UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut sociologických studií

Metodologický pohled na měření (anti)imigračních postojů

(kvalita měření se zaměřením na jeho ekvivalenci)

Diplomová práce

Praha 2019

Autorka práce: **Anna Šarapatková**

Vedoucí práce: **Mgr. Jiří Remr, Ph.D.**

Rok obhajoby: **2019**

Bibliografický záznam

ŠARAPATKOVÁ, Anna. Metodologický pohled na měření (anti)imigračních postojů (kvalita měření se zaměřením na jeho ekvivalenci). Praha, 2016. 79 s. Diplomová práce (Mgr.) Univerzita Karlova, Fakulta sociálních věd, Institut sociologických studií. Katedra sociologie. Vedoucí diplomové práce Mgr. Jiří Remr, Ph.D.

Rozsah práce

126 527 znaků

Abstrakt

Možnosti, které máme k dispozici v dnešním světě, se vyvíjí stále rychleji a svět se proměňuje přirozeně spolu s těmito změnami. Výrazně tento vývoj pozorujeme v otázce světového pohybu obyvatelstva, které ke dnešnímu dni dosáhlo změn zásadního rozsahu jak ekonomického či politického, tak sociálního. Dnešní tolik diverzifikovaná podoba migrace, která ztratila svou transparentnost, kterou oplývala dříve, je velmi aktuálním a diskutovaným tématem současnosti téměř celého světa. Vzhledem k této významnosti tématu migrace se stává čím dál tím častěji předmětem výzkumu a mnohých výběrových šetření. Jednou z nejčastěji zkoumaných oblastí v rámci této problematiky jsou postoje jedinců k imigraci a především k imigrantům, které jsou často zkoumány ve spojitosti s příčinami, které mohou konkrétní postoje ovlivňovat. Vzhledem k tomuto mezinárodnímu dosahu tématu tyto postoje často šetří mezinárodní výzkumy či výzkumy národní, které však data z mezinárodních šetření přebírají. Napříč evropskými státy existuje jasná nejednota v postojích obyvatel jednotlivých států k imigraci a především samotným příchozím. Vzhledem k této povaze výzkumů zkoumajících postoje obyvatel k imigrantům je důležité se věnovat kvalitě daného měření, které se stává v perspektivě mezinárodních výzkumů značně komplexnější. Je třeba brát v potaz provázanost dílčích chyb, které se v rámci výzkumného šetření objevují, ale především je pro tyto výzkumy zásadní otázka ekvivalence čili srovnatelnosti dat/měření. To znamená, že v případě je-li náš zkoumaný vzorek pestrý a obsahuje skupiny, které jsou odlišné, jako je tomu v případě mezinárodních výzkumů a odlišných zemí a kultur, je třeba si klást otázku, zda měření poskytuje data, která jsou následně srovnatelná, zda je měření ekvivalentní. Jinými slovy lze říci, že je žádoucí, aby byl daný měřící nástroj aplikovatelný ve všech skupinách a změřil ekvivalentní data. Jedním z nejvýznamnějších, zda ne nejvýznamnějším, nástrojem, který měří (anti)imigrační postoje, je vyvinut a měřen v prostředí instituce European Social Survey. Analýza ekvivalence měření tohoto instrumentu, která byla provedena prostřednictvím více skupinové konfirmační faktorové analýzy, demonstruje, že vybraná škála je ekvivalentní, přestože ne všechny položky jsou stejně ideálně vnímány respondenty. To znamená data o postojích obyvatel ČR k imigrantům (na základě vybraných položek) lze vzhledem k potvrzení dobré srovnatelnosti měření použít a dále interpretovat s oporou dobré kvality těchto dat.

Abstract

Opportunities that we have in today's world are sharply evolving, and the world is changing all together with these changes. This development is noticeably observed within the topic of global movement of (not only) population, which has changed fundamentally, both economically, politically and socially. Today's so much diversified form of migration, which has lost its transparency it used to has, is a very up to date and debated topic currently almost all over the world. Because of high importance of the topic "migration" it is often subject of research and number of surveys. One of the most examined area within the topic migration is attitudes of people towards immigration and immigrant, oftentimes together with investigating cause leading to particular attitude. Due to the international reach of the topic, these attitudes are often subject of cross-national research or national research, which, however, use data from international surveys. There is a clear disparity across European states in these attitudes towards immigration and, above all, the immigrants themselves. Given this nature of cross-national surveys measuring attitudes towards immigrants, it is important to focus on the measurement quality, which is becoming increasingly complex in the perspective of international research. It is necessary to take into account the interdependence of partial errors that appear in surveys, but above all focus on the question of equivalence (comparability) of data / measurement is crucial for this kind of studies. It means that if our sample is varied and contains different groups, as is the case with international research and different countries and cultures, the main question is whether the measurements provide data that are comparable, whether the measurement is equivalent. In other words, it is desirable that a given measuring tool is applicable in all groups and measure equivalent / comparable data. One of the most important, if not the most important, tools to measure (anti) immigration attitudes in Europe is developed and measured within the European Social Survey. The analysis of measurement equivalence of this instrument by means of MGCFA demonstrates that the selected scale is equivalent, although not all items are equally ideally perceived by respondents. This means data on attitudes of Czech citizens towards immigrants (based on selected items) can be used and further interpreted with the support of good quality of these data.

Klíčová slova

(anti)imigrační postoje, měření, kvalita dat, ekvivalence měření, škála, více skupinová konfirmační faktorová analýza, strukturální modelování, celková chyba šetření, výběrové šetření

Keywords

(anti)immigration attitudes, measurement, data quality, measurement equivalence, measurement invariance, scale, multi-group confirmation factor analysis, structural modelling, TSE, survey

Prohlášení

1. Prohlašuji, že jsem předkládanou práci zpracovala samostatně a použila jen uvedené prameny a literaturu.
2. Prohlašuji, že práce nebyla využita k získání jiného titulu.
3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne Anna Šarapatková

Poděkování

Chtěla bych poděkovat svému vedoucímu diplomové práce Mgr. Jiřímu Remrovi, Ph.D. za vedení a odborný dohled. Děkuji také PhDr. Ing. Petru Soukupovi, Ph.D. za pomoc a odborné rady při zpracování výzkumné části práce.

Obsah

1. Úvod ... 3
2. Proměna migrace – více, než multikulturalismus ... 4
3. Měření v sociálních vědách – Od konceptu k měřicímu nástroji ... 6
3.1. Měření postojů ... 8
4. Kvalita šetření ... 9
4.1. Celková chyba šetření a druhy chyb ... 9
4.2. Komplexnost a provázanost chyb v rámci šetření ... 12
5. Kvalita měření ... 15
5.1. Klasická Testová Teorie ... 15
5.2. Klasická testová teorie a systematická chyba ... 16
5.3. Typy měření provázených opakovaně ... 18
5.4. Mezi ideálním měřením a skutečným měřením - RELIABILITA ... 18
5.4.1. Odhad reliability ... 19
5.4.1.1. Opakované dotazování stejného vzorku respondentů ... 20
5.4.1.2. Mezipoložková reliabilita ... 22
5.5. Mezi konstrukty a měřením - VALIDITA ... 23
5.5.1. Typy validit a jejich dosahování ... 23
5.6. Efekt metody ... 24
5.7. Statistický odhad celkové kvality měření ... 25
6. MTMM design a odhad kvality měření ... 26
6.1. Survey Quality Predictor ... 27
7. Koncept ekvivalence v metodologii sociálních věd ... 28
7.1. Definice ekvivalence ... 29
7.2. Kategorizace ekvivalence ... 30
7.2.1. Ekvivalence na teoretické úrovni ... 31
7.2.2. Ekvivalence na úrovni měření ... 31
7.3. Dosahování/testování ekvivalence – úroveň teoretická ... 32
7.3.1. Etické x emické položky ... 33

