimentálního uspořádání je uspořádání se dvěma skupinami. Jednotlivci jsou do skupin rozděleni náhodně a následně je jedna ze skupin vystavena působení kauzálního faktoru. Následný rozdíl v průměrných výsledných hodnotách závislé proměnné pozorovaný mezi skupinami je odhadem průměrného kauzálního efektu *ATE* (např. pokud je vzorek náhodně vybírán z celé populace) nebo *ATET* (např. pokud jsou do experimentu zahrnuti pouze ti, kterým je kampaň určena). Mezisubjektové uspořádání se dvěma skupinami umožňuje zkoumat pouze vliv jednoho kauzálního faktoru s jednou úrovní intenzity (tj. jedna skupina je vystavena působení faktoru o určité intenzitě, zatímco druhá skupina je vystavena působení téhož faktoru o jiné intenzitě, anebo není tomuto faktoru vystavena vůbec - potom se jedná o kontrolní skupinu).

Pro řešení složitějších otázek o povaze kauzálního efektu je vhodné faktoriální uspořádání, které využívá jednu skupinu pro každou kombinaci úrovní faktorů (faktorem je v našem případě jakákoli proměnná, která charakterizuje kampaň.) Pokud nás například zajímá, jaký efekt má na mitigační chování podávání informací o rizicích a pozitivních dopadech klimatické změny, bude jedním faktorem přítomnost pozitivní informace (resp. její absence) a druhým faktorem přítomnost informace o riziku. Kombinace 2 (pozitivní informace: je vs. není) x 2 (informace o riziku: je vs. není) pak označuje tzv. kompletní faktoriální uspořádání zahrnující všechny kombinace úrovní faktorů.

Faktoriální uspořádání může mít mnoho podob od úplného faktoriálního uspořádání až po různě redukována faktoriální uspořádání (označované jako *fractional factorial design*). Důvodem pro redukci faktoriálního uspořádání je to, že s množstvím faktorů a jejich úrovní roste exponenciálně počet skupin nutných pro takový typ uspořádání. Pokud máme například 3 faktory a každý má dvě úrovně, potřebujeme $2 \times 2 \times 2$, tj. 8 skupin. Jak jsme již řekli výše, velikost skupin má vliv na přesnost odhadu kauzálních efektů a proto složité faktoriální designy vyžadují pro zachování přesnosti odhadu efektů velmi velké vzorky. Redukované faktoriální designy mohou být ze statistického hlediska efektivnější, tj. odhadnou určité efekty stejně přesně, ale s využitím menšího množství informace (tj. např. s využitím menšího vzorku). Nejznámějším redukováným faktoriálním uspořádáním je tzv. ortogonální faktoriální design, který umožňuje odhad hlavních efektů faktorů, ale ne jejich interakcí (základní informace o tomto uspořádání uvádí např. Cavazzuti, 2013). Redukované faktoriální uspořádání je ale možné navrhnout vždy tak, aby bylo možné testovat právě ty kauzální hypotézy, které nás zajímají nejvíce, a zvýšit efektivitu jejich odhadu za cenu toho, že pomineme zkoumání jiných efektů.

3.3.2 Kvaziexperimentální přístupy

Ne vždy je možné z praktických nebo etických důvodů náhodně vystavovat jedince působení kauzálních faktorů (což nastává, pokud jsou jedinci náhodně přiřazováni do jedné z experimentálních skupin). Alternativním postupem umožňujícím za určitých podmínek odhad kauzálního efektu jsou v takovém případě kvaziexperimentální přístupy.

Kvaziexperimentální přístupy nevyužívají randomizaci a tudíž nemohou spoléhat na předpoklad nepodmíněné zanedbatelnosti, o níž jsme hovořili dříve. Místo toho využívají předpoklad podmíněné zanedbatelnosti. Předpoklad podmíněné zanedbatelnosti znamená, že přiřazení jedinců do kvaziexperimentální a kvazikontrolní skupiny je podmíněně nezávislé na úrovni závislé proměnné (podmíněně vzhledem ke skupině kontrolních proměnných, které musejí být do analýzy explicitně zahrnuty). Jedná se tedy o podmíněnou nezávislost, která platí pouze tehdy, pokud jsou do modelu zahrnuty všechny prediktory, které ovlivňují jak nezávislou, tak závislou proměnnou. Tento předpoklad je netestovatelný (není možné zjistit, jestli jsou zahrnuty všechny takové faktory). Poznatky o kauzálních efektech, které je možné získat s využitím kvaziexperimentálních studií jsou tedy, za jinak podobných podmínek (velikost vzorku, uspořádání, různorodost skupin), slabší a snáze napadnutelné, než výsledky experimentálních studií.

Patrně nejznámějším kvaziexperimentálním přístupem je *propensity score matching* (základní úvod dává např. Angrist & Pischke, 2009; Guo & Fraser, 2010; netechnický úvod pro širokou veřejnost uvádí např. White & Sabarwal, 2014). Tento postup se v podstatě snaží o napodobení experimentálního mezisubjektového uspořádání s využitím statistického modelu, který bere v potaz

faktory, které ovlivňují to, jaké úrovni kauzálního faktoru je daný jedinec vystaven. Fungování *propensity score matching* si můžeme přiblížit na příkladu hodnocení dopadů kampaně Do práce na kole realizovaného v rámci jedné z demonstračních studií, jejíž popis uvádí Box 5.3 v kapitole 5.

Použití experimentálních přístupů s randomizací by bylo v případě hodnocení kampaně *Do práce na kole* obtížné. Tato kampaň trvá celý měsíc a účastníci této kampaně soutěží v pracovních týmech o nejdelší ujetou vzdálenost na kole při cestách do práce. Náhodné přiřazování účastníků do kontrolní (neúčastní se kampaně) a experimentální skupiny (účastní se kampaně) by bylo obtížné, přinejmenším kvůli zajištění motivace v účasti. Použití vnitrosubjektového uspořádání je zase limitující v tom, že využívání cyklistické dopravy k cestám do práce se mění v čase a je závislé na počasí. Změna chování pozorovaná s vnitrosubjektovým uspořádáním by tudíž mohla být vyvolána pouze změnami počasí (nebo jinými faktory, které se mění v čase). Kvaziexperimentální přístup *propensity score matching* tedy nabízí v případě hodnocení této kampaně vhodnou alternativu.

Před začátkem kampaně vybereme vzorek účastníků kampaně a kvazikontrolní vzorek jedinců, kteří se podobají účastníkům kampaně z hlediska faktorů, pravděpodobně ovlivňujících účast v kampani. Důležité je, že ve vzorku účastníků jsou jedinci, kteří se sami rozhodli kampaně zúčastnit, zatímco v kvazikontrolní skupině jsou jedinci, kteří se sami rozhodli nezúčastnit. Zdůrazněme tedy ještě jednou, že jedinci nejsou do těchto skupin přiřazováni náhodně.

Metoda *propensity score matching* následně sestavuje páry jednotlivců, kteří jsou si co nejvíce podobní z hlediska faktorů, které pravděpodobně ovlivňují účast v kampani. Existují různé způsoby párování (jeden na jednoho nebo více na jednoho) a mnoho různých algoritmů, které mohou být použity k párování. Jedinci, kteří jsou statisticky spárováni, tvoří dvě virtuální skupiny (experimentální a kontrolní), které mohou být porovnány podobným způsobem, jako experimentální skupiny u mezisubjektových randomizovaných experimentů.

Teoreticky mohou být s využitím *propensity score matching* odhadnuty míry kauzálních efektů *ATET* i *ATE*, v závislosti na síle přijatých předpokladů. Míra *ATET* může být odhadnuta, pokud platí, že byly do analýzy zahrnuty všechny podstatné prediktory členství v kvaziexperimentální skupině. Tím pádem je zajištěno to, že po jejich zahrnutí je členství ve skupině nezávislé na úrovni závislé proměnné před působením kauzálního faktoru (předpoklad absence zavádějícího efektu). Míra *ATE* může být odhadnuta po přijetí ještě dalšího dodatečného předpokladu, že totiž jedinci v kontrolní skupině (tj. ti, kteří se kampaně nezúčastnili) by se jí mohli zúčastnit (předpoklad překryvu; naplnění obou podmínek, tj. absence zavádějícího efektu a překryvu je označována jako předpoklad silné zanedbatelnosti, *strong ignorability*). Podmínky pro platnost kauzálních odhadů u *propensity score matching* jsou detailně popsány ve specializované literatuře (např. Guo & Fraser, 2010; Rosenbaum & Rubin, 1983).

Metoda *propensity score matching*, jako ostatně i jiné přístupy k odhadu kauzálního efektu, které jsou založeny na předpokladu podmíněné zanedbatelnosti (tj. regresní modelování anebo metoda instrumentálních proměnných; pro přehled těchto metod viz Angrist & Pischke, 2009) jsou často kritizovány proto, že dávají výzkumníkům do ruky nástroj, který mohou chybně použít k odhadu kauzálních efektů i v situaci, kdy nejsou naplněny podmínky pro použití těchto metod (např. podmíněná nebo silná zanedbatelnost, viz Holland, 1986; D. Rubin, 1974). I přes tuto kritiku je třeba si uvědomit, že kvaziexperimentální přístupy mohou v určitých situacích (tj. když není možné použít experimentální přístupy s nebo bez randomizace) poskytovat jediný prakticky využitelný nástroj k odhadu kauzálních efektů kampaní.

3.4. Shrnutí

V této kapitole jsme představili pojem efektu kampaně, který definujeme jako kauzální efekt ve smyslu Rubinova kauzálního modelu potenciálních výsledků. Tento model chápe kauzální efekt jako rozdíl mezi dvěma potenciálními stavy, které mohou nastat, pokud se změní nezávislá proměnná (tj. stavu kdy působí kampaň oproti stavu, kdy nepůsobí). Jen takový faktor (nezávislá proměnná), který splňuje podmínku vystavitelnosti (tj. můžeme jeho působení vystavit určitému jednotlivci), může být považován za kauzální faktor. V této kapitole jsme představili různé typy měr kauzálního efektu (*ATE*, *ATET*, *PATE*), které se liší tím, jak je na jejich základě možné zobecňovat existenci efektu pro celou populaci. Dále jsme v této kapitole představili dva základní typy měření efektů kampaní. První způsob odhadů efektů je postaven na využití experimentálních studií s vnitrosubjektovým (s jednou skupinou, nebo s kontrolní a experimentální skupinou), anebo mezisubjektovým uspořádáním (s kontrolní a experimentální skupinou nebo s faktoriálním uspořádáním). Výhodou experimentálních studií je to, že mohou pomocí randomizace odstínit vliv dalších proměnných, které potenciálně zkreslují odhad efektu kampaně. Naopak nevýhodou experimentálních přístupů může být to, že v některých situacích není prakticky možné zajistit manipulaci nezávislé proměnné u vybrané skupiny lidí (např. není z etických důvodů), která je předpokladem experimentálního uspořádání. V takových situacích je možné využít pro odhad efektu kvaziexperimentální přístupy. Kvaziexperimentální přístupy nevyužívají randomizaci, ale snaží se napodobit uspořádání randomizovaného experimentu tím, že kontrolují všechny potenciální prediktory závislé proměnné. Nevýhodou kvaziexperimentálních přístupů je to, že jejich odhady kauzálního efektu mohou být zkreslené, pokud se jim nepodaří zachytit všechny důležité prediktory závislé proměnné, přičemž je prakticky velmi obtížné takové zkreslení zjistit.

Rozšiřující četba

- Cox, D. R. (1992). *Planning of experiments* (Wiley Classics Library ed). New York: Wiley.
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Myers, A., & Hansen, C. H. (2006). *Experimental psychology* (6th ed). Belmont, CA: Thomson/Wadsworth.

Seznam literatury

- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 edition). Princeton: Princeton University Press.
- Bellemare, C., & Bissonnette, L. (2014). *Statistical power of within and between-subjects designs in economic experiments*. Bonn, Německo: Institute for the Study of Labor.
- Bonate, P. L. (2000). *Analysis of pretest-posttest designs*. Boca Raton: Chapman & Hall/CRC.
- Cavazzuti, M. (2013). Design of Experiments. In M. Cavazzuti, *Optimization Methods* (s. 13–42). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-31187-1_2
- Cox, D. R. (1992). *Planning of experiments* (Wiley Classics Library ed). New York: Wiley.
- Cox, D. R., & Reid, N. (2000). *The theory of the design of experiments*. Boca Raton: Chapman & Hall/CRC.
- Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: statistical methods and applications*. Thousand Oaks, Calif: Sage Publications.
- Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945–960. doi:10.2307/2289064
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8. doi:10.1016/j.jebo.2011.08.009
- Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467–475.
- Myers, A., & Hansen, C. H. (2006). *Experimental psychology* (6th ed). Belmont, CA: Thomson/Wadsworth.
- Otto, S., & Kaiser, F. G. (2014). Ecological behavior across the lifespan: Why environmentalism increases as people grow older. *Journal of Environmental Psychology*, 40, 331–338. doi:10.1016/j.jenvp.2014.08.004
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. doi:10.1093/biomet/70.1.41

- Rubin, D. (1974). Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies. *Journal of Education Psychology*, 66(5), 688–701.
- Rubin, D. B. (2005). Causal Inference Using Potential Outcomes: Design, Modeling, Decisions. *Journal of the American Statistical Association*, 100(469), 322–331.
doi:10.1198/016214504000001880
- Sekhon, J. (2007). The Neyman–Rubin Model of Causal Inference and Estimation via Matching Methods. In J. M. Box, H. E. B. Steffensmeier, & D. Collier (Ed.), *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- White, H., & Sabarwal, S. (2014). *Quasi-experimental design and methods*. Florencie, Italie: United Nations Children’s Fund (UNICEF). Získáno z https://www.unicef-irc.org/publications/pdf/brief_8_quasi-experimental%20design_eng.pdf
- Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist*, 70(1), 10–20. doi:10.1037/a0038208

4. Posuzování velikosti efektu kampaní

Tato kapitola představuje hodnocení velikosti efektu pomocí standardizovaných i nestandardizovaných měr a ukazuje, jakým způsobem je možné tyto velikosti efektů interpretovat tak, aby byly srozumitelné i širší veřejnosti.

4.1 Důvody pro hodnocení velikosti efektu

Často nás zajímá, jestli má nějaká kampaň nebo intervence zamýšlený efekt neboli účinek a jestli je tento efekt natolik velký, aby měl nějaký praktický význam. Tyto otázky se týkají velikosti efektu. Pojem velikosti efektu (*effect size*) je jedním z klíčových technických pojmů metodologie kvantitativního výzkumu a současně je to pojem, který propojuje výsledky výzkumu s praktickým životem. Schopnost vyjádřit velikost efektu a správně ji interpretovat je tedy zcela klíčová: například při předávání výsledků studií jejich zadavatelům, anebo široké veřejnosti, která zpravidla nemá podrobné znalosti metodologie kvantitativního výzkumu.

Domníváme se, že je nutné zabývat se problematikou měření a vyjádření velikosti efektu ještě z několik dalších důvodů. Prvním důvodem je to, že problematika měření velikosti efektu byla až donedávna v kvantitativním výzkumu opomíjena (Callahan & Reio Jr, 2006; Osborne, 2008) a dostává se do popředí zájmu kvantitativních metodologů teprve relativně nedávno (Cumming, 2011; srv. též American Psychological Association, 2008, 2010). Tento obrat souvisí mj. s tím, že výzkumníci si začínají uvědomovat nedostatečnost posuzování výsledků studií pouze z hlediska statistické významnosti (ta nám říká pouze to, jestli máme věřit tomu, že nějaký efekt existuje v populaci, z níž pochází náš vzorek) a začínají častěji brát do úvahy hledisko praktické významnosti (tj. jestli je zjištěný efekt dostatečně velký, aby měl vůbec nějaké využití v praktickém životě; viz Kirk, 1996).

Druhým důležitým důvodem, proč se zabývat velikostí efektu je to, že vyjádření velikosti efektu umožňuje akumulaci našeho poznání. Porovnání efektů různých kauzálních faktorů, např. formou meta-analýzy, nám umožní odhadnout pravděpodobný efekt určité kampaně, kterou se teprve chystáme realizovat. Jako příklad takové srovnávací studie můžeme uvést nedávno publikovanou metaanalýzu, která odhaduje průměrné efekty různých faktorů na představy o klimatické změně a dále vliv představ o klimatické změně na chování (Hornsey, Harris, Bain, & Fielding, 2016b).

Poslední důvod je metodologický a souvisí s tím, že znalost velikosti efektu kampaně umožňuje navrhovat efektivní evaluační studie pro zkoumání takových efektů v budoucnu. Přesněji řečeno, pokud známe (např. na základě studií v příbuzných oblastech) pravděpodobnou velikost efektu kampaně, kterou chceme teprve studovat, můžeme na základě analýzy statistické síly (power analysis,

viz Cohen, 1988; Ellis, 2010) odhadnout, jaké parametry (např. velikost vzorku) by měla mít evaluační studie, aby vůbec takové efekty mohla zachytit.

Ve zbývajících částech této kapitoly představíme různé míry velikosti efektu, jejich logiku, způsoby odhadu přesnosti velikosti efektů a na závěr budeme diskutovat interpretaci velikosti efektů.

4.2 Různé způsoby měření a vyjádření velikosti efektů

Ačkoli, jak jsme výše uvedli, se začala většina kvantitativních výzkumníků zaměřovat na analýzu velikosti efektu relativně nedávno, samotné míry efektu byly v kvantitativních studiích používány již dříve: často se totiž jedná o běžně používané statistiky, jako je například korelační koeficient. Různých měř velikosti efektu existuje více než 70 (Kirk, 2003) a jsou většinou vzájemně převoditelné. Míry efektu je možné rozdělit do dvou hlavních skupin: míry efektu typu *d* (tyto míry vycházejí z porovnávání rozdílů mezi skupinami; jejich název je odvozen od slova *difference*) a míry typu *r* (tyto míry vycházejí z měř asociace a patří mezi ně i korelační koeficient; jejich název je odvozen od slova *relationship*). Míry typu *d* se přirozeně hodí k tomu, abychom pomocí nich popisovali výsledky studií založených na porovnávání skupin, což je případ experimentálních a kvaziexperimentálních studií. Naopak míry *r* najdeme, jak jejich název napovídá, u korelačních studií, které sledují vztahy mezi proměnnými.

Pro účely porovnávání efektu napříč různými studiemi jsou zvláště vhodné míry efektu, které jsou nezávislé na jednotkách, v nichž jsou měřeny závislé proměnné. Tyto míry využívají místo původních jednotek variabilitu závislé proměnné, tj. vyjadřují velikost efektu vzhledem ke směrodatné odchylce závislé proměnné. Těmto mírám se budeme věnovat ve zbytku této kapitoly.

4.2.1 Míry typu *d*

Třemi nejpoužívanějšími mírami efektu typu *d* jsou *Cohenovo d*, *Glassovo Δ* a *Hedgesovo g*. Výpočet těchto měř shrnují vzorce 4.1-4.3.

$$\text{Cohenovo } d = \frac{M_1 - M_2}{SD_{\text{sloučené}}} \quad (4.1)$$

$$\text{Glassovo } \Delta = \frac{M_1 - M_2}{SD_{\text{kontrolní}}} \quad (4.2)$$

$$\text{Hedgesovo } g = \frac{M_1 - M_2}{SD_{\text{sloučené}}^*} \quad (4.3)$$

Jak si můžeme povšimnout, všechny tři míry jsou založeny na porovnávání rozdílu průměrné úrovně závislé proměnné ve dvou skupinách ($M1 - M2$), ale liší se tím, jakým způsobem tento rozdíl standardizují. *Cohenovo d* standardizuje tento rozdíl směrodatnou odchylkou sloučeného výběrového souboru, kterou je možné vypočítat podle následujícího vzorce:

$$SD_{\text{sloučené}} = \sqrt{((SD_1^2 + SD_2^2)/2)} \quad (4.4)$$

kde SD_1 a SD_2 jsou výběrové směrodatné odchylky obou souborů.

Pokud jsou rozdíly ve variabilitě obou skupin velké, doporučuje se použít jako míry efektu *Glassovo Δ*, které standardizuje velikost efektu směrodatnou odchylkou kontrolní skupiny ($SD_{\text{kontrolní}}$), protože variabilita v této skupině by neměla být ovlivněna intervencí. Jako alternativu je v případech rozdílných variabilit možné použít také *Hedgesovo g*, které standardizuje rozdíly mezi skupinami váženým průměrem směrodatných odchylek obou skupin ($SD^*_{\text{sloučené}}$).

4.2.2 Míry typu r

Měr typu r, které jsou založeny na popisu vztahu mezi proměnnými (tj. jejich asociacemi), je velká řada a zahrnují obecně známé statistiky popisující korelace (např. Pearsonovo r, Spearmanovo ρ, či Kendallovo τ) nebo sdílenou variabilitu mezi proměnnými (např. R^2 , η^2 , či ω^2). Protože jsou tyto statistiky běžně používány pro popis vztahu mezi proměnnými, neuvádíme na tomto místě vzorce pro jejich výpočet (Kirk, 2013).

Míry typu d a r jsou vzájemně převoditelné (vzorce pro převod mezi hlavními mírami d a r uvádí např. Ellis, 2010 k převodu je také možné použít online kalkulátor <https://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html>). Protože jsou tyto míry efektu, nezávislé na jednotkách, v nichž je vyjádřena úroveň závislé proměnné (tj. toho, co má daná kampaň ovlivnit), mělo by platit, že odhad efektu z jedné studie je přímo porovnatelný s odhadem efektu z jiné studie bez ohledu na to, jak konkrétní studie měří změnu úrovně závislé proměnné. Jak si ukážeme dále, takové srovnání má smysl v rámci určité oblasti (např. pro kampaně zaměřené na změnu znalostí o klimatické změně), ale může být zavádějící, pokud srovnáváme velikosti efektů mezi různými oblastmi (např. velikosti efektů u klimatických kampaní a v epidemiologických studiích).

4.3 Přesnost odhadu efektu

Kromě odhadu velikosti efektu je třeba znát také přesnost tohoto odhadu. Je třeba si uvědomit, že náš odhad efektu kampaně nebo intervence vzniká zpravidla na základě dat z výběrového vzorku a takový odhad je nutně nepřesný (zatížený výběrovou chybou), protože zahrnuje informaci pouze o části (často

relativně velmi malé) zkoumané populace (výběrovém vzorku). Tuto nepřesnost je možné vyjádřit uvedením intervalu spolehlivosti pro daný odhad (Ellis, 2010).

Interval spolehlivosti (*confidence interval*, *CI*) je dnes v kvantitativním výzkumu často doporučovanou mírou spolehlivosti statistik (viz např. Cumming, 2011; Wilkinson & Task Force on Statistical Inference, 1999). Interval spolehlivosti je interval pro určitou statistiku (myšleno odhad populačního parametru), u nějž máme velkou důvěru, že v něm leží skutečná velikost efektu. Nejčastěji používaným intervalem je 95% interval spolehlivosti, nicméně jedná se o arbitrární hodnotu. Studie, které chtějí mít velkou míru důvěry v odhad efektu, mohou odhadnout např. 99% interval spolehlivosti (anebo ještě spolehlivější interval), naopak malé studie zatížené velkou výběrovou chybou mohou odhadnout jen 90% interval spolehlivosti. Důležité ale je si uvědomit, že množství informace obsažené v určitých datech je vždy fixní, a proto zvýšení spolehlivosti intervalu vede, při použití stejných dat, k rozšíření intervalu spolehlivosti (máme velkou jistotu, že hledaný efekt leží v daném intervalu, ale tento interval je širší). Z tohoto krátkého vysvětlení je zřejmé, že čím širší interval spolehlivosti je, tím nepřesnější je náš odhad.

Technicky vzato souvisí interval spolehlivosti s představou opakování studie: pokud bychom mnohokrát opakovali určitou studii a při každém opakování odhadli 95% CI pro velikost efektu, tak 95 % těchto intervalů bude obsahovat skutečnou hodnotu velikosti efektu. Pravděpodobnostní označení intervalu spolehlivosti tedy neříká, jaká je pravděpodobnost, že v daném intervalu leží skutečná hodnota velikosti efektu (protože tato hodnota buď v tomto intervalu leží, nebo ne⁶, nejedná se tedy o pravděpodobnost). Procentuální označení intervalu spolehlivosti říká, jak spolehlivý daný interval je, neboli jak velkou důvěru bychom měli mít v to, že obsahuje skutečnou hodnotu efektu. Spodní hranice CI je současně jakýmsi „konzervativním“ odhadem efektu (tj. můžeme mít velkou důvěru v to, že skutečný efekt dané intervence je alespoň tak velký, jako tato hodnota).

Zatímco výpočet konfidenčního intervalu pro některé míry velikosti efektu, jako je např. Pearsonovo r , je relativně snadný a je možné ho provést prakticky v jakémkoli statistickém programu, odhady konfidenčních intervalů pro jiné míry efektu, jako je např. Cohenovo d , jsou obtížné (nemají analytické řešení a není znám způsob jednoduché aproximace rozdělení dané statistiky; viz Cumming & Finch, 2001) a řeší se numericky ve statistických programech, např. v rozšířeném a volně přístupném statistickém prostředí R (<https://cran.r-project.org/web/packages/effsize/effsize.pdf>).

4.4 Presentace a interpretace velikosti efektu

⁶ To, jestli "skutečná" (tj. populační) velikost efektu leží v daném intervalu spolehlivosti, navíc není možné z výběrových dat (tj. z dat pocházejících z výběrového vzorku) zjistit, protože odhad efektu je zatížen výběrovou chybou.

Poslední důležitou otázkou, které se budeme v této kapitole věnovat je, jakým způsobem interpretovat velikost efektu a sdělit tuto informaci lidem, kteří nemají hlubší znalosti kvantitativních výzkumných metod. Důležité je totiž uvědomit si, že ačkoli jsou míry velikosti efektu jako je korelační koeficient nebo Cohenovo d velmi užitečné při porovnávání výsledků studií a umožňují akumulaci našeho vědění o efektech kampaní, širší veřejnost jim zpravidla nerozumí. Interpretace velikosti efektu je navíc hodnotovým soudem (Kirk, 2001) závislejícím mj. na kontextu studie (který zná nejlépe výzkumník sám). Z těchto důvodů je povinností výzkumníka poskytnout nejen odhad velikosti efektu a jeho přesnosti, ale také sdělit tyto výsledky jazykem, který je obecně srozumitelný a poskytnout interpretaci těchto výsledků (Ellis, 2010).

Existuje řada způsobů, jak efektivně komunikovat velikosti efektů jazykem, který je srozumitelný široké veřejnosti. Často je možné popisovat efekty kampaní v přirozených jednotkách, kterým lidé rozumí (např. průměrná absolutní nebo relativní změna množství ujetých kilometrů nebo množství ušetřeného paliva). U diskrétních stavů se dá často využít pojmu pravděpodobnosti (který lidé obecně intuitivně chápou, ačkoli za určitých mezních okolností jen zkresleně) a hovořit o absolutní nebo relativní změně pravděpodobnosti výskytu nějakého jevu.