7.3.2. Překlad dotazníku ... 33
7.3.3. Kognitivní metody ... 34
7.4. Testování/ odhad ekvivalence – úroveň měření ... 35
7.4.1. Strukturní modelování ... 36
7.4.2. Více skupinová konfirmační faktorová analýza ... 37
8. Měření (anti)imigračních postojů ... 39
8.1. European Social Survey ... 40
8.2. Metodologie ESS ... 40
8.2.1. Statistický odhad kvality měření v rámci analýz ESS ... 41
8.2.2. Odhad kvality měření otázek z modulu imigrace v rámci všech kol šetření ESS skrze MTMM experimenty ... 42
9. Statistické testování ekvivalence v SPSS a AMOS ... 50
9.1. (Anti)imigrační postoje obyvatel ČR ... 52
9.2. Explorační faktorová analýza ... 54
9.3. Více-skupinová konfirmační faktorová analýza ... 54
10. Diskuze a závěr ... 59
11. Literatura ... 63
12. Přílohy ... 70

1. Úvod

Jednotlivé společnosti a kultury nejsou statické prvky, ale naopak velmi dynamické entity prostupující neustálou změnou. Sociologie je pak vědou, která se těmito proměnlivými a rozmanitými útvary a problémy s nimi spojovanými zabývá a je založena na jejich porovnávání. Za pomoci této srovnávací metody je možné jednotlivé systémy lépe popsat. Změna však neprobíhá pouze na úrovni kultur a společností, ale (mimo další) také v oblasti kvantitativního výzkumu a především jeho metodologie. Proto máme-li v úmyslu provést kvalitní výzkum, respektive získat kvalitní data, musíme reagovat na proměnlivost, která prostupuje zkoumanými společnostmi a kulturami, ale současně také mít na paměti vývoj, ke kterému dochází v rámci metodologie výzkumů sociálních věd. Je totiž velmi nešťastným přístupem ve výzkumu aplikování již existujícího měřícího nástroje bez zohlednění zmíněných rozdílů, změn či vývoje ve společnosti i metodologii. Stejně tak sekundární využití a interpretace dat bez zhodnocení širšího kontextu problematiky a rozlišností zkoumaných skupin je poměrně limitovaný vědecký přístup. Z důvodu zásadní významnosti, kterou přisuzuji tématu kvality měření, se věnuji problematice měření v rámci výzkumných výběrových šetření se zaměřením na postojové škály. Koncentruji se na hlavní aspekty (psychometrické ukazatele) v procesu dosahování kvality měření, které jsou nezbytnými prvky v rámci statistického odhadování celkové chyby měření, která je jedním z nejvýznamnějších ukazatelů kvality měření, potažmo získaných dat. Mimo zmíněné ukazatele se práce zaměřuje na srovnatelnost měření v rámci především mezinárodních komparativních studií.

Nastíněný metodologický rámec práce je provázán s jedním z nejaktuálnějších a významným sociálním jevem, který proměnil svět a má na naše současné (a pravděpodobně i budoucí) životy značný vliv. Tímto činitelem je změna světového pohybu obyvatelstva, jinými slovy otázka migrace a aktérů jednajících v rámci ní. Jelikož je práce metodologického charakteru, toto téma migrace a imigrantů je v práci diskutováno pouze okrajově a především jako prostředek pro usnadnění pochopení aktuální situace, jako pomocné vodítko pro porozumění obsahu a podstaty práce, která je metodologického charakteru, jak bylo naznačeno již v prvním odstavci. Především je však toto téma migrace v práci pomyslnou základnou pro interpretaci dílčích výsledků analýz. V práci bude několikrát ilustrováno a zdůrazněno jak důležitou roli hraje kontext. Konkrétně je v práci analyzováno měření postojů (anti)imigračních a to v prostředí European Social Survey, které je mezinárodním komparativním šetření, a proto je celá problematika měření a dosahování jeho kvality řešena současně v mezinárodním kontextu se zohledněním elementů typických pro tento druh studií. Především se jedná o otázku

srovnatelnosti dat, která je v tomto typu výzkumu svou významností mezi ukazateli kvality měřícího nástroje stejně významným indikátorem jako je reliabilita či validita. Bohužel však srovnatelnosti měření není věnována taková pozornost jako zmíněným dvěma zbylým ukazatelům. Tento nedostatek se diplomová práce pokouší napravit a zvýšit pozornost přikládanou tomuto ukazateli kvality měření.

Metodologii tvorby měřících nástrojů a jejich kvalitu zde zkoumám ve spojitosti s (anti)imigračními postoji ze dvou hlavních důvodů. Zaprvé, jak již bylo zmíněno výše, se jedná o fenomén, který je velice aktuální a významný v oblasti nejen výzkumu. Druhým důvodem je, že data o pohybu obyvatelstva (případně měřící nástroje) a jeho vnímání obyvateli v konkrétních částech světa, jsou velmi často přebírána v prostředí jiných výzkumů pro sekundární analýzu dat. Nicméně následná interpretace dat, zhodnocení aplikovatelnosti měření pro specifickou nově zkoumanou skupinu či zohlednění podmínek, za kterých byla data sbírána, není bohužel samozřejmostí. Proto tato práce shrnuje zásadní procesy a kroky nutné pro dosažení kvalitního měření a to nejenom v prvotním výzkumu, v rámci kterého je konstruován měřící nástroj, ale také v případě sekundární analýzy dat či převzetí měřícího nástroje. Toto shrnutí se snaží být aktuální, komplexní (optika mezinárodního výzkumu a provázanost chyb) a upozornit na možné nedostatky některých stále praktikovaných technik, pro které je současně navrženo alternativní řešení či realističtější model jeho provedení. Výše zmíněné dílčí cíle jsou pod záštitou jednoho většího, kterým je objasnění problematiky kvalitního měření postojů v sociálních vědách, které je aktuální a velice komplexní. Druhý hlavní cíl této práce je analytický a spočívá ve zhodnocení srovnatelnosti měření (anti)imigračních postojů v prostředí European Social Survey. Prvotní je zde nalezení odpovědí především pro Českou republiku, tj. jaká je kvalita měření vybraných položek v našem prostředí, respektive zda je instrument aplikovatelný v prostředí ČR a jak se postoje obyvatel k tomuto fenoménu mění nejenom s ohledem na kulturu a stát, ale také v průběhu času.

2. Proměna migrace – více, než multikulturalismus

V rámci stručné demonstrace současné migrační situace bych ráda využila termínu Stevena Vertoveca, kterým je „*superdiversity*“ či superdiverzita [Vertovec 2007]. Vznik tohoto pojmu se totiž pojí s novou situací ve společnosti, respektive ve světě, který se tak zásadně proměnil s ohledem na pohyb obyvatelstva, jednotlivé životy či komunikaci, a termíny do té doby existující se proto staly nedostačující pro popis dané situace, která se stala o mnoho komplexnější, než tomu bylo kdy dříve. Superdiverzita je pojem, který se často pojí s Velkou

Británií, jelikož ta je, a především byla, v době vzniku pojmu, symbolem komplexnosti a diverzity dnešní doby a je skvělým ilustračním příkladem pro vysvětlení aktuální problematiky. Jedná se o místo, kde se, jako v jednom z prvních, ve velkém měřítku (ne)mísilo množství odlišných kultur. Toto množství diverzity tak dalo nejprve vznik dodnes populárnímu pojmu multikulturalismus, který obecně řečeno prosazuje toleranci s ohledem na kulturní diverzitu a podporuje mírumilovné soužití skupin z různých kulturních prostředí se svými kulturními odlišnostmi a jejich respektování [Guttman 1994; Modood 2007]. Vesměs můžeme říci, že je tento přístup protipólem teorie „melting potu", dle které se naopak kultury mísí dohromady [Alba, Nee 1997]. Nicméně i přes značnou popularitu tohoto pojmu v současné době a jeho frekventované používání je multikulturalismus fenomén, o němž, respektive o jeho přítomnosti, by bylo vhodnější mluvit do období devadesátých let minulého století. Následný přelom desetiletí znamenal změnu, která vyžadovala úpravu konceptualizace této problematiky a považuji za důležité zde právě tuto proměnu zdůraznit a aktualizovat tak povědomí o migrační situaci. Stručně řečeno do devadesátých let byla migrace a (i)migranti považováni za záležitost transparentní a relativně uspořádanou, jelikož existovaly stálé neměnné proudy migrantů, proudící ze stejných zemí původu na podobné cílové lokace, a jednalo se také o migranty s podobnými charakteristikami, kteří se pohybují po stejných trajektoriích. Proto vezmeme-li v potaz změnu v době rozrůstající se komplexnosti současné diverzity, kterou přinesl konec devadesátých let a mizení oné transparentnosti, je logické, že bylo potřeba termínu, respektive přístupu, který bude tuto situaci studovat a mapovat. Tímto konceptem je onen pojem superdiverzita, který zdůrazňuje mnoho-vrstevnatost současné migrace. Už se nesetkáme pouze s diverzitou, nýbrž diverzitou diverzity. Tento proces můžeme zaznamenat jak na výčtu zemí, ze kterých migranti putují či zemí, které je přijímají, ale také je diverzita bohatší a strukturovanější na úrovni socioekonomických charakteristik jedinců, kulturních charakteristik či diferencí jazykových nebo náboženských. Současně je třeba brát v potaz diverzifikaci trajektorií a skutečnost, že lidé se nyní zapojují do více mezinárodních pohybů (ať fyzických či abstraktních) najednou [Vertovec 2007]. Shrneme-li výše zmíněné, je důležité vidět dynamickou interakci výše zmíněných i dalších proměnných v rámci pohybu (nejen) obyvatelstva ve světě, která aktuálně probíhá. Tuto situaci můžeme ilustrovat za pomoci porovnání migrace ve dvou časových bodech jejího historického vývoje. V úvodu textu jsem mluvila o Velké Británii. V souvislosti s ní si představme typickou migraci z britských kolonií jako příklad klasické transparentní migrace (či pohyb mezi jinými bývalými koloniemi a kolonizátory). Následně, pro porovnání, si uvědomme změny posledních přibližně dvaceti let. Nežijeme sice v době „válečnější" s ohledem na historii, ale čelíme více útokům a problémům