Jinou možností pro komunikaci velikosti efektu je tzv. společný jazyk velikosti efektu (*common language effect size*; McGraw & Wong, 1992) označovaný též jako pravděpodobnost převahy (probability of superiority; Grissom, 1994). Tento způsob vyjádření velikosti efektu je založený na tom, že velikost efektu je popsána jako pravděpodobnost, že náhodně vybraný zástupce z jedné skupiny (experimentální) bude mít nižší (nebo vyšší) úroveň sledované závislé proměnné, než náhodně vybraný zástupce vybraný z druhé skupiny (kontrolní). Pokud bychom např. v případě kampaně zaměřené na znalosti klimatické změny zjistili 50% pravděpodobnost, že jedinec vystavený informační kampani bude mít vyšší znalost klimatické změny než jedinec, který této kampani nebyl vystaven (a zároveň 50% pravděpodobnost, že bude mít nižší znalost), znamenalo by to, že kampaň neměla žádný efekt. Pokud by tato pravděpodobnost byla 80%, znamenalo by to, že kampaň měla velký efekt, protože je slušná šance, že člověk vystavený kampani bude mít vyšší znalost než člověk, který nebyl kampani vystaven. Přepočet standardizovaných měr velikosti efektu (Cohenova d či korelačních koeficientů) na tyto pravděpodobnosti je velmi jednoduchý a je popsán podrobně jinde (Grissom, 1994; McGraw & Wong, 1992), a navíc může být snadno vypočítán ve volně dostupném statistickém softwaru, jako je např. statistické prostředí R (<https://cran.r-project.org/web/packages/compute.es/compute.es.pdf>).

Dalším způsobem, jak můžeme obecně srozumitelnou formou přiblížit velikost efektu je tzv. binominální zobrazení velikosti efektu (*binomial effect size display*, BESD, Rosenthal & Rubin, 1982). BESD je vlastně kontingenční tabulka, v níž jsou zobrazeny pravděpodobnosti úspěchu (zlepšení, dosažení požadovaného pozitivního výsledku apod.) pro dvě srovnávané skupiny (např. kontrolní a

experimentální), které odpovídají určité velikosti standardizované míry efektu. Výhodou BESD je, že pravděpodobnosti úspěchu mohou být snadno vypočítány z korelačního koeficientu, kdy

$$P_{\text{úspěch}} = 0,5 - \frac{r}{2} \quad (4.5)$$

Tabulka 1 ukazuje příklad takového binominálního zobrazení pro modelový příklad, kdy korelace mezi nezávislou proměnnou (např. vystavení efektu informační kampaně) a závislou proměnnou (např. ochotou snížit v zimě v obytné místnosti průměrnou teplotu o 1°C) je $r = 0,1$. Na první pohled se jedná o malou korelaci, zejména pokud bychom tento efekt interpretovali s využitím koeficientu determinace (R^2), tedy že vystavení informační kampani ovlivní jen 1 % chování souvisejícího s regulací teploty. Nicméně jak příklad uvedený v tabulce binominálního zobrazení ukazuje, i takový efekt může být prakticky významný, protože se jedná o zvýšení úspěšnosti o 10 procentních bodů. Je třeba si však uvědomit, že pravděpodobnosti zobrazené v Tabulce 1 jsou jen ilustrací velikosti efektu; skutečné pravděpodobnosti se budou od těchto pravděpodobností lišit tím více, čím větší budou průměrné rozdíly mezi pravděpodobnostmi úspěchu a pravděpodobnostmi neúspěchu (viz Randolph & Edmondson, 2005).

Tabulka 1: Příklad binominálního zobrazení efektu pro $r = 0.1$.

	$P_{\text{neúspěch}}$	$P_{\text{úspěch}}$	Celkem
Kontrolní skupina	0.55	0.45	1.00
Experimentální skupina	0.45	0.55	1.00
Celkem	1.00	1.00	–

Kromě výše zmíněných způsobů komunikace velikosti efektů, které by měly být srozumitelné i pro neškolené publikum, je samozřejmě možné, a někdy nutné, interpretovat přímo velikosti standardizovaných měr efektu, jako je Cohenovo d nebo nejpoužívanější korelační koeficient, *Pearsonovo* r . Tyto míry však většinou nemají intuitivní interpretaci; výjimkou je korelační koeficient r , jehož druhá mocnina (koeficient determinace, R^2) odpovídá podílu variability sdílené dvěma proměnnými a bývá proto interpretován jako podíl vysvětlené variability závislé proměnné vysvětlený nezávislou proměnnou. Jak uvidíme dále, někdy může takto interpretovaný R^2 dávat mylnou představu o praktickém významu efektu.

Protože často nemají samy o sobě intuitivní interpretaci, měly by být míry efektu interpretovány s ohledem na konkrétní oblast, resp. na výsledky studií, které byly provedeny v té které oblasti (anebo oblastech co nejpříbuznějších; Wilkinson & Task Force on Statistical Inference, 1999, s. 599). Výzkumník by měl být schopen sdělit svému obecenstvu – buď na základě vlastní rešerše, anebo na základě publikované meta-analýzy (jakou prezentují pro oblast klimatické změny např. Hornsey et al., 2016b) –, jestli je jím naměřený efekt srovnatelný s efekty jiných studií a zhodnotit slovně jeho relativní velikost (např. „je velmi malý“) a jeho praktický význam (např. „není prakticky významný“).

Pokud danou a ani příbuznou oblast neexistují žádné odhady efektů intervencí, které by mohly posloužit jako srovnání při interpretaci vlastní studie, je možné použít orientační kritéria pro interpretaci velikosti efektu navržená Cohenem (1988), která reprodukuje v Tabulce 2.

Tabulka 2: Orientační kritéria pro interpretaci velikosti efektu (*Cohenovo d*, biseriální korelační koeficient).

Interpretace efektu	<i>Cohenovo d</i>	Korelace (r_{pb})
Malý efekt	0.2	0,100
Střední efekt	0.5	0,243
Velký efekt	0.8	0,371

Pozn: tato tabulka je vytvořena podle kritérií, která navrhl Cohen (1988: 25 - 27).

Tato kritéria udávají intervaly standardizovaných měr efektu (r_{pb} , d) pro malé, střední a velké efekty. Původní logika této klasifikace byla taková, že střední efekty jsou dostatečně velké na to, aby byly patrné na první pohled. Od nich se liší malé i velké efekty. Ačkoli jsou tato kritéria často nekriticky využívána v empirickém výzkumu, je třeba si uvědomit, že sám Cohen (1988, s. 25–27) je chápal jako poslední východisko v situaci, kdy neexistuje možnost interpretovat velikosti efektu s ohledem na velikosti efektů typické pro určitou oblast. Nevhodné používání kritérií pro posuzování standardizovaných měr efektů bývá právě z tohoto důvodu některými autory kritizováno (Glass, McGaw, & Smith, 1981, s. 104; Shaver, 1993, s. 303; Thompson, 2008, s. 258).⁷

⁷ Zajímavou ilustraci toho, jak by mohlo být nekritické používání kritérií pro interpretaci velikosti efektu zavádějící, ukazuje často zmiňovaný (viz např. Rosnow & Rosenthal, 2003) příklad výsledků testování Salkovy vakcíny proti obrně v 50. letech, jehož se zúčastnilo přes 400 tisíc amerických dětí a jednalo se o jeden z největších epidemiologických experimentů. Výsledky této studie byly následující: zatímco v experimentální skupině dětí, které dostaly vakcínu, onemocnělo obrnou 33 dětí z 200 712 (0,016%), v kontrolní skupině onemocnělo 115 dětí z 201 114 (0,057%). Tento polní experiment vyhodnocen jako velmi úspěšný a díky jeho výsledkům se začala tato první vakcína proti dětské obrně používat v USA. Vypočítáme-li ovšem na základě

4.5 Shrnutí kapitoly

V této kapitole jsme představili koncept velikosti efektu, který je důležitý pro zhodnocení praktického významu kampaně (tj. vedle statistického efektu) a navíc nám umožňuje akumulovat naše poznání o vlivu různých nástrojů použitých v kampaních. V této kapitole jsme si řekli, že existují dva základní typy míry velikosti efektu, totiž míry typu d (založené na měření rozdílu mezi skupinami) a míru typu r (založené na mírách asociace). Míry, které spadají do těchto dvou skupin, jsou vzájemně převoditelné. Dále jsme upozornili na to, že kromě odhadu velikosti efektu je třeba brát do úvahy i přesnost odhadu velikosti efektu vyjádřenou např. jako interval spolehlivosti pro odhad velikosti efektu. V neposlední řadě jsme upozornili i na to, že míry efektu nemají většinou intuitivní interpretaci, a proto by měly být vyjadřovány způsobem, který je obecně srozumitelný. V krajním případě, kdy neexistuje možnost srovnání velikosti efektu s podobnými studiemi, je možné použít arbitrární kritéria navržená pro hodnocení standardizovaných efektů.

zmíněných statistik míry velikosti efektu, zjistíme, že jsou velmi malé ($r = 0,011$, $R^2 = 0,0001$, $d = 0,011$) a podle obecných kritérií zmíněných výše (Cohen, 1988) by byly považovány za zanedbatelné z hlediska praktické významnosti. Takový závěr by však byl chybný: vakcína více než trojnásobně snížila riziko onemocnění dětí tak vážnou chorobou, jakou je dětská obrna a zavedení této vakcíny mělo za následek to, že se průměrně předešlo 41 onemocněním na každých 100 000 dětí. Vysvětlením velmi nízkých hodnot velikosti efektu v podobných případech je to, že pravděpodobnost sledované události je velmi nízká, a proto i úspěšná intervence změní absolutní hodnoty této pravděpodobnosti jen velmi málo. Právě z tohoto důvodu je třeba kriticky vyhodnocovat odhadnuté velikosti efektů vzhledem ke kontextu studie, tzn. vzhledem k velikosti efektů alternativních intervencí (pokud existují) nebo obdobných studií v příbuzných oblastech (Wilkinson & Task Force on Statistical Inference, 1999, s. 599). Teprve pokud žádná taková kritéria neexistují, má smysl využívat velmi obecná kritéria pro hodnocení efektu navržená Cohenem (1988) nebo jiná podobná doporučení.

Rozšiřující četba

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 edition). Hillsdale, N.J: Routledge.

Cumming, G. (2011). *Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York: Routledge.

Ellis, P. D. (2010). *The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results*. Cambridge ; New York: Cambridge University Press.

Seznam literatury

American Psychological Association. (2008). *Publication manual of the American Psychological Association* (5. ed., 8. print). Washington, DC: American Psychological Association.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed). Washington, DC: American Psychological Association.

Callahan, J. L., & Reio Jr, T. G. (2006). Making subjective judgments in quantitative studies: The importance of using effect sizes and confidence intervals. *Human Resource Development Quarterly*, 17(2), 159–173. doi:10.1002/hrdq.1167

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 edition). Hillsdale, N.J: Routledge.

Cumming, G. (2011). *Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York: Routledge.

Cumming, G., & Finch, S. (2001). A Primer on the Understanding, Use, and Calculation of Confidence Intervals that are Based on Central and Noncentral Distributions. *Educational and Psychological Measurement*, 61(4), 532–574. doi:10.1177/0013164401614002

Ellis, P. D. (2010). *The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results*. Cambridge ; New York: Cambridge University Press.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills: Sage Publications.

Grissom, R. J. (1994). Probability of the superior outcome of one treatment over another. *Journal of Applied Psychology*, 79(2), 314–316. doi:10.1037/0021-9010.79.2.314

Hornsey, M. J., Harris, E. A., Bain, P. G., & Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change*, 6(6), 622–626. doi:10.1038/nclimate2943

Kirk, R. E. (1996). Practical Significance: A Concept Whose Time Has Come. *Educational and Psychological Measurement*, 56(5), 746–759. doi:10.1177/0013164496056005002

- Kirk, R. E. (2001). Promoting Good Statistical Practices: Some Suggestions. *Educational and Psychological Measurement*, 61(2), 213–218. doi:10.1177/00131640121971185
- Kirk, R. E. (2003). The importance of effect magnitude. In S. E. Davis (Ed.), *Handbook of research methods in experimental psychology* (s. 83–105). Malden, MA: Blackwell.
- Kirk, R. E. (2013). *Experimental design: procedures for the behavioral sciences* (4th ed). Thousand Oaks: Sage Publications.
- McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, 111(2), 361–365. doi:10.1037/0033-2909.111.2.361
- Osborne, J. W. (2008). Sweating the small stuff in educational psychology: how effect size and power reporting failed to change from 1969 to 1999, and what that means for the future of changing practices. *Educational Psychology*, 28(2), 151–160. doi:10.1080/01443410701491718
- Randolph, J. J., & Edmondson, R. S. (2005). Using the Binomial Effect Size Display (BESD) to present the magnitude of effect sizes to the evaluation audience. *Practical Assessment, Research, and Evaluation*, 10(14).
- Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 74(2), 166–169. doi:10.1037//0022-0663.74.2.166
- Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 57(3), 221–237. doi:10.1037/h0087427
- Shaver, J. P. (1993). What Statistical Significance Testing Is, and What It Is Not. *The Journal of Experimental Education*, 61(4), 293–316. doi:10.1080/00220973.1993.10806592
- Thompson, B. (2008). Computing and interpreting effect sizes, confidence intervals, and confidence intervals for effect sizes. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (s. 246–262). Thousand Oaks, CA: Sage.
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594–604. doi:10.1037/0003-066X.54.8.594

5. Náročnost a kvalita hodnocení dopadu kampaní

V této kapitole představujeme různé typy evaluací kauzálního efektu kampaně z hlediska jejich odborné a finanční náročnosti. Evaluační studie, které využívají složitější metodologické postupy, obvykle produkují průkaznější a přesnější odhady kauzálních efektů kampaní, avšak současně bývají nákladnější a jejich provedení je odborně náročnější. V této kapitole proto navrhuje obecné pravidlo pro přiměřenost nákladů na evaluační studie.

5.1 Finanční náklady evaluací a jejich přiměřenost nákladům kampaně

Zpracovávání evaluačních studií není obvykle rutinní a opakující se činností. Zkoumané otázky se mezi jednotlivými evaluacemi mohou výrazně lišit a spolu s tím se liší i vhodný výzkumný design a použité metody. V návaznosti na to nelze vytvořit přesný a neměnný návod, který by byl krok po kroku využitelný při evaluaci rozdílných aktivit a projektů (Sempik, Becker, & Bryman, 2007).