ve světě, především nové podoby, které jsou bez pochyb odlišné. Tyto nové a četnější útoky a nové problémy vedou k rozšíření migrace z mnohonásobně většího množství světových lokalit a tudíž výraznému nárůstu komplexnosti celého fenoménu. To znamená, nová migrace s sebou přináší také nové migranty.

Pohyb světového obyvatelstva se proměnil nejenom s ohledem na komplexitu, ale v prostředí Evropy zaznamenal také enormní nárůst oněch příchozích [Hooghe at al. 2008]. Migrace, především nárůst imigrantů v Evropě, je velmi aktuálním tématem a často také velmi vyhroceným. Mnoho výzkumů demonstruje, že značná část populace, která spadá do významných hostitelských zemí EU, shledává imigraci, respektive imigranty, jako negativní prvek nebo jako prvek, který s sebou přináší negativní dopady na hostitelskou společnost [Anderson 1996; Lubbers, Gijsberts, Scheepers 2002]. Česká republika patří k těmto státům, ačkoli v porovnání s jinými evropskými státy má zkušenosti s imigranty výrazně menší, a stejně tak v současné době s imigranty obyvatelé ČR nepřijdou do kontaktu tak často jako je tomu v jiných státech Evropy, které jsou k otázce migrace a především samotným příchozím přívětivější. Z důvodu specifičnosti ČR s ohledem na vnímání imigrace a imigrantů a obecně rozdílnosti tohoto vnímání napříč státy, je více, než důležité se věnovat kvalitě měření (anti)imigračních postojů napříč státy, respektive srovnatelnosti měření, aby bylo jasné, zda nesrovnáváme jablka a hrušky a výsledky jsou spolehlivé a vypovídající.

3. Měření v sociálních vědách – Od konceptu k měřicímu nástroji

Jelikož je hlavním předmětem této práce metodologie a konstrukce či kvalita měřících nástrojů v rámci sociologických šetření, považuji za nezbytné jednotlivé kroky dosahování oné kvality uvést alespoň stručným popisem a vysvětlením samotného měření postojů v sociálních vědách. Proces měření, který můžeme jinak označit jako pozorování nebo sběr dat, je sběrem informací o cílové populaci, na kterou chceme vztahovat zjištění, kterých jsme se dobrali analýzou a interpretací těchto dat získaných prostřednictvím pozorovaného vzorku cílové populace [Babbie 2013; Řehák 1971, Berka 1971, Illner 1972]. Například Groves definuje měření v této oblasti jako různé způsoby, kterými jsou získávána data o jednotlivých konstruktech (co jsou to konstrukty, je vysvětleno níže) [Groves 2011: 43]. Toto vysvětlení měření je samozřejmě velmi obecné, proto čtenáře, který nenalezne v této práci veškeré podrobnosti o měření v sociálních vědách, odkazuji na práce Hynka Jeřábka, Karla Berku či Jana Řeháka z českých autorů, ze zahraničních pak na Aarona van Cicourela, Roberta Grovese či Willema E. Sarise a Irmtrauda N Gallhofera, kteří jsou v současnosti velmi významnými

autory v oblasti metodologie a významně přispívají svým praktickým a statistickým zaměřením a zdokonaleními v prostředí současných výzkumů. Z definice měření logicky vyplývá, že měřící nástroje jsou prostředky, díky kterým může být tato činnost (měření/pozorování) vykonávána. V sociologii, stejně jako v jiných oborech, existuje mnoho způsobů jak data získávat. V této práci je pozornost věnována pouze měření kvantitativnímu, ke kterému dochází v rámci výběrových šetření. Konkrétně se jedná o měření postojů jednotlivců k danému tématu (v této práci je tímto tématem imigrace a imigranti).

Pozorování či měření v kontextu sociálních věd se liší od měření ve vědách přírodních v jedné zásadní věci, kterou je měření něčeho, co ve skutečnosti jako takové neexistuje (neexistuje co do smyslu objektivnosti a „nezkonstruovanosti“). Proto prvním z konceptů, kterému je nutno v rámci procesu měření důkladně porozumět, je konstrukt, respektive: *koncept a konstrukt*. Vysvětlení těchto pojmů zde poskytuji skrze Kaplanovo rozlišení tří oblastí vyčleněných na základě toho, co je měřeno [Kaplan 2017]. Znovu uvádím, že pro sociologii je typická třetí uvedená skupina právě zmíněných konstruktů.

- Přímo pozorovatelné (př.: barva svetru, co bylo zaškrtnuto v dotazníku)
- Nepřímo pozorovatelné (př.: knihy o historii poskytují nepřímé pozorovaní sociálních akcí v minulosti)
- Konstrukty (založeno na pozorování, ale nemohou být pozorovány přímo ani nepřímo, samy o sobě neexistují, vytvořili jsme je; př.: rasismus)

Příkladem konstruktu může být IQ. Tento příklad je často pro ilustraci uváděn, pravděpodobně také z důvodu jeho souvislosti se vznikem tzv. latentních proměnných, které jsou právě oněmi konstrukty (blíže o vzniku latentních proměnných viz následující kapitola o měření postojů). Analogii můžeme pozorovat také v terminologii Sarise a Gallhofera, kteří rozlišují intuitivní koncepty (*concepts by intuition*) a koncepty na základě deklarování (*concepts by postulation*) [Saris, Gallhofer 2007 :15-20]. Intuitivními koncepty jsou myšleny ty, jejichž význam je okamžitě zřetelný, v druhém případě (concepts by postulation) se jedná o koncepty, které jsou výše zmíněnými konstrukty a je nutná jejich přesnější definice [Saris, Gallhofer 2007:15]. V případě otázky měření (anti)imigračních postojů je proto zřejmé, že (anti)imigrační postoj je konstruktem či konceptem na základě deklarace, a intuitivními koncepty by v tomto případě byly například dílčí informace jako pocit ohledně bydlení vedle imigranta, zda jedinec volil pro určitá (proti)imigrační opatření a podobně.

Z výše uvedeného vyplývá, že ke změření konstruktu, který zastupuje určitý teoretický koncept (zkoumaný fenomén, který se ve společnosti objevuje a je předmětem výzkumu), a který měřitelný není, je třeba nalézt prostředek, jak informace o konstruktu získat. Tímto prostředkem jsou *indikátory*. Indikátory měřitelné jsou, odpovídají tomu, co jsme výše popsali jako koncepty intuitivní a představují proto dílčí kroky v podobě jednotlivých otázek vedoucích ke změření latentní proměnné, čili daného konstruktu. Je třeba si však uvědomit, že každý jedinec je jiný, každý má zcela autentické a specifické vnímání světa. S tímto vnímáním se přirozeně pojí i subjektivní vnímání jednotlivých konstruktů. Z tohoto důvodu je nezbytné, aby v průběhu designu výzkumu, a především konstrukce měřícího nástroje, byla věnována dostatečná pozornost jednotlivým konceptům či konstruktům, které by v žádném případě neměly vznikat z vlastní vůle výzkumníka, bez jakéhokoli podložení literaturou, empirií či konsensem s ostatními členy výzkumného týmu a důkladného popisu daného konstruktu. Tato praxe je obsahem procesu zvaného *konceptualizace a operacionalizace*.

3.1. Měření postojů

Když se mluví o postojích, mají se dle definice na mysli „naučené predispozice k odpovídání konzistentně příznivým nebo nepříznivým způsobem vzhledem k danému předmětu“ [Fishbein and Ajzen 1975: 6]. V případě této práce je předmětem „imigrace“, ke které se respondenti staví pozitivně či negativně. Tyto postoje si jedinci utváří na základě vjemů, které získávají ze svého okolí. Patří mezi ně nejčastěji informace od známých či média, ale také osobní zkušenost. Avšak přirozeně ne všechny vjemy a získané informace o daném předmětu musí být pravdivé. A ačkoli uvedená definice zmiňuje pojem „konzistentnosti“, v praxi lze pozorovat to, co dokazují i četné studie, že ať už si utvoří respondent názor na daný předmět jakýkoli, není striktně konzistentní [Krosnick, Alwin 1987; Tourangeau, Rasinski 1988] Přestože existují případy, kdy jsou postoje konzistentnějšími, například v případě osobní zkušenosti, stále je třeba si uvědomovat, že na tvorbu a změnu postoje má významný vliv mnoho dalších faktorů.

K významnému momentu v historii měření v sociálních vědách došlo na začátku minulého století, když Charles Spearman v rámci svého měření různých vlastností či schopností jedinců (toho, jak jsou dobří v jazycích, matematice atp.) odhalil, že jsou tyto jednotlivé schopnosti velmi silně korelovány [Spearman 1904]. Snažil se proto zjistit, co se skrývá v pozadí za těmito proměnnými, které hromadně měří to, co dnes již běžně označujeme jako latentní proměnná. Zde navazujeme na téma konstruktů, které bylo vysvětleno v úvodu. Spearman došel k závěru,

že v jeho případě je touto latentní proměnnou (či konstruktem) inteligence [Spearman 1910, 1904]. Paul F. Lazarsfeld zastával podobný přístup, avšak nepředpokládá identicky jako Spearman cosi skrytého za daty, ale usiluje o určitou typologii – „latentní strukturu“ [Lazarsfeld 1968]. Tímto přístupem nastavil Lazarsfeld základy pro dnes poměrně běžně aplikovanou analýzu latentních tříd. Významné pro metodologii však bylo zjištění Lazarsfelda s Mertonem, o měření určité latentní proměnné, které je mnohem přesnější, je-li při jejím měření použito více otázek, respektive položek [Lazarsfeld, Merton 1950].

Více takových položek, které měří jeden konstrukt, jsou označovány škálami, jejichž prostřednictvím jsou postoje kvantitativně měřeny. V souvislosti s rozvojem teorií o latentních proměnných a jejich identifikování či měření byly škály vystaveny značnému vývoji a dnes se setkáme s poměrně širokým spektrem podob těchto měřících instrumentů. Jak bylo uvedeno již v předchozím odstavci, použitím více otázek, položek, či indikátorů, je dosaženo vyšší přesnosti měření, jelikož pouhá přímá otázka na postoj jedince k určitému jevu by vedla k zásadnímu vychýlení v datech.

4. Kvalita šetření

Předchozí kapitola stručně rekapituluje a popisuje měření typické pro sociální vědy, přesněji měření postojů respondentů. Tato pozornost, kterou při vývoji měřících nástrojů vyvíjíme je tolik vyžadována, jelikož data, která nezbytně potřebujeme pro náš výzkum, by měla být co nejkvalitnější. Jakými způsoby, metodami a technikami tohoto požadavku v rámci výzkumu dosahujeme, objasňuji v kapitole o kvalitě měření a jejích jednotlivých podkapitolách. V této kapitole se blíže zaměřuji na zasazení chyb měření do celkového kontextu chyb vznikajících v průběhu šetření se zohledněním nejenom národních, ale také mezinárodních šetření.

4.1. Celková chyba šetření a druhy chyb

Nejčastějším způsobem, kterým je vysvětlována kvalita dat šetření, je její odvození od tzv. celkové chyby šetření (dále již TSE z anglického překladu Total Survey Error). TSE jak již název napovídá, implikuje veškeré odchylky a chyby v datech, které ohrožují naše šetření. Nicméně jakkoli je toto vysvětlení přesné a výstižné, je značně obecné a do vypovídající hodnoty poněkud omezené. Představím proto na následujících řádcích tento koncept detailněji v kontextu nejenom národních šetření, ale i mezinárodních. V závěru kapitoly současně

poukazuji na to, proč odvození kvality dat šetření pouze od TSE by bylo poněkud omezené, jelikož by takové odvození nebylo zcela kompletní.

Není v možnostech této práce jít do takové hloubky teorie TSE, aby byl popsán celý historický vývoj v této oblasti se všemi změnami a postupnými zdokonaleními, tudíž přestože zde uvádím zásadní milníky a pojmy, zdůrazňuji, že se jedná pouze o zásadní body. Samotný vývoj TSE probíhal v intenzivnějším rozsahu od poloviny čtyřicátých let minulého století, kdy Edwards Deming v roce 1944 publikoval v americkém sociologickém časopise článek, ve kterém upozorňuje na pestrost chyb, respektive zdrojů chyb, které se vyskytují ve výběrových šetření - konkrétně se jednalo o seznam 13 faktorů, které ovlivňují kvalitu dat [Deming 1944]. Pro ilustraci zde uvádím tabulku těchto faktorů, po jejímž prohlédnutí lze zřetelně pozorovat, že mimo chyby pokrytí populace, obsahuje tento seznam, který vznikl v polovině minulého století, většinu okruhů chyb, se kterými se pracuje dodnes, avšak dnes již ve strukturovanější a detailnější podobě.

1. Variability in response;
2. Differences between different kinds and degrees of canvass;
   (a) Mail, telephone, telegraph, direct interview;
   (b) Intensive vs. extensive interviews;
   (c) Long vs. short schedules;
   (d) Check block plan vs. response;
   (e) Correspondence panel and key reporters;
3. Bias and variation arising from the interviewer;
4. Bias of the auspices;
5. Imperfections in the design of the questionnaire and tabulation plans;
   (a) Lack of clarity in definitions; ambiguity; varying meanings of same word to different groups of people; eliciting an answer liable to misinterpretation;
   (b) Omitting questions that would be illuminating to the interpretation of other questions;
   (c) Emotionally toned words; leading questions; limiting response to a pattern;
   (d) Failing to perceive what tabulations would be most significant;
   (e) Encouraging nonresponse through formidable appearance;
6. Changes that take place in the universe before tabulations are available;
7. Bias arising from nonresponse (including omissions);
8. Bias arising from late reports;
9. Bias arising from an unrepresentative selection of date for the survey, or of the period covered;
10. Bias arising from an unrepresentative selection of respondents;
11. Sampling errors and biases;
12. Processing errors (coding, editing, calculating, tabulating, tallying, posting and consolidating);
13. Errors in interpretation;
   (a) Bias arising from bad curve fitting; wrong weighting; incorrect adjusting;
   (b) Misunderstanding the questionnaire; failure to take account of the respondents' difficulties (often through inadequate presentation of data); misunderstanding the method of collection and the nature of the data;
   (c) Personal bias in interpretation.

Obrázek č. 1: 13 faktorů ovlivňujících kvalitu fakt dle Deminga (Deming 1994)

Mezi zmíněná propracovanější a již členitější schémata můžeme zařadit například to Roberta Grovese. Toto schéma, respektive jeho jednotlivé komponenty, o pár řádků níže používám pro vysvětlení hlavní typologie chyb, které se ve výběrovém šetření vyskytují a je nejvyšším cílem výzkumníků je minimalizovat. Je známo, že vezmeme-li v potaz tento typ šetření v rámci sociálních věd, existují dva druhy jakési *inference*, či jinými slovy usuzování či

odvozování. Inferencí je myšlen popis určitého nepozorovaného fenoménu na základě pozorovaného. V prvním případě je možné mluvit o inferenci na úrovni, kde onou pozorovanou složkou je náš vzorek respondentů a nepozorovanou, na kterou je usuzováno, je cílová populace [Groves 2011: 63; Groves, Lyberg 2010]. V druhém případě inference odvozujeme na základě proměnných, indikátorů či dotazníku, obecné charakteristiky respondenta – či latentní koncepty - v takovém případě hovoříme o inferenci konceptuální [Groves, Lyberg 2010].