Optimální podíl nákladů na evaluační studie vzhledem k ceně kampaně, závisí na typu a účelu kampaně. Zatímco pilotní kampaně mohou být svým rozsahem malé, ale může být důležité hodnotit jejich dopady co nejpřesněji (a tudíž náklady evaluace budou tvořit větší podíl nákladů), u rozsáhlých kampaní mohou náklady na provedení evaluace představovat relativně menší podíl nákladů kampaně. Obecně by však mělo platit, že náklady na evaluaci kampaně jsou úměrné nákladům kampaně. Jako orientační pravidlo pro stanovení přiměřenosti nákladů evaluace doporučujeme rozmezí 5 % - 15 % nákladů kampaně. Toto pravidlo by mělo zajistit, že u skutečně velmi malých kampaní nebude evaluace prováděná a naopak u velkých a nákladných kampaní by měla být provedena relativně zevrubná evaluace jejich dopadů. Finanční náklady pro jednotlivé typy evaluací popisujeme ve zbytku této kapitoly. Pravidlo přiměřenosti se vztahuje k nákladům realizované kampaně (vč. přímo souvisejících personálních nákladů, intervence), které budou tvořit jen část nákladů projektu. Je tomu tak proto, že právě náklady kampaně mohou být považovány za ztracené, pokud kampaň nemá očekávaný efekt. Každá evaluace dopadů kampaně, bez ohledu na rozsah a způsob provedení, by měla splňovat řadu podmínek, které jsou shrnuty v Tabulce 5.1.

Tabulka 5.1: Zásady přípravvy evaluace.

1. Objektivita	Evaluace by měla být v maximální možné míře nestranná. Popis faktů musí být odlišen od subjektivních hodnotících soudů. Výsledky, jejich interpretace a doporučení mají být založené na konkrétních datech a musí být srozumitelné.
2. Nezávislost hodnotitelů	Hodnotitelé by měli mít dostatečné odborné znalosti. Důvěryhodnost je zvýšena, pokud jsou nezávislí na subjektu nebo subjektech realizujících kampaň.
3. Účast všech zúčastněných stran v procesu	Výhodou při evaluaci kampaně je, pokud evaluátor může získávat informace o průběhu kampaně od realizátora kampaně, resp. pokud je průběh kampaně dobře popsán.
4. Transparentnost a zaměření	V projektové dokumentaci kampaně má být jasně určena podoba intervence a její zamýšlené kauzální dopady. V optimálním případě by měly být definované konstrukty, které lze operacionalizovat do podoby úzce zaměřených závislých proměnných. Evaluátor by měl v hodnotící zprávě jasně popsat metodu evaluační studie.
5. Spolehlivost (reliabilita)	Pro dosažení dostatečné spolehlivosti výsledků je zapotřebí zvolit optimální výzkumný design hodnocení a používat spolehlivé nástroje pro měření závislých a nezávislých proměnných. Pro prokázání kauzálních efektů kampaně je důležité, aby evaluace měla dostatečnou statistickou sílu.
6. Ochrana zájmů zúčastněných stran	Během evaluace je třeba zajistit, aby nebyla porušena bezpečnost, důstojnost a práva účastníků evaluační studie (např. účastníků experimentálních a kvaziexperimentálních studií).

(Upraveno dle Austrian Development Agency, 2009.)

5.2 Hierarchie výzkumných designů pro evaluaci kauzálních dopadů

Podoba výzkumného designu, tedy komplexu využitých metod, se běžně považuje za nejsilnější indikátor síly prokázání kauzálního efektu kampaně. Přes dílčí rozdíly v klasifikaci jednotlivých metod, se většina autorů shoduje na hierarchii metod z hlediska jejich průkaznosti při zkoumání kauzálních vztahů (Bagshaw & Bellomo, 2008; Nutley, Powell, & Davies, 2013; Petticrew & Roberts, 2003); z této hierarchie vycházíme i my a upravujeme ji pro specifické účely této metodiky. Tito autoři se shodují na tom, že nejméně prokazatelné pro existenci kauzálních vztahů jsou analýzy postavené na případových studiích a na studiích, které nevyužívají kontrolní skupinu a randomizaci (průřezové,

panelové nebo kohortní studie).⁸ Větší sílu při prokazování kauzálního efektu mají randomizované experimenty.⁹ Největší sílu při prokazování kauzálního efektu mají rozsáhlé terénní experimenty a metaanalytické studie.¹⁰

V praxi bývají dosud často při hodnocení kampaní využívané korelační studie. Korelační analýzy však poskytují relativně slabé důkazy o existenci kauzálního efektu, protože jsou postaveny na zkoumání statistických vztahů; ty jsou nutnou, ale ne postačující podmínkou existence kauzálních vztahů (Bagshaw & Bellomo, 2008; Nutley et al., 2013; Petticrew & Roberts, 2003). Vzhledem k relativní nákladnosti korelačních studií doporučujeme jejich nahrazení experimentálními nebo kvaziexperimentálními studiemi, které je možné realizovat s podobnými náklady, ale přitom přinášejí daleko silnější evidenci o existenci kauzálních efektů.

Na základě výše uvedené hierarchie výzkumných designů navrhujeme tři typy evaluačních přístupů: orientační, základní a pokročilé evaluace (jejich shrnutí obsahuje Tabulka 5.3). Tyto tři typy evaluací mají vzrůstající schopnost prokazovat kauzální efekty kampaní a liší se také rostoucími náklady a odbornou náročností svého provedení. U každé z těchto úrovní popíšeme provedení takové evaluace na příkladech demonstračních studií realizovaných v letech 2016-2017 v rámci projektu TAČR „Nástroje hodnocení dopadu kampaní zaměřených na snížení emisí skleníkových plynů“ (TD03000282).

5.2.1 Orientační evaluace

Mezi orientační evaluace zahrnujeme přehledy odborné literatury nebo nákladnější, nicméně z hlediska prokázání kauzality nepříliš silné, korelační studie.

Úvaha o tom, jaké dopady může zamýšlená kampaň mít, nejlépe podpořená přehledem relevantní literatury, by měla být součástí návrhu jakékoli kampaně. Taková úvaha by měla zahrnovat i vymezení toho, jaké závislé proměnné má kampaň ovlivnit, jaký je mechanismus kampaně a jak velký efekt můžeme zhruba očekávat. Tato úvaha zároveň pomůže vylepšit parametry kampaně a zvýšit její dopady. Pokud má orientační evaluace poskytnout přesnější informace o velikosti potenciálního efektu kampaně, měla by být postavena na systematickém přehledu dříve realizovaných studií, anebo dokonce na metaanalýze.

Publikované studie vhodné pro přehledy a metaanalýzu je možné vyhledávat v databázích odborných článků (např. *Scopus*, www.scopus.com; *Web of Science*, www.webofknowledge.com).

⁸ V naší klasifikaci se objevují tyto postupy v rámci *orientačních evaluací* a částečně též u základních evaluací.

⁹ Tyto postupy se v naší klasifikaci objevují v rámci *základních evaluací*.

¹⁰ Terénní studie se v naší klasifikaci objevují v *pokročilých evaluacích*. Metaanalýzy řadíme však do orientačních evaluací, a to z toho důvodu, že existuje jen málo primárních studií (srv. Hornsey, Harris, Bain, & Fielding, 2016), které hodnotí efekty klimatických kampaní. Proto mohou metaanalýzy v této oblasti poskytnout jen orientační informaci o očekávaných efektech konkrétních kampaní.

Vyhledávání v databázích odborných článků je vhodné doplnit o vyhledávání volně publikovaných textů například v *Directory of Open Access Journals* (www.doaj.org) a zdroje z tzv. šedé literatury, které nemusely projít náročným recenzním procesem, a které je možné najít např. pomocí vyhledávače *Google Scholar* (scholar.google.cz). Vzhledem k rychlému vývoji poznání v oblasti klimatických kampaní je vhodné vyhledávat co nejaktuálnější odbornou literaturu.

Box 5.1: Příklad orientační evaluace

Kampaň: Evaluace se zaměřovala na možnosti využití jednoduchých informačních nástrojů pro změnu znalostí o klimatické změně a jejich prostřednictvím vedoucích ke změně postojů ke klimatické změně. Průběh evaluace: Byla provedena rešerše studií publikovaných během v posledních 15 letech v odborných časopisech, které jsou citovány v databázích *Scopus* a *Web of Science*. Výsledky rešerše byly doplněny o šedou literaturu vyhledanou pomocí vyhledávače *Google Scholar*.

Výsledky evaluace: Na základě publikovaných studií bylo zjištěno, že informace o některých aspektech klimatické změny mají slabý nebo střední efekt na změnu postojů ke klimatické změně. Jedná se o následující proměnné: znalost vědeckého konsensu o klimatické změně, znalost mechanismu klimatické změny a znalost dopadů klimatické změny. Informační kampaně, které dávají příjemcům chybějící informace o těchto aspektech klimatické změny, mohou mít slabý až střední efekt na změnu postojů ke klimatické změně. Evidence o tom, že poskytování těchto informací by mělo vliv na změnu chování, je zatím slabá. Na základě existujících studií a obecných modelů chování je možné předpokládat, že efekt na chování bude spíše slabý, maximálně střední.

Rozpočet: Provedení rešerše a její vyhodnocení trvalo přibližně 2 měsíce, na provedení rešerše se podíleli na částečný úvazek dva kvantitativní výzkumníci a odhadované náklady evaluace činí přibližně 60 tisíc Kč.

5.2.2 Základní evaluace

Základní evaluace je postavena na testování efektu kampaně prostřednictvím webových experimentů nebo laboratorních experimentů, které umožňují odhalit bezprostřední efekt kampaně ve zjednodušených podmínkách. Existence takového efektu je nutnou, avšak ne dostačující, podmínkou toho, aby intervence fungovala v reálných podmínkách a měla prakticky významný efekt. Evaluace dopadů v podmínkách laboratorních nebo webových experimentů je zpravidla méně finančně a časově náročná, než testování kampaně ve skutečných podmínkách. Laboratorní a webové experimenty však mohou mít slabou externí validitu (tj. zobecnitelnost na reálnou situaci; viz Campbell, Shadish, & Cook, 2001), i když u testování kauzálních dopadů kampaní nemusí být tento problém tak významný (Mitchell, 2012; Stone-Romero, 2002).

Náklady na získání informací od účastníků základních (i pokročilých) evaluačních studií se budou lišit v závislosti na použitém způsobu sběru dat, množství získávaných informací (např. délce dotazníku nebo experimentu, počtu opakování měření), či na tom, o jakou cílovou populaci se jedná.¹¹ Příklad provedené základní evaluace shrnuje Box 5.2.

Box 5.2: Evaluace kauzálního efektu informačního videa Jak dochází ke globálnímu oteplování.

Kampaň: Nástrojem kampaně je čtyřminutové informační animované video „Jak dochází ke globálnímu oteplování“, které popisuje fyzikální a chemický mechanismus skleníkového efektu. Původní video vytvořené a testované ve Spojených státech (Ranney & Clark, 2016a) bylo přeloženo a nadabováno do češtiny.

Očekávaný dopad kampaně: Záměrem kampaně je zvýšení znalosti mechanismu klimatické změny a sekundárně také změna postojů ke klimatické změně (Ranney & Clark, 2016a).

Průběh evaluace: Objektivní znalost mechanismu klimatické změny je měřena pomocí 32 položkové škály, která je kalibrována jako dichotomická Raschova škála; postoje ke klimatické změně jsou měřeny prostřednictvím jedno- a více položkových škál převzatých z literatury.

Výsledky evaluace: Výsledky evaluace ukazují, že informační video má středně silný vliv na znalosti mechanismu klimatické změny. Zvýšení znalostí má však jen relativně slabý efekt na postoje ke klimatické změně a zanedbatelný efekt na mitigační chování.

Rozpočet: Náklady ve výši přibližně 180 tisíc Kč zahrnovaly rekrutaci respondentů z proprietárního internetového panelu (80 tis. Kč) a náklady na přípravu experimentů a analýzu dat (100 tis. Kč).

Provedení základní (a pokročilé) evaluace, které zahrnuje vlastní výzkum, vyžaduje plán provedení evaluační studie (vč. časového plánu výzkumu, výzkumného uspořádání, měření proměnných, sběru dat a plánovaných analytických procedur), který shrnuje tabulka 5.3.

¹¹ Významnou část nákladů evaluační studie tvoří rekrutace účastníků evaluačních studií. Podle zkušeností, které máme z vlastní výzkumné praxe, se v současnosti (2015-2017) náklady na externí rekrutaci účastníka z panelu marketingové agentury (tj. externího dodavatele) pohybují u 20 minutového dotazníku přibližně od 90 Kč za účastníka výše. Pokud se jedná o složitější výzkumné uspořádání, účast je časově náročnější, anebo jsou s rekrutací spojeny další náklady (např. zvýšené odměny pro účastníky) mohou se náklady pohybovat až po několik set korun nebo dokonce několik tisíc korun.

Tabulka 5.3: Plán přípravy evaluační studie.

Aktivita	Popis
1. Konceptuálně-teoretická příprava	Na základě požadavků kladených na evaluační studii jsou identifikovány základní proměnné klíčové pro popis kauzálních dopadů kampaně, je provedena jejich operacionalizace (vytvořeny konstrukty) a jsou připraveny nástroje měření.
2. Příprava sběru dat	Přestože většina intervencí ve svém plánu operuje s koncepty cílových skupin, pro evaluaci je zapotřebí jejich přesné vymezení. Je třeba naplánovat, jak budou zástupci těchto skupin rekrutováni do evaluační studie.
3. Provedení měření před intervencí (pretest).	Před intervencí může být proveden pretest. Pretest nemusí být součástí všech studií.
4. Provedení intervence (simulace dopadu intervence)	Je provedena intervence, která simuluje působení kampaně (např. jsou účastníkům zobrazovány informace, které byly součástí původní kampaně).
5. Provedení měření po intervenci (posttest)	V některých případech bude provedeno pouze toto měření, v jiných případech (v závislosti na zvoleném výzkumném uspořádání), budou porovnávány výsledky pretestu a posttestu.
6. Vyhodnocení a interpretace výsledků	Je provedena analýza dat a odhadnut efekt intervence. Tento efekt je interpretován způsobem, který je obecně srozumitelný.