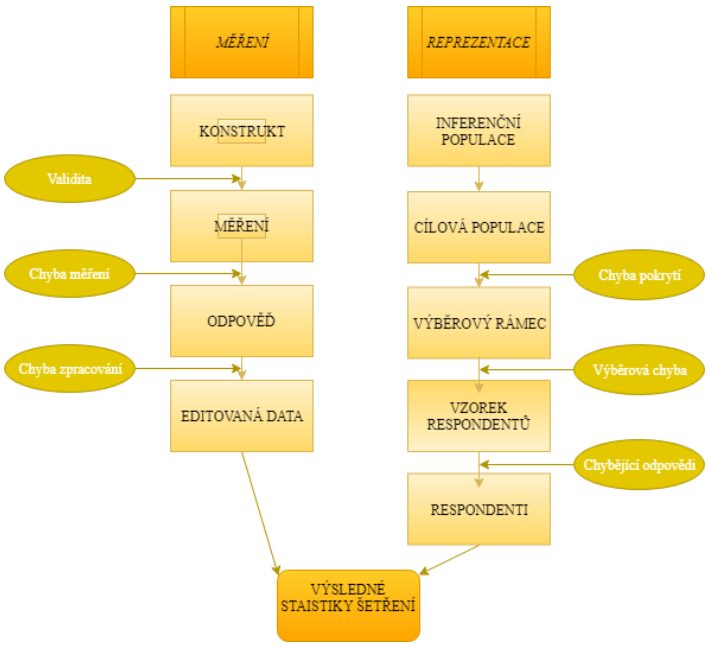


Schéma 1: Schéma komponentů TSE propojených s jednotlivými inferenčními procesy [Groves at al 2004]

O inferenci se zde zmiňuji proto, že právě tyto dva (inferenční) procesy zároveň zrcadlí hlavní skupiny chyb TSE, se kterými se ve výběrových šetřeních setkáváme a vypořádáváme. V Grovesově rozdělení lze pozorovat, že schéma obsahuje dělení na dvě hlavní části [Groves 1989]. Jedna ze skupin chyb se týká celého procesu měření (často jsou tyto chyby analogicky označovány jako chyby pozorování, jelikož vznikají již v samotném procesu měření/pozorování nebo v souvislosti s ním), v tomto případě se chyby pojí s inferencí, kterou jsme výše nazvali konceptuální. Druhá oblast chyb pojímá chyby reprezentace. V rámci této sekce chyb je možné zaznamenat další dílčí skupiny chyb. Zaprvé se jedná o chyby, které se pojí s naším výběrem, tzv. chyby výběrové, jelikož nemáme-li data od celé cílové populace, ale

pouhý vzorek, budeme se potýkat s nepříjemnou avšak nevyhnutelnou situací, kdy se naše data neshodují svými hodnotami s daty v populaci. Zadruhé pak v této skupině chyb reprezentace čelíme chybám, které jsou zapříčiněny skutečností, že ne všichni respondenti, které oslovíme, budou na výzkumu také reálně participovat, tudíž čím více takových potenciálních respondentů ztratíme, tím je vyšší pravděpodobnost, že některá ze skupin bude méně či více zastoupena. Současně může nastat situace, že respondenti odmítnou odpovědět na určité otázky, a v případě, že se budou tomuto odpovídání vyhýbat systematicky, způsobí opět vychýlení našich dat a následně i výsledků. Ať už mluvíme o kterémkoli ze zmíněných typů chyb v rámci druhé skupiny, mluvíme o chybách, které jsou označovány jako chyby nepozorování, jelikož na rozdíl od první skupiny chyb (chyby měření či pozorování), vycházejí ze skutečnosti, že některé elementy/jednotky jsme vůbec nepozorovali [Groves 2011]. Pro tuto práci je zásadní první skupina chyb, to znamená chyby pozorování či jinak měření. Groves dále v rámci této sekce chyb rozlišuje 4 hlavní zdroje, které mohou chybu způsobit. Těmi jsou: tazatel, respondent, měřící nástroj a metoda sběru dat.

Dalším dělením chyb, které je nezbytné především pro statistické odhadování kvality měření, je rozlišení chyb systematických a náhodných. Systematické chyby, označované anglickým výrazem „bias“, jsou takové, které jsou přímo směrované a jsou způsobeny chybou daného měřícího nástroje nebo špatně zvolené metody a modifikují tak odhady průměrů [Smith 2011]. V případě identifikace takové chyby je možné ji odstranit, nebo její vliv na výslednou hodnotu odhadnout, což je otázka především reliability a validity, o kterých je pojednáno v samostatných kapitolách. Se systematickými chybami tedy můžeme do značné míry pracovat, ale provézt měření nezatížené žádnou chybou je téměř, ne-li zcela, nemožné, jelikož vždy bude přítomna chyba náhodná. Náhodné chyby není možné nikdy zcela odstranit, jelikož jsou způsobeny náhodnými faktory, jako jsou například fyzikální podmínky působící při dotazování, paměť a její neideální fungování či selhání, či selhání jiných smyslů.

4.2. Komplexnost a provázanost chyb v rámci šetření

Přes neopomenutelný přínos problematice TSE a obecně kvalitě dat Roberta Grovese, je přirozené, že se setkáme s určitými nedostatky či omezeními výše popsaného, ať už na úrovni strukturální, v podobě schématu a řešení dimenziálnosti typologie, či převažující pozornosti pouze národním jednorázovým šetřením. V tomto ohledu je možné jeho poznatky doplnit a stručně rozvinout těmi Tomase W. Smitha. Zásadními přínosy v tomto směru jsou především detailnější propracovanost a přehlednost schématu. Důležitým příspěvkem Smitha je důraz na

tzv. „interactions“, které pro potřeby této práce označuji českým ekvivalentem „vzájemné působení“ [Smith 2011]. Jelikož o tomto vzájemném působení hovoříme v souvislosti s TSE, jedná se o působení jednotlivých chyb. To dlouhé roky bylo (dodnes bohužel částečně je) často zcela opomíjeno. Smith vysvětluje, že ačkoli je například chyba měření zjišťována dílčí chybou pocházející od tazatele, respondenta či otázek šetření, nejsou brány v potaz vzájemné působení těchto chyb, ani efekt vzniklý jejich kombinací [Smith 2011].

Velká část literatury na téma kvality dat výběrových šetření, se vztahuje především k jednorázovým šetřením probíhajícím v rámci jedné skupiny v jeden časový bod. Vzájemné ovlivňování chyb je však přítomné i v panelových šetřeních a šetřeních vícenásobných, které jsou prováděny v dnešní době stále častěji. Chyby a jejich interakce v těchto případech dosahují výrazně vyšší komplexnosti. Panelová šetření obsahují chyby, které se v rámci jednoduchých neobjevují [Alwin 2007]. Patří mezi ně například zkušenost respondenta z první vlny dotazování ovlivňující následné dotazování, respektive porozumění, a odpovídání respondenta. Takovými chybami jsou konkrétně falešná konzistence pramenící z odpovídání respondenta stejným způsobem jako v první vlně šetření, ačkoli se nejedná o pravdivou odpověď či chyba pramenící ze znovu-dotazování chybného respondenta, chyba autokorelace aj. [Smith 2009, Jung 2005]. Skutečná výzva pro výzkumníka poté nastává v případě vícenásobných šetření, které porovnávají více dílčích šetření a nezkoumají chybu měření s ohledem pouze na jednu studii [Smith 2011]. Tyto vícenásobné šetření mohou mít několik forem, respektive aspektů, ve kterých se liší:

a) Kdo sbírá data – kdo řídí výzkum (house effect) [Smith 1978, 1982]
b) Kdy byla data sbírána (nejčastěji studia sociální změny) [Smith 2007]
c) Cílová populace či od koho jsou data sbírána (nejčastěji mezinárodní či mezi-kulturní studie) [Smith 2003, 2004, 2010]

Pro tuto práci je zásadní třetí možnost, tedy problematika mezinárodních srovnávacích studií a chyby v rámci nich, jelikož nejvýznamnější zdroje dat o anti-imigračních postojích v ČR jsou právě tohoto charakteru. Respektive, jsou pro tuto práci stěžejní uvedené možnosti všechny tři, jelikož mezinárodní šetření probíhají často současně v různých časových bodech a liší se i jednotlivými kolektory dat, včetně šetření European Social Survey analyzovaného v rámci této práce. Je třeba připomenout, že čím jsou (porovnávané) jednotky odlišnější, tím je šetření a porovnatelnost komplikovanější a jedná se skutečně o metodologickou výzvu analytikům a analytičkám ve spolupráci s metodology, a to z důvodu o pár řádků výše zmíněné

mnoho-dimenziálnosti a komplexnosti šetření a násobící se míry možných interakcí jednotlivých chyb.