5.2.3 Pokročilá evaluace

Pokročilá evaluace je taková evaluace, která je schopná zachytit alespoň krátkodobý efekt kampaně na chování lidí, anebo jejich motivaci (tj. představy, postoje, apod.) v podmínkách, které se blíží podmínkám skutečného působení kampaně. Pokročilé evaluace budou obvykle nákladnější a jejich realizace časově náročnější, než je tomu u základních nebo orientačních evaluací. Síla evidence o kauzálních dopadech kampaně je u tohoto typu hodnocení vyšší.

Některé pokročilé evaluace budou usilovat o hodnocení delšího než jen okamžitého efektu, a proto mohou provádět opakovaná nebo odložená měření. V takovém případě budou muset být

účastníci hodnotících studií kontaktováni opakovaně. Důležitým faktorem, který ovlivňuje možnost opakovaných měření, je míra odmítnutí, resp. panelová úmrtnost, kdy účastníci opakovaných šetření z výzkumu postupně odcházejí. Míru odmítnutí a panelovou úmrtnost je možné ovlivnit vhodným využitím motivace pro respondenty, přičemž problematická může být jak příliš malá, tak příliš velká odměna pro respondenty (více informací o této problematice uvádí např. Dillman, Smyth, & Christian, 2014).

Box 5.3: Příklad evaluace efektu soutěže „Do práce na kole 2016“.

Kampaň: Kampaň Do práce na kole organizovaný každoročně spolkem Auto*mat má za cíl motivovat, formou soutěže mezi pracovními týmy, k dojíždění do práce na kole. V rámci této kampaně si pracovníci sestaví na svých pracovištích týmy o 3-5 soutěžících. Tyto týmy i jednotlivci soutěží například o největší počet najetých kilometrů do práce ve vymezeném časovém období.

Očekávaný dopad kampaně: Cílem kampaně bylo zvýšit využívání cyklistiky (a jiných nízkouhlíkových dopravních módů) k cestám do práce a propagovat tento způsob dopravy u širší veřejnosti.

Nástroj měření: Výzkum zahrnoval webový dotazník, který obsahoval jednoduchou kalkulačku uhlíkové stopy, otázky zjišťující dopravní chování a otázky zjišťující postoje k životnímu prostředí a dopravě.

Průběh evaluace: Úvodní fáze zahrnovala přípravu dotazníku a zajištění kontrolní skupiny prostřednictvím placeného panelu společnosti pro výzkum veřejného mínění. Dotazováno bylo 400 soutěžících a 400 respondentů v kontrolní skupině. Měření byla získána před začátkem kampaně, 14 dní po jejím skončení a pět měsíců po skončení kampaně. Změna chování byla odhadnuta pomocí metody rozdíl v rozdílech (viz Kapitola 3 této metodiky). Pozorování z kvaziexperimentální a kvazikontrolní skupiny byla párována pomocí metody *propensity score matching* (viz Kapitola 3 této metodiky). Párování pozorování pomocí této metody umožňuje zmenšovat chybu měření a tudíž zpřesňovat odhad kauzálního efektu.

Výsledky evaluace: Kampaň měla střední až velký efekt ve svém průběhu, ale již 14 dní po skončení začal efekt oslabovat. U dopravního chování byl nalezený znatelný nárůst jízd na kole, a to i v období několika týdnů po skončení kampaně. Realizátorovi kampaně bylo doporučeno, aby posílil nástroje vnitřní motivace účastníků, které prodlouží efekt kampaně po skončení samotné soutěže.

Rozpočet: Náklady zahrnovaly především zajištění rekrutace respondentů kontrolní skupiny (120 000 Kč) a osobní náklady související s organizační a analytickou prací, včetně provozních nákladů realizátora (300 000 Kč).

Jako náklady kampaně jsou v této metodice označovány přímé náklady související s realizací veřejné kampaně. Náklady kampaně proto zahrnují typicky náklady související s hmotnými či duchovními vstupy nutnými pro realizaci kampaně (například tiskové náklady, náklady realizace veřejných akcí či osobní náklady přímo spojené s realizací výstupů podstatných v intervenční logice kampaně). Naopak za náklady kampaně nelze považovat nepřímé náklady, administrativní náklady či osobní náklady nesouvisející bezprostředně s přípravou kampaně. Nákladové údaje jsou uvedené v cenách roku 2017 (vč. DPH) a měly by být chápány jako orientační.

Tabulka 5.4: Typy evaluací, jejich náročnost a vhodné užití.

Typ evaluace	Použité metody	Náročnost evaluace	Pro jaké kampaně by tato evaluace měla být zpracována
Orientační evaluace	Na základě přehledu literatury nebo meta-analýzy je zjištěno, jestli daná kampaň (nebo daný typ kampaně) má nebo alespoň může mít kauzální efekt a na jaké proměnné. Může být stanovena orientační velikost takového efektu, jeho bodový odhad a/nebo interval spolehlivosti pro velikost efektu.	Finanční náročnost: do 100 tis. Kč; obvykle 20-60 tis. Kč. Časová náročnost: týdny nebo několik měsíců. Odborná náročnost: jednoduché přehledy literatury může zpracovat prakticky kdokoli, složitější metaanalýzy vyžadující zpravidla statistické modelování by měl zpracovat odborník, který tyto postupy zná.	Vhodné pokud náklady kampaně přesáhnou 400 tis. Kč.
Základní evaluace	Je proveden odhad velikosti efektu na základě výsledků experimentální nebo kvaziexperimentální studie. Cílem základní evaluace je odhadnout okamžitý efekt ve zjednodušených podmínkách (např. s využitím elektronických dotazníků nebo laboratorních experimentů). Existence takového efektu je nutnou podmínkou toho, aby kampaň	Finanční náročnost: 100 tis. Kč a více v závislosti na výzkumném uspořádání a trvání studie. Časová náročnost: obvykle několik měsíců. Odborná náročnost: zpracování pokročilé evaluace vyžaduje znalosti kvantitativních metod a aplikace experimentálních nebo kvaziexperimentálních přístupů.	Vhodné pokud náklady kampaně přesáhnou 2 mil. Kč.

	mohla mít efekt i v reálných podmínkách.	Proto bude tento tyto evaluace realizovat zpravidla specializovaný evaluátor nebo kvantitativní výzkumník.	
Pokročilá evaluace	Pokročilá evaluace odhaduje velikost efektu kampaně za podmínek, které se blíží skutečným podmínkám působení kampaně, často také sleduje dlouhodobější efekt kampaně. Tento typ evaluace tedy bude využívat terénní experimenty a longitudinální kvaziexperimentální studie. Efekty kampaně budou studovány buď přímo na příjemcích kampaně, anebo na skupinách, které jsou jim velmi podobné.	Finanční náročnost: obvykle 200 tis. Kč a více v závislosti na výzkumném uspořádání a trvání studie. Časová náročnost: obvykle několik měsíců až několik let podle cílů a uspořádání studie. Odborná náročnost: zpracování pokročilé evaluace vyžaduje znalosti kvantitativních metod a aplikace experimentálních nebo kvaziexperimentálních přístupů. Proto bude tento tyto evaluace realizovat zpravidla specializovaný evaluátor nebo kvantitativní výzkumník.	Vhodné pokud náklady kampaně přesáhnou 4 mil Kč.

5.3 Shrnutí kapitoly

V této kapitole jsme představili tři úrovně evaluace kauzálních dopadů kampaní a ilustrovali jsme je na třech příkladech evaluačních studií. Orientační hodnocení je založeno na porovnání výsledků existujících studií formou přehledu nebo meta analýzy, anebo korelačních studií. Takové hodnocení dává odpověď na otázku, jestli daný typ kampaně může mít očekávané kauzální dopady a jak velké mohou být. Základní hodnocení je již postaveno na evaluačním výzkumu a využívá experimentální a kvaziexperimentální výzkumné postupy zpravidla aplikované v ideálních nebo umělých podmínkách (např. webové a laboratorní experimenty). Základní hodnocení dává odpověď na otázku, jestli daná kampaň má alespoň okamžitý efekt za ideálních podmínek. Nejvyšším stupněm evaluace jsou pokročilé evaluace, které zkoumají dopady kampaní v reálných podmínkách, v nichž kampaň působí. Takové evaluace dávají odpověď na otázku, jaké jsou skutečné dopady kampaně. Obecně platí, že

složitější evaluace přinášejí přesnější a průkaznější informace o kauzálním efektu kampaně, ale jsou náročnější z hlediska finančních prostředků, odborných nároků na hodnotitele a také z hlediska času nutného pro jejich realizaci.

Rozšiřující četba

Chen, Huey-Tsyh. (2004). Practical program evaluation: Assessing and improving planning, implementation, and effectiveness. Thousand Oaks: SAGE Publications.

Stufflebeam, D. L., Shinkfield, A. J. (2011). Evaluation theory, models , and applications. New Jersey: Jossey-Bass.

Seznam literatury

Austrian Development Agency. (2009). Guidelines for project and programme evaluations.

Organizace pro hospodářskou spolupráci a rozvoj. Získáno z

<https://www.oecd.org/development/evaluation/dcdndep/47069197.pdf>

Bagshaw, S. M., & Bellomo, R. (2008). The need to reform our assessment of evidence from clinical trials: A commentary. *Philosophy, Ethics, and Humanities in Medicine*, 3, 23.

doi:10.1186/1747-5341-3-23

Campbell, D. T., Shadish, W. R., & Cook, T. D. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (2. vyd.). Boston: Wadsworth Publishing.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. New York: John Wiley & Sons.

Hornsey, M. J., Harris, E. A., Bain, P. G., & Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change*, 6(6), 622.

doi:10.1038/nclimate2943

Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science*, 7(2), 109–117.

doi:10.1177/1745691611432343

Nutley, S., Powell, A., & Davies, H. (2013). *What counts as good evidence?* London: Research Unit for Research Utilisation.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *J*

Epidemiol Community Health, 527–529. doi:10.1136/jech.57.7.527

Ranney, & Clark, D. (2016). Climate change conceptual change: Scientific information can transform attitudes. *Topics in Cognitive Science*, 8(1), 49–75. doi:10.1111/tops.12187

Sempik, J., Becker, S., & Bryman, A. (2007). The quality of research evidence in social policy: consensus and dissension among researchers. *Evidence & Policy: A Journal of Research, Debate and Practice*, 3(3), 407–423. doi:10.1332/174426407781738047

Stone-Romero, E. F. (2002). The relative validity and usefulness of various empirical research designs. In *Handbook of research methods in industrial and organizational psychology* (s. 77–98). New Jersey: Blackwell Publishing.

6. Subjekty provádějící hodnocení kauzálních dopadů kampaní

V této kapitole popíšeme výběr realizátora evaluační studie (hodnotitele). Hodnotitelem může být odborný specializovaný evaluátor, výzkumník zabývající se kvantitativním výzkumem, nebo interní evaluátor, který je zaměstnancem organizace podílející se na tvorbě hodnocené kampaně. Každý hodnotitel by přitom splňovat následující mezinárodně uznávané podmínky (American Evaluation Association, 2013): (i) měl by mít zkušenost s výzkumnou činností orientovanou na analýzu dat; (ii) jeho odbornost by měla vyplývat z patřičného vzdělání v metodách kvantitativního výzkumu; (iii) měl by mít předpoklady pro zajištění nestrannosti a morální integrity evaluačního procesu. V této kapitole představíme základní typy evaluátorů a následně úlohy, kterým se mohou v rámci projektového cyklu věnovat.

6.1 Externí specializovaní evaluátoři

Dosavadní česká evaluační praxe byla spojena zejména s projekty podpořenými z evropských fondů, především pak projekty sociálního a vzdělávacího charakteru. Evaluace kauzálních dopadů kampaní dosud nebyla systematicky vyžadována a to ani u rozsáhlých kampaní. Metodické evaluační materiály (např. Ministerstvo práce a sociálních věcí, 2015; Ministerstvo pro místní rozvoj ČR, 2017) se primárně nezaměřovaly na hodnocení kauzálních dopadů kampaní, které jsou ale pro oblast klimatických kampaní důležité.

V České republice jsou evaluátoři kampaní a intervencí sdružení především v České evaluační společnosti. Jedná se o neziskové profesní sdružení odborníků, kteří se věnují evaluaci veřejných i soukromých rozvojových programů a projektů. Sdružení dohlíží na dodržování profesních etických standardů dle vlastního etického kodexu. Česká evaluační společnost rovněž formulovala své Formální standardy provádění evaluací a spolupracuje s časopisem *Evaluační teorie a praxe*. Více informací je dostupných na webových stránkách společnosti (www.czecheval.cz).

Evaluacemi se zabývá také celá řada zahraničních organizací, které mohou poskytnout metodická doporučení, oponentury evaluací a evaluace kampaní s mezinárodním dosahem. V Tabulce 6.1 uvádíme seznam vybraných asociací a sítí spojujících odborné hodnotitele, kteří jsou potenciálně vhodní pro realizaci evaluačních studií zaměřených na hodnocení dopadů klimatických kampaní.

Tabulka 6.1: Relevantní zahraniční asociace a sítě sdružující evaluátory programů a projektů.

Název a popis asociace nebo sítě	Webové stránky
<i>Mezinárodní asociace rozvojové evaluace</i> je dobrovolnou organizací, která sdružuje hodnotitele především z oblasti	www.ideas-global.org/IDEAS

rozvojové a environmentální mezinárodní spolupráce. Tato organizace má vlastní etický kodex.	
<i>Síť národních evaluačních společností</i> je volným sdružením zaštitěným Evropskou evaluační společností. Prostřednictvím této sítě je možné získat kontakt na evaluátory v dalších evropských zemích.	www.europeanevaluation.org/partnerships/network-european-evaluation-societies-nese
<i>Slovenská evaluační společnost</i> byla založena v roce 2006 a úzce spolupracuje s českými evaluátory v rámci evaluace programů a projektů.	www.evaluacia.sk
<i>Mezinárodní iniciativa pro evaluaci dopadů</i> nabízí velké množství case-studies, odborných materiálů a zdrojů relevantních při přípravě evaluačního programu.	www.3ieimpact.org/en/evaluation/resources/impact-evaluation-resources/
<i>Global Environment Facility - Evaluation office</i> je sdružení orientované na monitorování a evaluaci intervencí zaměřených na adaptaci a mitigaci změny klimatu. Nezaměřuje se primárně na evaluaci kauzálních efektů.	www.thegef.org/gef/eo_office
<i>Gesellschaft für Evaluation</i> je německá síť zaměřená na evaluaci projektů.	www.degeval.de
<i>American Evaluation Association</i> je hlavní americká asociace sdružující evaluátory.	www.eval.org

6.1.1 Externí akademičtí evaluátoři

Evaluaci kauzálního efektu mohou provádět akademičtí výzkumníci orientovaní na experimentální nebo obecně kvantitativní výzkum v oblasti psychologie, sociologie, nebo jim příbuzných společenských věd. Nejvýraznější rozdíl při zapojení evaluátorů z akademického prostředí oproti specializovaným evaluátorům je, že akademičtí pracovníci budou evaluace zpravidla provádět vedle svých výzkumných a pedagogických aktivit. Na jednu stranu může být výhodou jejich vysoká odborná erudovanost, na druhou stranu může být nevýhodou jejich malá časová flexibilita a fakt, že evaluace je doplňkem jejich akademických aktivit, což může mít vliv na volbu metod a stanovování cílů evaluací. Externí akademičtí mají také zájem a příležitost zveřejňovat výsledky evaluací ve vědeckých výstupech. Tabulka 6.2 uvádí seznam některých pracovišť zabývajících se sociálněvědním kvantitativním výzkumem, z nichž se hodnotitelé mohou rekrutovat.

Tabulka 6.2: Seznam některých pracovišť, která mohou evaluace provést.

Univerzitní vědecká pracoviště

Katedra psychologie, Filozofická fakulta, Univerzita Karlova (Praha)

Katedra psychologie, Filozofická fakulta, Masarykova univerzita (Brno)

Katedra psychologie, Fakulta sociálních studií, Masarykova univerzita (Brno)

Katedra psychologie, Filozofická fakulta, Univerzita Palackého (Olomouc)

Katedra psychologie, Filozofická fakulta, Ostravská univerzita (Ostrava)

Katedra sociologie, Fakulta sociálních věd, Univerzita Karlova (Praha)

Katedra sociologie, Filozofická fakulta, Univerzita Karlova (Praha)

Společenskovední modul - Fakulta humanitních studií, Univerzita Karlova (Praha)

Katedra sociologie, Fakulta sociálních studií, Masarykova univerzita (Brno)

Katedra environmentálních studií, Fakulta sociálních studií, Masarykova univerzita (Brno)

Katedra sociologie, andragogiky a kulturní antropologie - Filozofická fakulta, Univerzita Palackého (Olomouc)

Katedra sociologie, Filozofická fakulta, Ostravská univerzita (Ostrava)

Katedra statistiky a pravděpodobnosti, Fakulta informatiky a statistiky, Vysoká škola ekonomická (Praha)

Výzkumné instituty a ústavy

Sociologický ústav AV ČR, v.v.i.; www.soc.cas.cz

Psychologický ústav AV ČR, v.v.i.; www.psu.cas.cz

Ústav výzkumu globální změny AV ČR, v. v. i.; www.czechglobe.cz

Institut evaluací a sociálních analýz; www.inesan.eu

Centrum pro otázky životního prostředí UK; www.czp.cuni.cz

CERGE - Economics Institute; cz.cerge-ei.cz

6.1.2 Interní evaluátoři

Ačkoli z hlediska kvaklity a integrity evaluace by měla být preferována externí evaluace, v některých situacích je možné provést evaluaci efektů kampaně interně, tedy prostřednictvím pracovníků organizace provádějící kampaň. Využití interní evaluace je vhodné při omezených zdrojích určených na evaluace. Interní evaluátoři však mají zpravidla omezené možnosti sami realizovat složitější evaluační studii (Bourgeois, Whynot, & Thériault, 2015). Interní evaluace by proto měla být založená především na přehledech a rešerších předchozích publikovaných studií, jejichž nevýhodou však je, že poskytují jen hrubé odhady efektu konkrétní kampaně. Interní evaluace by měla být uskutečněná pouze za podmínky, že negativní dopad na evaluaci vyplývající ze střetu zájmů evaluátora je v porovnání s přínosem evaluace relativně malý. Chyba v hodnocení v důsledku střetu zájmů evaluátora

může potenciálně nastat v celém spektru situací a je poměrně obtížné se jí vyhnout či ji kontrolovat (May, 2017).

6.2 Úlohy evaluátora

Hlavními úlohami evaluátora při evaluaci kauzálních dopadů je: (i) správně identifikovat klíčové proměnné pro popis dopadů kampaně; (ii) zvolit správné výzkumné uspořádání pro evaluační studii; (iii) zajistit či zprostředkovat sběr dat; (iv) vyhodnotit a interpretovat data z evaluační studie. Evaluační aktivity však mohou zahrnovat i jiné aktivity, které se týkají všech fází projektového cyklu a zahrnují přípravu projektové výzvy a samotných kampaní, průběžné a závěrečné hodnocení, či oponenturu evaluace (Kazimirski, 2014; Rogers, Peersman, & Macfarlan, 2016). Tyto dodatečné aktivity nyní stručně popíšeme.

6.2.1 Příprava projektové výzvy

Role evaluátora je podstatná již při nastavování podmínek projektové výzvy zadavatelem. Hodnocení dopadů kampaní by mělo být provázáno s projektovou výzvou. Vyhlašovaná projektová výzva by proto měla obsahovat výčet konceptů a jejich indikátorů, na jejichž změnu se má kampaň zaměřit. Takový výčet je důležitý nejen pro správné nastavení cílů kampaní, ale také pro pozdější hodnocení jejich kauzálního dopadu. Při hodnocení projektů je následně možné bodově hodnotit, zda navrhovaná kampaň cílí na tyto jevy, jestli návrh projektu zahrnuje evaluaci dosažení předpokládaného kauzálního dopadu, a jak kvalitně má být tato evaluace provedena.

6.2.2 Role hodnotitele při přípravě kampaní

Při přípravě návrhu kampaně se může hodnotitel podílet na následujících úkolech: (i) na identifikaci a hodnocení hlavních cílů kampaně; (ii) na výběru nástrojů použitých v kampani k dosažení plánovaných kauzálních dopadů; (iii) na přípravě evaluačního plánu a výběru indikátorů pro hodnocení dopadů kampaně (Patton, 2011; Tannahill & Sridharan, 2013).

V případě kampaní s delším průběhem, anebo nákladných kampaní, je vhodné hodnotit kauzální dopady už v průběhu realizace kampaně. Spolupráce mezi evaluátorem a realizátorem kampaně by v těchto případech měla probíhat již během přípravy kampaně (Uitto, Juha I., van den Berg, Rob D., & Puri, Jyotsna, 2017; Urban, Hargraves, & Trochim, 2014). Úvodní konzultace zároveň umožňuje nastavit sběr dat v rámci projektu s ohledem na zamýšlenou podobu vyhodnocení a reportování dopadů kampaně.

6.2.3 Průběžná a závěrečná evaluace projektů

Rozsáhlé programy či projekty by měly být rozdělené do časově oddělených fází, určených například jako roční cykly. Pro lepší dosažení cílů celého projektu je přínosné provést evaluaci efektu jedné fáze a na základě této evaluace upravit nastavení fáze následující (Ensminger, Kallemeyn, Rempert, Wade, & Polanin, 2015). Evaluátor identifikuje nedostatky dle zjištěných kauzálních vazeb a případně navrhne možné úpravy designu kampaně (Hughes, & Baumgartl, 2005). Vzhledem k vysokým nárokům na finanční a personální zabezpečení lze průběžnou evaluaci očekávat především u nákladných a dlouhodobých programů. Není-li z finančních či časových důvodů možné zajistit průběžnou evaluaci kampaně, uskuteční se základní či pokročilá evaluace kauzálních efektů po skončení kampaně.

6.2.4 Oponentní hodnocení evaluace

Zprávy o hodnocení dopadů kampaně mohou být dále hodnoceny nezávislým oponentem nebo oponenty. Takový oponentní posudek je vhodný pokud může mít hodnocení potenciálně velké dopady pro realizátora kampaně, anebo pokud bude mít zásadní vliv na kampaně prováděné v budoucnu. Cena vypracování oponentního posudku představuje jen zlomek obvyklé ceny evaluace a může přitom fungovat jako pojistka proti chybně provedeným evaluacím. Oponentní hodnocení je cenné zvláště v případech, kdy se realizátor přikloní k vnitřní evaluaci. Oponentura může zmírnit řadu nevýhod vnitřní evaluace, včetně její nižší věrohodnosti. Oponentní hodnocení evaluace je možné, pokud evaluátor zachová datové soubory a zaznamená podrobnou informaci o využitých metodách práce s daty a využitých statistických operacích. Pokud jím využití metody práce s daty a statistické operace nejsou běžné či standardní, je zapotřebí zaznamenat popis celého jejich mechanismu anebo odkaz na jeho vysvětlení.

6.3 Shrnutí kapitoly

V této kapitole jsme popsali výběr evaluátora a výhody a nevýhody provedení evaluační studie externími a interními evaluátory. Dále jsme v této kapitole diskutovali jednotlivé fáze evaluace a místo evaluace v projektovém cyklu. Externí evaluace by měla být preferovaným způsobem zadávání evaluace, pokud tomu nebrání finanční nebo jiná omezení. Interní evaluace je naopak realizována subjektem, který se podílel na provedení kampaně. Vhodná je pouze v ojedinělých případech: pro evaluace s malým rozpočtem, které mají přinést jen základní informaci o kauzálním efektu kampaně, a u nichž je zadavatel evaluace ochoten přijmout zvýšené riziko, že evaluace může být ovlivněna potenciálním střetem zájmů.

Rozšiřující četba

Hendl, J., Remr, J. (2017). *Metody výzkumu a evaluace*. Praha: Portál.

Newcomer, K. E., Hatry, H. P., Wholey, J. S. (eds.) (2015). *Handbook of practical program evaluation*. 4. vydání. New Jersey: Jossey-Bass.

Smith, Nick L., Brandon, Paul R. (2007). *Fundamental issues in evaluation*. New York: Guilford Press.

Seznam literatury

American Evaluation Association. (2013). Guiding principles for evaluators. American Evaluation Association. Získáno z <http://www.eval.org/p/cm/ld/fid=51>

Bourgeois, I., Whynot, J., & Thériault, É. (2015). Application of an organizational evaluation capacity self-assessment instrument to different organizations: Similarities and lessons learned. *Evaluation and Program Planning*, 50, 47–55. doi:10.1016/j.evalprogplan.2015.01.004

Ensminger, D. C., Kallemeyn, L. M., Rempert, T., Wade, J., & Polanin, M. (2015). Case study of an evaluation coaching model: Exploring the role of the evaluator. *Evaluation and Program Planning*, 49, 124–136. doi:10.1016/j.evalprogplan.2015.01.002

Hughes, J., & Baumgartl, B. (2005). *Project manager's guide to evaluation*. Wien: Navreme.

Kazimirski, A. (2014). *Building your measurement framework: NPC'S four pillar approach*. Získáno z <http://www.impact-value.com/attachments/File/NPCs-four-pillar-approach-FINAL.pdf>

May, A. (2017). Conflict of interest in research. In *The SAGE encyclopedia of communication research methods*. Thousand Oaks: SAGE.

Ministerstvo práce a sociálních věcí. (2015). Základní pojmy pro přípravu a realizaci evaluace inovačního projektu. Získáno z <https://www.esfcr.cz/documents/21802/797034/Z%C3%A1kladn%C3%AD+pojmy+k+evaluaci+inova%C4%8Dn%C3%ADch+projekt%C5%AF/9d30263d-3c57-49b0-a441-da03c6935722>

Ministerstvo pro místní rozvoj ČR. (2017). Evaluace evropských strukturálních a investičních fondů. Ministerstvo pro místní rozvoj ČR. Získáno z <http://dotaceeu.cz/cs/Fondy-EU/Narodni-organ-pro-koordinaci/Evaluace>

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York: Guilford Press.

Rogers, P., Peersman, G., & Macfarlan, A. (2016). Approaches for better evaluation. Získáno z <http://www.betterevaluation.org/en/approaches>

Tannahill, C., & Sridharan, S. (2013). Getting real about policy and practice needs: Evaluation as a bridge between the problem and solution space. *Evaluation and Program Planning*, 36(1), 157–164. doi:10.1016/j.evalprogplan.2012.03.002

- Uitto, Juha I., van den Berg, Rob D., & Puri, Jyotsna. (2017). *Evaluating climate change action for sustainable development*. New Delhi: Springer Open. Získáno z <http://www.gefio.org/sites/default/files/ieo/documents/files/cc-action-for-sustainable-development.pdf>
- Urban, J. B., Hargraves, M., & Trochim, W. M. (2014). Evolutionary evaluation: implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate. *Evaluation and Program Planning*, 45, 127–139. doi:10.1016/j.evalprogplan.2014.03.011

Přílohy

Součástí příloh metodiky jsou popisy škál měřících znalosti klimatické změny (Příloha 1), postoje ke klimatické změně (Příloha 2), přesvědčení o aspektech klimatické změny (Příloha 3), záměr mitigačního a adaptačního chování (Příloha 5), podporu politik (Příloha 6) a individuální uhlíkovou stopu (Příloha 6).