Pro více informací o vzájemném působení chyb a jejich propojenosti odkazuji na práci Toma Smitha a dalších, kteří se touto problematikou zabývají a popisují příklady této provázanosti [Alwin 2007; Groves et al., 2004; Schuman & Presser 1981; Smith 2005; Tourangeau, Rips, & Rasinski, 2000].

V souvislosti s tématikou provázanosti chyb šetření je vhodné zmínit, že přestože se ve výzkumu zabývám chybami měření a dosahováním kvality měření v rámci výběrových šetření, je důležité nahlížet i na tuto problematiku komplexněji v širším kontextu. V předchozí kapitole bylo vysvětleno dělení chyb na dvou úrovních pojících se se dvěma typy inference. Chyby měření se vyskytují v rámci skupiny chyb pozorování. Do druhé skupiny chyb, která zohledňuje reprezentaci, jsou zařazeny mimo jiné chyby, které vznikají z důvodu nezodpovězení otázky respondentem (item nonresponse). Ačkoli se nejedná o chyby ze skupiny chyb pozorování (měření) jsou v otázce kvality měření zásadní, jelikož čím nižší návratnost v našem výzkumu bude, to znamená, čím více bude chybějících hodnot v datech, tím více budou zkresleny výsledky analýzy, které nebudou plně reprezentovat zkoumanou populaci [Saris, Gallhofer 2007: 186].

Závěrem této kapitoly bych chtěla důrazně upozornit na skutečnost, že TSE není jediným ukazatelem kvality dat. Tento koncept, jakožto významný ukazatel kvality, byl použit z důvodu jeho vazby na psychometrické aspekty kvality měření, které jsou v této práci rozebírány. Například U. S. Key National Indicators Initiative však uvedli schéma, které skvěle demonstruje jak široké spektrum faktorů ovlivňujících celkovou kvalitu dat je. Většina příspěvků (nejen) metodologů v rámci TSE a metodologie šetření spadá pod položku „data estimator“ a „data quality“, nicméně ve schématu lze pozorovat, že existují i další skupiny, které mají na kvalitu dat vliv. Mezi takové patří například: kredibilita (tj., jak moc uživatelé důvěřují produktu na základě jejich mínění o zprostředkovateli dat, například tedy vize určité značky) či relevance (tj., zda jsou získané informace ohledně problému či daného fenoménu významné a důležité pro uživatele). Tyto ukazatele však nejsou dále v této práci rozebírány. Pozornost je věnována pouze chybám měření a možnostem jejího odhadu a korekce, na základě psychometrických ukazatelů jako je reliabilita, validita či efekt metody, a následně také ekvivalence.

5. Kvalita měření

Jsou situace, kdy v rámci výzkumu pracujeme s proměnnými, pro něž je možné získat skutečné hodnoty (již existující). Příkladem může být evidovaná účast na konkrétní události, například volební účast. V takovýchto případech je pro měření zvolena metoda, která poskytuje svým měřením odpovědi (očištěné o náhodnou chybu) nejvíce se blížící skutečné hodnotě [Saris, Gallhofer 2007: 187]. Nicméně jak již bylo řečeno v předchozím textu, v sociálních vědách měříme často něco, co ve skutečnosti neexistuje. Pracujeme tak s proměnnými, u kterých skutečné hodnoty k dispozici nejsou. Z tohoto důvodu je možné pouze pozorovat výsledky měření prostřednictvím různých metod, které v případě lišících se výsledků, značí systematické vychýlení minimálně jedné z těchto použitých metod. Toto systematické vychýlení je jedním ze základních kritérií, která musí být v dnešní době zohledněna v otázce kvality měření. Existují však kritéria, která vysvětlují vztahy mezi indikátory, které by měly měřit stejnou proměnnou, a rozdíly ve vztazích mezi proměnnými měřenými různými metodami. Těmito kritérii jsou reliabilita, validita a efekt metody. Právě tato kritéria, jakožto ukazatele kvality měření jsou předmětem výkladu následujících kapitol, ve kterých se mimo popis jednotlivých ukazatelů, věnuji technikám dosahování kvality měření s ohledem na tyto ukazatele, respektive jejich odhadování. Jelikož se jedná o aspekty vycházející z psychometrie, je výklad založen na objasnění Klasické testové teorie spolu s jejím rozšířením o systematickou složku, která dodává modelům realističtější tvář a vyšší přesnost. Toto rozšíření Klasické testové teorie je důležité nejenom z důvodu informování čtenáře o vývoji v metodologii, ale také kvůli strukturálnímu modelování, které je aplikováno v rámci analytické části práce, kde má systematický vliv zásadní význam.

5.1. Klasická Testová Teorie

Zmínkou v úvodním odstavci této kapitoly o dosahování skutečné hodnoty s ohledem na vliv chyby v procesu měření ve své podstatě přímo odkazuji na Klasickou testovou teorii (KTT), potažmo teorii skutečné hodnoty (zkratka CTST z anglického překladu Clasical True Score Theory). Jedná se o průkopnickou teorii Lorda a Novicka, která je od 60. let minulého století rozvíjena a tvoří základ pro veškeré odhadovaní chyb v rámci měření [Lord Novick 1966]. Vychází z psychometrie, přičemž toto převzetí přístupu z oboru psychologie v prostředí sociologie se jeví jako zcela logický počin s ohledem na abstraktnost a nesnadnou uchopitelnost či komplikovanější měřitelnost konceptů, která je oběma oborům společná. Klasická testová teorie Novicka a Lorda byla své doby průkopnickou teorií a skutečně zásadním

posunem v rámci metodologie výzkumných šetření. Schéma č. 2 ilustruje základní myšlenku Lorda a Novicka. Tato teorie je slovně vyjádřena následovně: rozptyl měření (y) se skládá z rozptylu dané měřené proměnné (t) a náhodné chyby (e) [Novick 1966].

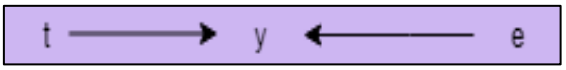


Schéma č. 2: Klasický model KTT s náhodnou chybou

5.2. Klasická testová teorie a systematická chyba

Klasická testová teorie je základem veškerého odhadování chyb měření, avšak je nutné si uvědomit, že počítá s chybou náhodnou, stejně jako techniky odhadů kvality měření, které z ní vychází. To, jaké problémy toto omezení technik přináší, je detailněji popsáno v kapitole pojednávající o reliabilitě. Nyní se zaměřuji pouze na vztah, a jeho následnou změnu, popsaný v rámci KTT respektive CTST, se zohledněním chyby systematické. Vysvětlení podstaty systematických chyb je poskytnuto v rámci kapitoly o celkové chybě měření a druzích chyb, ze které vyplývá důležitost povědomí o rozlišení náhodných a systematických chyb a jejich zohlednění při analýze dat. Přítomnost systematických chyb a jejich vliv lze zřetelně ilustrovat také za pomoci jednoho z ukazatelů kvality měření, který stojí po boku reliabilitě a validitě - *efektem metody*. Určit, zda vzniklý rozptyl v rámci proměnných je výsledkem náhodné chyby či chyby systematické za podmínek, že daný jev měříme pouze jedním způsobem (jednou metodou) je nemožné [Saris, Gallhofer 2007: 174]. Proto je-li pro naše měření použito vícero metod a následně jsou porovnány kontingenční tabulky hodnot proměnných, lze pozorovat, že vzniklá inkonzistence v odpovědích je do určité míry způsobena náhodnou chybou, avšak současně lze vypozorovat určitý vzorec v těchto hodnotách, který odkazuje na systematický efekt způsobený konkrétní metodou. Schéma č. 3, které je rozšířením schématu č. 2, ilustruje popsanou situaci, kdy je zohledněn vliv systematické chyby (u).