1. Měření znalostí klimatické změny

1.1 Subjektivní znalost

1.1.1 Škála k měření subjektivní znalosti

Jak byste ohodnotil/a svou znalost problematiky globálního oteplování? (Urban et al., připravováno k publikaci)

Možnosti odpovědí: 11 bodová škála s olabelovanými polárními kategoriemi 0 (velmi špatná znalost) – 10 (vynikající znalost).

1.1.2 Škála k měření subjektivní znalosti

Kolik si myslíte, že víte o globálním oteplování? (Boyes et al., 2009)

Možnosti odpovědí:

- a) O globálním oteplování toho vím hodně.
- b) O globálním oteplování toho vím něco.
- c) O globálním oteplování toho vím málo.
- d) O globálním oteplování nevím nic.

1.2 Objektivní znalost

1.2.1 Škála k měření objektivní znalosti (Urban et al., připravováno k publikaci)

Jedná se o škálu, která obsahuje 32 výroků o mechanismu klimatické změny. Škála je kalibrována na základě Raschova modelu. Možnosti odpovědí: pravda, nepravda, nevím.

- i. Metan je skleníkový plyn. (Pravda. Dále uváděno jako „P“.)
- ii. Oxid uhličitý je skleníkový plyn. (P)
- iii. Skleníkové plyny jsou podobné sklu ve skleníku v tom, že omezují unikání tepelné energie ze zemské atmosféry. (P)
- iv. Současná klimatická změna souvisí s nárůstem koncentrace skleníkových plynů v atmosféře. (P)
- v. Průměrná teplota při zemském povrchu v uplynulých desetiletích roste. (P)

- vi. Mezi hlavní lidské zdroje skleníkových plynů patří spalování paliv, zemědělská činnost a odlesňování. (P)
- vii. Klima popisuje dlouhodobé vzorce počasí v určitém místě. (P)
- viii. Skleníkový efekt je přirozený atmosférický jev. (P)
- ix. Záření přicházející od Slunce k Zemi se liší vlnovou délkou od záření, které odchází ze Země do vesmíru. (P)
- x. Záření, které dopadá ze Slunce na Zemi, není stejné jako záření, které odchází ze Země do vesmíru. (P)
- xi. Hlavním zdrojem energie pro planetu Zemi je energie záření ze Slunce. (P)
- xii. Skleníkový efekt umožňuje na Zemi život v podobě, v jaké ho známe. (P)
- xiii. Skleníkové plyny jsou tvořeny molekulami, které zachytávají infračervené záření. (P)
- xiv. Ze Země je do vesmíru vyzařováno převážně infračervené záření. (P)
- xv. Atmosféra pohlcuje tepelné záření vyzařované povrchem Země. (P)
- xvi. Když na Zemi dopadne více energie, než z ní odejde, planeta se otepluje. (P)
- xvii. Skleníkové plyny způsobují smog ve městech. (Nepravda. Dále uváděno jako „N“.)
- xviii. To, čemu se dnes říká klima, se dříve nazývalo počasím. (N)
- xix. Ozonová vrstva způsobuje globální oteplování. (N)
- xx. Oxid siřičitý je řazen mezi významné skleníkové plyny. (N)
- xxi. Atmosférický dusík je skleníkový plyn. (N)
- xxii. Ke skleníkovému efektu dochází proto, že skleníkové plyny urychlují vyzařování tepelné energie ze Země do vesmíru. (N)
- xxiii. Ze Slunce přichází na Zemi převážně infračervené záření a ze Země do vesmíru odchází především viditelné světlo. (N)
- xxiv. Skleníkové plyny se vyznačují tím, že silně pohlcují viditelné světlo. (N)
- xxv. Všechny plyny v atmosféře jsou skleníkovými plyny. (N)
- xxvi. Všechny skleníkové plyny způsobují úbytek ozonové vrstvy. (N)
- xxvii. Skleníkové plyny získaly svůj název od toho, že je najdeme hlavně ve sklenících. (N)
- xxviii. Ke skleníkovému efektu může docházet pouze ve skleníku. (N)
- xxix. Ke globálnímu oteplování dochází proto, že se snižuje množství skleníkových plynů v atmosféře. (N)
- xxx. Globální oteplování je způsobené hlavně vulkanickou činností Země. (N)
- xxxi. Globální oteplování je způsobeno hlavně zvýšenou aktivitou Slunce. (N)
- xxxii. Změna klimatu je způsobena střídáním ročních období. (N)

1.2.2 Škála k měření objektivní znalosti (Ashworth et al., 2011)

Jedná se o škálu se dvěma položkami (někdy bývají analyzovány odděleně jako indikátory různých dimenzí znalostí).

- I. Skleníkové plyny způsobují změnu klimatu.
- II. Díra v ozónové vrstvě přispívá ke změně klimatu.

Možnosti odpovědí: pětibodová škála s olabelovanými polárními kategoriemi a střední kategorií 1 (rozhodně nesouhlasím), 2, 3 (ani souhlasím, ani nesouhlasím), 4, 5 (rozhodně souhlasím).

1.3 Představy o vědeckém konsensu

1.3.1 Škála k měření představ o vědeckém konsensu (Ding et al., 2011; Leiserowitz et al., 2014)

Který výrok se nejvíce blíží vašemu názoru?

Výběr jedné z následujících možností odpovědí:

- a) Většina vědců si myslí, že globální oteplování probíhá.
- b) Vědci se neshodují na tom, zda globální oteplování probíhá či neprobíhá.
- c) Většina vědců si myslí, že globální oteplování neprobíhá.
- d) Nevím dost, abych mohl/a odpovědět.

1.3.2 Škála k měření představ o vědeckém konsensu (Linden et al., 2014)

Zkuste prosím podle svého nejlepšího vědomí odhadnout, jaké procento klimatologů se domnívá, že dochází ke klimatické změně způsobené člověkem.

Svůj odhad uveďte prosím v procentech (%).

2. Postoje ke klimatické změně

2.1 Škála pro měření skepticismu (Brügger et al., 2016; tato studie uvádí $\alpha = ,79$)

Do jaké míry souhlasíte s následujícími výroky?

Možnosti odpovědí: pětibodová škála s olabelovanými polárními kategoriemi 1 (rozhodně nesouhlasím) - 5 (rozhodně souhlasím).

- I. Klimatická změna je způsobena pouze přírodními procesy.
- II. Experti se shodují, že klimatická změna je skutečný problém. (rekódovat)
- III. Média jsou často příliš alarmistická o záležitostech týkajících se klimatické změny.
- IV. Důkazy o klimatické změně jsou nespolehlivé.
- V. Nejsem si jistý/ jistá, zda dochází ke klimatické změně.
- VI. Nevěřím, že klimatická změna je skutečný problém.

3. Přesvědčení

3.1 Přesvědčení o existenci klimatické změny

3.1.1 Škála k měření představ o existenci klimatické změny (Linden et al., 2014)

Do jaké míry jste přesvědčen(a), že v současnosti dochází nebo nedochází ke klimatické změně (tj. bez ohledu na její příčiny)?

Možnosti odpovědí: 11 bodová škála s olabelovanými polárními kategoriemi a střední kategorií 0 (jsem rozhodně přesvědčen(a), že nedochází ke klimatické změně) – 5 (nejsem si jistý/ jistá, že dochází ke klimatické změně) – 10 (jsem rozhodně přesvědčen/a, že dochází ke klimatické změně).

3.1.2 Škála k měření představ o existenci klimatické změny (Leiserowitz et al., 2014)

Co si myslíte? Probíhá globální změna klimatu?

Možnosti odpovědí: ano, ne.

3.1.3 Škála k měření představ o existenci klimatické změny (Poortinga et al., 2011)

Myslíte si osobně, že světové klima se mění nebo ne?

Možnosti odpovědí: ano, ne, nevím.

3.2 Přesvědčení o antropogenních příčinách klimatické změny

3.2.1 Škála k měření přesvědčení o antropogenních příčinách klimatické změny (Linden et al., 2014)

Za předpokladu, že dochází ke klimatické změně, do jaké míry je vyvolána lidskými aktivitami, přirozenými změnami, nebo kombinací obou těchto faktorů?

Možnosti odpovědí: participant odpovídají na 11-bodové škále [0] - [10]

[0] = Věřím, že klimatická změna je zcela vyvolána přirozenými změnami prostředí

[5] = Věřím, že klimatická změna je vyvolána stejným dílem přirozenými změnami prostředí a lidskou činností.

[10] = Věřím, že klimatická změna je zcela vyvolána lidskou činností.

3.2.2 Škála k měření přesvědčení o antropogenních příčinách klimatické změny (Bolsen et al., 2013)

Pokud dochází ke globálnímu oteplování, do jaké míry si myslíte, že je způsobeno lidskou činností, na rozdíl od přirozených změn v životním prostředí?

Možnosti odpovědí:

a) Určitě lidskou činností

b) Velmi pravděpodobně lidskou činností

c) Pravděpodobně lidskou činností

- d) Ani lidskou činností, ani přirozeně
- e) Pravděpodobně přirozeně vyvolané
- f) Velmi pravděpodobně přirozeně vyvolané
- g) Rozhodně přirozeně vyvolané

3.2.3 Škála k měření přesvědčení o antropogenních příčinách klimatické změny (Aitken et al., 2011)

Do jaké míry se domníváte, že lidská činnost přispívá ke změně klimatu?

Možnosti odpovědí: 1 (vůbec), 2, 3, 4, 5 (hodně).

3.3 Přesvědčení o načasování dopadů klimatické změny

3.3.1 Škála k měření přesvědčení o načasování dopadů klimatické změny (McCright et al., 2013, 2014)

Který z následujících výroků odráží váš názor na to, kdy se začnou projevovat dopady globálního oteplování?

Možnosti odpovědí:

- a) již se začaly projevovat
- b) začnou se projevovat během několika let
- c) začnou se projevovat v průběhu Vašeho života
- d) nebudou se projevovat v průběhu Vašeho života, ale budou mít vliv na budoucí generace
- e) nikdy nenastanou

3.3.2 Škála k měření přesvědčení o načasování dopadů klimatické změny (Ding et al., 2011)

Kdy si myslíte, že globální oteplování začne poškozovat lidi v České republice (příp. ve světě)?

Možnosti odpovědí:

- a) Nikdy
- b) za 100 let
- c) za 50 let
- d) za 25 let
- e) za 10 let
- f) poškozuje je už nyní

3.4 Přesvědčení o dopadech klimatické změny

3.4.1 Škála k měření přesvědčení o dopadech klimatické změny (Ashworth et al., 2011; Ding et al., 2011; Leiserowitz et al., 2014)

Jak moc podle Vašeho názoru globální změna klimatu poškodí...

Možnosti odpovědí:

a) vůbec, b) málo, c) středně, d) hodně

... svět jako celek

... Vás osobně

... lidi na celém světě

... přírodu

... oblast kde se nacházíte

... budoucí generace

... Českou republiku

3.4.2 Škála k měření přesvědčení o dopadech klimatické změny (Zhao et al., 2011)

Jak moc podle Vašeho názoru globální změna klimatu poškodí...

Možnosti odpovědí:

a) vůbec, b) málo, c) středně, d) hodně

... Vaši rodinu

... lidi v České republice

... lidi v dalších rozvinutých zemích

... lidi v rozvojových zemích

... rostlinné a živočišné druhy

3.4.3 Škála k měření přesvědčení o dopadech klimatické změny (Dietz et al., 2007; tato studie uvádí $\alpha = ,92$)

Jak je podle Vás pravděpodobné, že klimatická změna s sebou přinese tyto následky?

Možnosti odpovědí: 1 (velmi nepravděpodobné), 2 (poněkud nepravděpodobné), 3 (poněkud pravděpodobné), 4 (velmi pravděpodobné)

- I. Zdravotní problémy v České republice se zvýší.
- II. Zdravotní problémy ve světě se zvýší.
- III. Životní standard mnoha lidí v České republice poklesne.
- IV. Životní standard mnoha lidí ve světě poklesne.
- V. Množství vyhynulých druhů v České republice se zvýší.
- VI. Množství vyhynulých druhů ve světě se zvýší.

3.4.4 Škála k měření přesvědčení o dopadech klimatické změny (Hornsey & Fielding, 2016; tato studie uvádí $\alpha = ,90$)

Do jaké míry souhlasíte, či nesouhlasíte s následujícími výroky?

Možnosti odpovědí: 1 (silně nesouhlasím), 2, 3, 4, 5, 6 (silně souhlasím)

- I. Změna klimatu bude mít zřetelně negativní dopad na mé zdraví v příštích 25 letech
- II. Změna klimatu bude mít zřetelně negativní dopad na mou ekonomickou a finanční situaci v příštích 25 letech
- III. Změna klimatu bude mít zřetelně negativní dopad na životní prostředí, v němž žije moje rodina a já

4. Znepokojení klimatickou změnu

4.1 Škála k měření znepokojení klimatickou změnou (Linden et al., 2014)

Uvedte prosím na škále od 0 do 10, do jaké míry se obáváte klimatické změny.

Možnosti odpovědí: participanté odpovídají na 11-bodové škále [0] - [10]

[0] = Vůbec se neobávám.

[5] = Středně se obávám.

[10] = Velmi se obávám.

4.2 Škála k měření znepokojení klimatickou změnou (McCright et al., 2013, 2014)

Do jaké míry se vy osobně obáváte globálního oteplování?

Možnosti odpovědí:

- a) vůbec
- b) málo
- c) středně
- d) velmi

4.3 Škála k měření znepokojení klimatickou změnou (Corner et al., 2011; tato studie uvádí $\alpha = ,83$)

- I. Do jaké míry, jestli vůbec, jste znepokojeni klimatickou změnou?
- II. S ohledem k případným možným dopadům změny klimatu, které by se mohly dopadat na vás osobně, jak jste znepokojeni, pokud vůbec, změnou klimatu?
- III. S ohledem k případným možným dopadům změny klimatu, které by se mohly dopadat na společnost, jak jste znepokojeni, pokud vůbec, změnou klimatu?

Možnosti odpovědí:

- a) Vůbec nejsem znepokojen/a
- b) Ne příliš znepokojen/a
- c) Docela znepokojen/a
- d) Velmi znepokojen/a
- e) Nevím
- f) Žádný názor

5. Záměr chování

5.1 Škála pro měření záměru mitigačního chování (Bolsen et al., 2013)

Do jaké míry jste ochotni podniknout osobní kroky ke snížení vlastních emisí skleníkových plynů?

Možnosti odpovědí:

- a) Mimořádně neochotný
- b) Středně neochotný
- c) Trochu neochotný
- d) Ani ochotný, ani neochotný
- e) Trochu ochotný
- f) Středně ochotný
- g) Mimořádně ochotný

5.2 Škála pro měření záměru mitigačního chování (Brügger et al., 2016; tato studie uvádí $\alpha = ,85$)

Jak je pravděpodobné, nebo nepravděpodobné, že v budoucnu učiníte následující kroky v boji proti změně klimatu?

Možnosti odpovědí: 1 (velmi nepravděpodobně), 2, 3, 4, 5 (velmi pravděpodobně)

- i. Vybrat si auto s dobrou spotřebou paliva.
- ii. Nainstalovat (více) zateplení ve své domácnosti.
- iii. Sdílení automobilu/Využití spolujízdy.
- iv. Častější využívání veřejné dopravy.
- v. Více chodit pěšky nebo jezdit na kole.
- vi. Vyměnit starší spotřebiče za energeticky efektivnější nové modely (např. lednici).
- vii. Připojit se k environmentální skupině.
- viii. Offsetovat své lety.¹²

¹² Příklad vysvětlení položky pro respondenty: Offsetování je způsob, jak vyvážit své emise přispívající ke změně klimatu odpovídajícím opatřením, které ji zmírňuje. Ptáme se vás tedy na to, zda své konkrétní letecké cesty vyvažujete dodatečnými na ně přímo navázanými opatřeními. Může jít například o platbu organizaci, která se offsetováním zabývá nebo třeba vysazením odpovídajícího počtu stromů, které by se jinak nevysadili.

- ix. Jíst méně masa.
- x. Snížit množství nakupovaných nových věcí.
- xi. Strávit dovolenou v ČR spíše než v zahraničí.
- xii. Požádat svého voleného zástupce, aby podpořil silný zákon v boji proti změně klimatu.

5.3 Škála pro měření záměru adaptačního chování (Brügger et al., 2016; tato studie uvádí $\alpha = ,77$)

Jak je pravděpodobné, nebo nepravděpodobné, že v budoucnu učiníte následující kroky k adaptaci na změnu klimatu?

Možnosti odpovědí: 1 (velmi nepravděpodobně), 2, 3, 4, 5 (velmi pravděpodobně)

Přebarvit si svůj (budoucí) příbytek ve světlejší barvě (menší absorpce tepla v létě).

Zakoupit si pojištění proti povodním pro svojí (budoucí) domácnost.

Nainstalovat systém opětovného využití vody ve své domácnosti (vyhnout se nedostatku vody během sucha).

Darovat peníze na záchranu druhů ohrožených změnou klimatu.

Přesvědčit příbuzné nebo přátele, aby se přestěhovali ze záplavové oblasti.

Nainstalovat na svou toaletu zařízení pro úsporu vody.

Přečíst si o tom, jak se vyhnout stresu z horka v průběhu vln horka.

Zjistit, jak moc je Váš příbytek ohrožen záplavami.

Darovat peníze na projekty v rozvojových zemích, které přemísťují obytné čtvrti z ohrožených oblastí.

6. Podpora politik

6.1 Škála pro měření podpory mitigačních politik v ČR (Urban et al., připravováno k publikaci)

Tato škála je složená ze 13 výroků, z nichž některé jsou převzaty z jiných publikovaných škál. V rámci škály je použito několik různých formátů odpovědí, proto je vhodné použít pro kalibraci např. partial credit model.

(i) Lidé mají rozdílné názory na to, jestli a jaké politiky na ochranu klimatu by se měly zavádět. Tyto názory občanů by měly být zohledněny. Politici však často tyto názory neznají, protože se na ně občanů nikdo neptá. Do jaké míry souhlasíte nebo nesouhlasíte s následujícími politikami?

Možnosti odpovědí:

- a) Rozhodně nesouhlasím
- b) Spíše nesouhlasím
- c) Ani nesouhlasím, ani souhlasím
- d) Spíše souhlasím

e) Rozhodně souhlasím

f) Nevím/nedovedu odpovědět

(ii) Souhlasíte s tím, že by Česká republika měla snížit své emise skleníkových plynů (oxidu uhličitého, metanu a jiných), o nichž se tvrdí, že způsobují globální oteplování? (podle Leiserowitz, 2006)

(iii) Na pařížské konferenci OSN o ochraně klimatu COP 21 se Česká republika zavázala, že sníží do roku 2030 své emise skleníkových plynů o 40% oproti úrovni roku 1990. Souhlasíte s přijetím tohoto závazku Českou republikou? (upraveno podle Leiserowitz, 2006)

(iv) V rámci Evropské uni funguje trh s povolenkami na vypouštění skleníkových plynů. V rámci tohoto systému je stanoven limit pro celkový objem skleníkových emisí vypouštěných všemi zeměmi EU dohromady. Každá země dostane povolení vypouštět část tohoto celkového objemu. Pokud firmy v určité zemi vypustí více emisí, než kolik jim jejich stát umožňuje prostřednictvím přidělených povolenek, musí si zakoupit práva na vypouštění emisí od jiných firem, anebo zaplatit vysokou pokutu. Souhlasíte nebo nesouhlasíte s existencí takového evropského trhu umožňujícího firmám prodávat a nakupovat uhlíkové povolenky a současně omezit celkový objem emisí v celé EU? (upraveno podle Ding et al., 2011)

(v) Česká republika se společně s dalšími členskými zeměmi Mezinárodní organizace pro civilní letectví v roce 2016 zavázala ke snižování emisí z letecké dopravy, které v současnosti odpovídají 2% globálních emisí skleníkových plynů. V důsledku povedou zavedená opatření k tomu, že celkové emise skleníkových plynů z letecké dopravy klesnou do roku 2050 o 50% vzhledem k roku 2005. Je pravděpodobné, že toto opatření současně povede k nárůstu cen letecké dopravy. Souhlasíte s takovým závazkem České republiky?

(vi) Do jaké míry souhlasíte nebo nesouhlasíte se zvýšením daně na benzín naftu a jiná fosilní paliva do automobilů o 40%? Cílem tohoto zvýšení daně by bylo motivovat lidi k omezování jízdy automobilem a snižování emisí oxidu uhličitého. (upraveno podle Leiserowitz, 2006)

(vii) Někteří lidé navrhují, aby se zavedla dodatečná energetická daň pro podniky, která by podniky motivovala k šetrnějšímu využívání energie. Taková daň by zvýšila ceny většiny věcí, které kupujete, včetně jídla a oblečení, o 3%. Souhlasil/a byste se zavedením takové energetické daně pro podniky? (upraveno podle Leiserowitz, 2006)

(viii) Souhlasil/a byste s tím, aby byly v ČR zavedeny dotace pro občany na nákup elektromobilů a plug-in hybridů do 3,5 tuny ve výši 200 000 Kč na zakoupený automobil? Finanční prostředky na tyto dotace by byly získány zvýšením daně na benzín, naftu, plyn a jiná fosilní paliva. (upraveno podle Ding et al., 2011)

(ix) Souhlasil/a byste se zavedením dodatečné uhlíkové daně na benzín, naftu a další fosilní paliva do automobilů ve výši 15%, kdyby tyto prostředky byly následně využity na snížení daně z příjmu fyzických osob? (upraveno podle Ding et al., 2011)

(x) V současné době se diskutuje o tom, že systém emisních povolenek bude rozšířen i na fosilní paliva do automobilů. V praxi by takové opatření znamenalo, že ceny benzínu, nafty a dalších fosilních paliv pro automobily by vzrostly přibližně o 25%. Souhlasil/a byste se zavedením takového opatření?

(xi) Domácnosti v ČR v současnosti platí 4950 Kč (bez DPH) za každou odebranou megawatthodinu elektřiny na podporu obnovitelné energie. Tento příplatek činí více než 10% průměrné ceny elektřiny dodávané domácnostem. Pál/a byste si, aby...

Možnosti odpovědí:

- a) ...se tento příplatek snížil
- b) ...zůstal stejný
- c) ...zvýšil se
- d) Nevíte/ nedovedete odpovědět.

(xii) V současnosti je cca 45% elektrické energie v ČR vyráběno v uhelných elektrárnách. Jakou z následujících variant byste preferoval pro příštích 5 let...

Možnosti odpovědí:

- a) Snížit podíl energie vyráběné z uhlí.
- b) Ponechat podíl energie vyráběné z uhlí na současné úrovni.
- c) Zvýšit podíl energie vyráběné z uhlí.
- d) Nevíte/ nedovedete odpovědět.

(xiii) Dal(a) byste v senátních volbách hlas kandidátovi, který by zpochybňoval, že klimatická změna je způsobená člověkem? Takového kandidáta/kandidátku byste...

Možnosti odpovědí:

- a) Rozhodně nepodpořil(a)

- b) Spíše nepodpořil(a)
- d) Spíše podpořil(a)
- e) Rozhodně podpořil(a)
- f) Nevíte/ nedovedete odpovědět

7. Měření uhlíkové stopy (uhlíková kalkulačka)

K odhadu individuální uhlíkové stopy doporučujeme použít kalkulačku vytvořenou v rámci projektu TAČR „Nástroje pro hodnocení dopadu kampaní zaměřených na snížení individuálních emisí skleníkových plynů" (č. TD03000282), která je zdarma dostupná na internetové adrese:
<http://uhlikovastopa.cz/kalkulacka>.

Seznam literatury citované v přílohách:

- Aitken, C., Chapman, R., & McClure, J. (2011). Climate change, powerlessness and the commons dilemma: Assessing New Zealanders' preparedness to act. *Global Environmental Change*, 21(2), 752–760. doi:10.1016/j.gloenvcha.2011.01.002
- Ashworth, P., Jeanneret, T., Gardner, J., & Shaw, H. (2011). *Communication and climate change: What the Australian public thinks*. Canberra: Csiro publishing. Retrieved from
- Bolsen, T., Leeper, T. J., & Shapiro, M. A. (2013). Doing what others do: Norms, science, and collective action on global warming. *American Politics Research*, 42(1), 65–89. doi:10.1177/1532673X13484173
- Boyes, E., Skamp, K., & Stanisstreet, M. (2009). Australian secondary students' views about global warming: Beliefs about actions, and willingness to act. *Research in Science Education*, 39(5), 661–680. doi:10.1007/s11165-008-9098-5
- Brügger, A., Morton, T. A., & Dessai, S. (2016). "Proximising" climate change reconsidered: A construal level theory perspective. *Journal of Environmental Psychology*, 46, 125–142. doi:10.1016/j.jenvp.2016.04.004
- Corner, A., Venables, D., Spence, A., Poortinga, W., Demski, C., & Pidgeon, N. (2011). Nuclear power, climate change and energy security: Exploring British public attitudes. *Energy Policy*, 39(9), 4823–4833. doi:10.1016/j.enpol.2011.06.037
- Dietz, T., Dan, A., & Shwom, R. (2007). Support for climate change policy: Social psychological and social structural influences. *Rural Sociology*, 72(2), 185–214. doi:10.1526/003601107781170026
- Ding, D., Maibach, E. W., Zhao, X., Roser-Renouf, C., & Leiserowitz, A. (2011). Support for climate policy and societal action are linked to perceptions about scientific agreement. *Nature Climate Change*, 1(9), 462–466. doi:10.1038/NCLIMATE1295
- Hornsey, M. J., & Fielding, K. S. (2016). A cautionary note about messages of hope: Focusing on progress in reducing carbon emissions weakens mitigation motivation. *Global Environmental Change*, 39, 26–34. doi:10.1016/j.gloenvcha.2016.04.003
- Leiserowitz, A. (2006). Climate change risk perception and policy preferences: The role of affect, imagery, and values. *Climatic Change*, 77(1–2), 45–72.
- Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinberg, G., Rosenthal, S., & Marlon, J. (2014). *Climate change in the American mind: Americans' global warming beliefs and attitudes in November, 2013*. New Haven, CT: Yale Project on Climate Change.

- Linden, S. L. van der, Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2014). How to communicate the scientific consensus on climate change: plain facts, pie charts or metaphors? *Climatic Change*, 126(1–2), 255–262. doi:10.1007/s10584-014-1190-4
- McCright, A. M., Dunlap, R. E., & Xiao, C. (2013). Perceived scientific agreement and support for government action on climate change in the USA. *Climatic Change*, 119(2), 511–518. doi:10.1007/s10584-013-0704-9
- McCright, A. M., Dunlap, R. E., & Xiao, C. (2014). Increasing influence of party identification on perceived scientific agreement and support for government action on climate change in the United States, 2006–12. *Weather, Climate, and Society*, 6(2), 194–201. doi:10.1175/WCAS-D-13-00058.1
- Poortinga, W., Spence, A., Whitmarsh, L., Capstick, S., & Pidgeon, N. F. (2011). Uncertain climate: An investigation into public scepticism about anthropogenic climate change. *Global Environmental Change*, 21(3), 1015–1024. doi:10.1016/j.gloenvcha.2011.03.001
- Zhao, X., Leiserowitz, A. A., Maibach, E. W., & Roser-Renouf, C. (2011). Attention to science/environment news positively predicts and attention to political news negatively predicts global warming risk perceptions and policy support. *Journal of Communication*, 61(4), 713–731. doi:10.1111/j.1460-2466.2011.01563.x