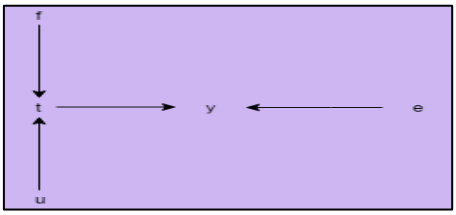


Schéma č. 3: Model měření s náhodnou a systematickou chybou

Na problém identifikace chyb reagovali Campbell a Fiske, kteří již v roce 1959 navrhli využití více, než jedné metody pro měření znaků, aby mohlo být rozeznáno, do jaké míry se jedná o chybu náhodnou a jaké množství rozptylu proměnných je vysvětleno chybou systematickou [Campbell, Fiske 1959]. Uvádím schéma č. 4, kde můžeme pozorovat reakci respondenta na konkrétní metodu. Současně jsou v tomto schématu znázorněny koeficienty reliability či validity, kterým věnuji pozornost níže v textu, stejně jako samotnému efektu metody. Ze statistického hlediska však můžeme na základě uvedeného schématu konstatovat, že systematický efekt metody, může být popsán jako vztah mezi faktorem metody (který reprezentuje reakci respondenta na danou metodu) a skutečnou hodnotou. K tomuto schématu se pozornost vrací v rámci kapitoly vysvětlující statistický odhad tzv. celkové chyby měření. Stejně tak modelu vyvinutému Campbellem a Fiskem (MTMM design) je pozornost věnována v samostatné kapitole.

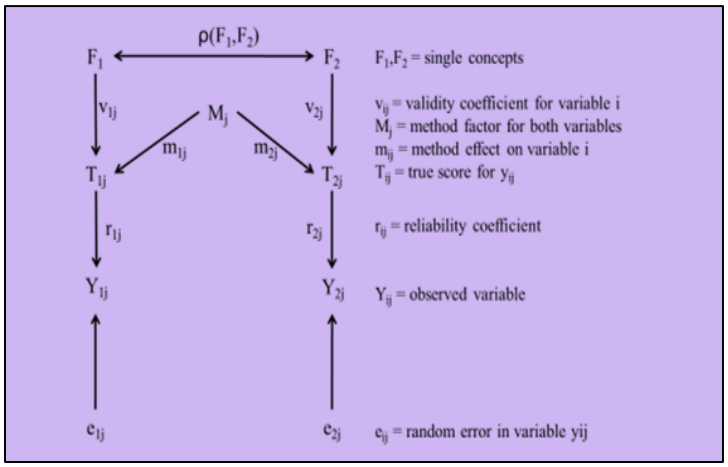


Schéma č. 4: Model měření pro dva znaky měřené stejnou metodou

5.3. Typy měření provázených opakovaně

Předtím, než bude pozornost věnována zásadním ukazatelům či kritériím v otázce kvality měření, především v souvislosti s reliabilitou, je užitečné připomenout rozlišení typů měření, která jsou provázena opakovaně. Tato užitečnost spočívá v jejich klíčovosti pro odhadování náhodných chyb, které vychází z Klasické testové teorie. Porozumění jim je důležité, jelikož každý z typů opakovaného měření stejných proměnných má jiné vlastnosti. Typy měření, které zde rozlišuji, popisuje Kenneth Bollen [Bollen 1989]. Jejich porovnání je možné pokud přepokládáme, že skutečná hodnota se mezi jednotlivými měřeními nezměnila a náhodné chyby nejsou korelované.

Prvním z těchto typů měření je *měření paralelní*. V tomto případě jsou regresní koeficienty závislosti měřené proměnné a skutečné hodnoty stejné a současně rozptyly náhodné chyby jsou taktéž shodné. Tento typ měření je předpokládán v rámci odhadu reliability nejčastěji, nicméně v praxi můžeme pozorovat, že se jedná o měření, které je využívané nejméně [Chylíková 2017]. Druhý typ měření se nazývá *Tau – ekvivalentní měření*, kde jsou regresní koeficienty také shodné, nicméně rozptyly v tomto případě se již liší. V rámci posledního typu měření, kterým je *měření Kongenerické*, se nesetkáme se shodou regresních koeficientů a stejně tak ani rozptyly se v této formě měření neshodují. Jedná se o typ, který je pro odhady nejfrekventovaněji využíván, avšak "teoretickým předpokladům metod odhadu reliability zpravidla nevyhovuje" [Chylíková 2017].

5.4. Mezi ideálním měřením a skutečným měřením - RELIABILITA

Majoritní pozornost v rámci dvojice reliabilita a validita jakožto hlavních psychometrických ukazatelů kvality měřícího nástroje, bývá věnována reliabilitě. To je pravděpodobně z důvodu jejího lepšího statistického uchopení a objektivnosti. Reliabilitu lze popsat jako míru měření konzistentních či ekvivalentních skórů mezi jednotlivými měřeními [Schutt, Engell 2012: 90]. Jinými slovy lze také říci, že v otázce odhadů reliability měření nás zajímá, jak spolehlivý daný měřící nástroj je, tedy zda při replikovaném použití stejného nástroje získáme ekvivalentní hodnoty. Analogicky, jako tomu bylo v případě odvození rovnice KTT, rozsah hledané spolehlivosti měření získáme analyzováním vztahu mezi skutečnou hodnotou a hodnotou, kterou jsme reálně naměřili, přičemž v zájmu kompletnosti tohoto vztahu je třeba vzít v potaz také proměnnou v podobě chyby měření. V tuto chvíli, spolu s níže uvedenou rovnicí pro výpočet reliability a výše uvedenou rovnicí CTST, je zřejmá souvislost

mezi nimi. Statisticky lze koncept reliability popsat jako „sílu vztahu mezi naměřenou odpovědí a její skutečnou hodnotou“ a popsat ji rovnicí Y=X+ E.

Důležité je opět připomenout, že rovnice, vycházející z KTT má 3 předpoklady [Řehák 1998]:

1. chyba *E* není závislá na hodnotě *T*, $r(E,T) = 0$;

2. chyba *E* neobsahuje systematické vychýlení, $\mathbf{E}E = 0$ (průměr chyb při opakovaných

měření je nulový);

3. při různých měřeních jsou chyby nekorelované.

První předpoklad určuje, že střední hodnota (náhodné) chyby se rovná nule, což má za následek, že tato chyba nevychyluje průměr proměnné (X). Druhý předpoklad jasně vypovídá o absenci systematického vychýlení, tedy, že skutečná hodnota je nezávislá na chybě měření – tj. není mezi nimi lineární vztah. Poslední předpoklad se týká nezávislosti jednotlivých měření a tedy následné nekorelovanosti náhodných chyb těchto opakovaných měření.

5.4.1. Odhad reliability

Existuje poměrně velké množství technik vycházejících ze vztahů popsaných v KTT, které umožňují reliabilitu testovat. V základu však můžeme pozorovat dvě hlavní skupiny technik těchto odhadů spolehlivosti měření a to na základě rozlišení dvou typů reliability. Groves poskytuje dělení technik s ohledem na ty, které jsou použity v rámci opakovaného dotazování stejného vzorku respondentů jednotlivými otázkami, a na techniky, které testují reliabilitu otázek sestávajících se z více položek, tedy škál – tzn. odhad reliability nástroje, který používá více indikátorů měřeného konceptu [Groves et al 2004]. K odhadu obou těchto typů reliability je možné využít poměrně širokého spektra přístupů. V rámci první skupiny uvádím z nejfrekventovaněji diskutovaných: test-retest či kvazi-simplex model. Do druhé skupiny pak řadíme techniky jako: Split Half, True Score MTMM model (spojený s odhadem efektu metody i validity) nebo Cronbachovo alfa.

5.4.1.1. Opakované dotazování stejného vzorku respondentů

Jednou z nejoblíbenějších a především, nejčastěji aplikovaných technik je metoda *test-retest*, v rámci které je stejný vzorek respondentů dotazován stejnými otázkami. Použití této metody předpokládá, že jsou opakovaná měření paralelní a její užití je vhodné za splnění daných podmínek, které již byly naznačeny v rámci pojednání o typech opakovaných měření v souvislosti s KTT. Pozorovaný znak se mezi měřeními nesmí změnit, to znamená, že ve skutečné hodnotě nenastala mezi měřeními změna. Dále pak, okolnosti dotazování jsou shodné a jednotlivá měření na sobě nejsou závislá – tedy ideálně, respondent si nepamatuje svou odpověď uvedenou v předešlém dotazování [Engell, Schutt 2012: 90]. Stejně tak bylo zmíněno, že měření by na sobě neměla být závislá, to znamená, že dotazování v čase 1 by nemělo ovlivnit dotazování v čase 2 a dalších. Nicméně již základní zhodnocení popisu této techniky zdravým rozumem může vést k určité skepsi vůči tomuto testu. Popsaný přístup je ve své podstatě nerealistickým modelem z následujících důvodů. Dodržení předpokladu o neměnnosti názoru respondenta mezi jednotlivými měřeními je výrazně ohroženo, pokud mezi těmito měřeními existuje příliš dlouhá doba. Naopak pokud však bude doba mez těmito měřeními krátká, je vysoké riziko, že si respondent zapamatuje odpověď uvedenou v prvním dotazování a tu v rámci dotazování následujícího pouze replikuje, přestože se ve skutečnosti změnila. Schéma č. 5 ilustruje graficky tuto techniku. Schéma, stejně jako všechna ostatní v této kapitole, jsou vytvořena na základě schémat uvedených Sarisem a Gallhoferem v jejich knize [Saris, Gallhofer 2007]. F(i) je proměnná, která má být měřena a y(i1) a y(i2) jsou odpovědi na otázky měřící zmíněnou proměnnou. Vzhledem k podmínkám, které tato metoda musí dodržet (shodné měření jednou metodou), lze předpokládat, že hodnoty koeficientů reliability budou shodné, to znamená, že hodnota reliability odpovídá korelačnímu koeficientu testu-retestu [Saris, Gallhofer 2007: 191-192].

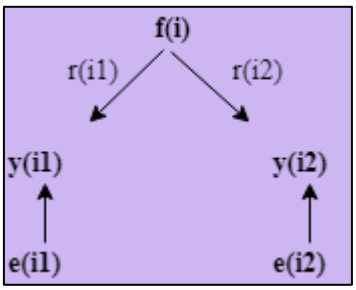


Schéma č. 5: standartní test-retest model

Realističtější model této metody, který navrhují Saris a Gallhofer, rozlišuje mezi latentní proměnnou pro první a druhé měření, jelikož počítají s možnou změnou, která mezi jednotlivými měřeními mohla nastat [Saris, Gallhofer 2007: 191]. Současně je v rámci tohoto upraveného modelu možné pozorovat korelaci mezi jednotlivými chybami kvůli problému s možným zapamatováním si předchozí odpovědi respondenta.

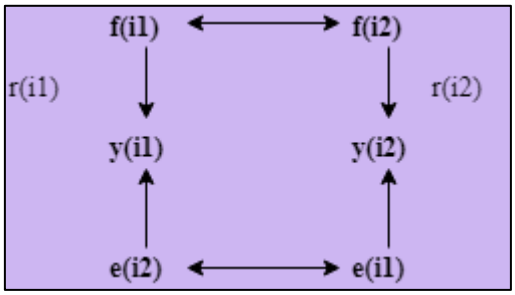


Schéma č. 6: realističtější test-retest model

Druhý přístup v rámci tohoto typu reliability je označován jako *quasisimplex (model)*. Dotazování probíhá oproti předchozímu případu nejen dvakrát, ale alespoň třikrát. Rozdíl oproti předchozímu modelu můžeme pozorovat na schématu č. 7, kde jsou zřetelná jednotlivá (tři) měření.

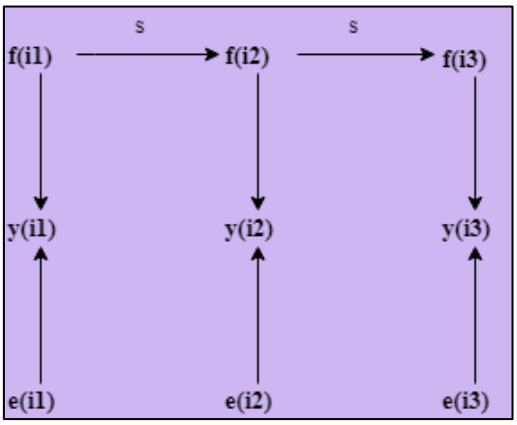


Schéma č. 7: Quasisimplex model pro tři opakovaná měření

Zásadním přínosem tohoto přístupu je možnost nedodržení předpokladu o neměnnosti skutečné hodnoty mezi měřeními a zároveň je možné prodloužení doby mezi jednotlivými pozorováními a tudíž docílení minimalizace „efektu zapamatování si předchozí odpovědi“

respondenta [Saris, Gallhofer 2007: 192]. Nicméně i tento model není dokonalý a má dvě zásadní nevýhody. Zaprvé model předpokládá, že není možné, aby uvažování respondenta spojené s danou proměnnou zmizelo v následujícím (druhém) dotazování a ve třetím se opět objevilo, jelikož by to odkazovalo na vliv proměnné f1 na f3, a tyto vlivy by vedly k nepřesným odhadům [Saris, Gallhofer 2007: 192]. Detailně o tomto problému pojednává Germà Coenders [Coenders et al. 1999]. Druhým zásadním problémem, se kterým se v rámci této techniky potýkáme, je špatné uchopení specifických komponentů, které se vyskytují pouze v některých z po sobě jdoucích měření, respektive, které jsou variabilní. Tímto komponentem může být například nálada respondenta, která se v průběhu dotazování měnila, a při dotazování například na postoj k imigrantům získáme odlišné skóry, avšak přestože se jedná o důležitý komponent pro měřenou proměnnou, je uchopen jako chyba měření (více by tato problematika byla zřetelná například v případě spokojenosti respondenta v životě) [Saris, Gallhofer 2007: 192]. Abychom změnu tohoto komponentu eliminovali, bylo by nutné provést měření v těsné návaznosti, což opět zvýší riziko efektu zapamatování si předchozí odpovědi, viz předchozí technika odhadu reliability. Kvůli povaze a vlivu komponentu proměnné se (chybně) podíl chyby měření zvýší na úkor hodnoty jeho reliability.

5.4.1.2. Mezipoložková reliabilita

Druhým zmíněným typem reliability, potažmo metody jejího odhadu, je mezipoložková reliabilita, tj. reliabilita položek s více indikátory. V tomto případě je středem našeho zájmu vnitřní konzistence měřících nástrojů, konkrétně škál. Existuje opět několik přístupů jak tuto vnitřní konzistenci odhadovat. Prvním z nich je metoda *split-half*. Jak již název této techniky napovídá, princip spočívá v rozdělení škály na dvě části (jakým způsobem závisí na uvážení výzkumníka), následném sečtení skórů pro každého respondenta v rámci skupin, a poté jsou pozorovány dílčí korelace [Engel, Schutt 2012: 91]. Tento přístup se může jevit jako velmi podobný paralelním metodám zmíněným výše, avšak v tomto případě se provádí dotazování ve stejný okamžik a není nutné, aby byla měření ekvivalentní.

Další přístup velmi úzce souvisí s předchozím, respektive z něj vychází. Jedná se o dnes jeden z nejpopulárnějších, ne-li nejpopulárnější, co se testování vnitřní konzistence položek týče. Konkrétně se jedná o přístup, který využívá *Cronbachova alfa*. V praxi si tato metoda klade za cíl nalezení koeficientu, kterým je průměrný skór všech možných kombinací, které je možné vytvořit metodou split-half, čili všech možností rozdělení dané škály na dvě části [Engel, Schutt 2012: 91]. Stejně jako v předchozím případě, měří tato technika ve stejný časový

okamžik, tudíž nehrozí změna skutečné hodnoty. Metodu je také možné využít jak pro paralelní, tak pro tau-ekvivalentní měření.

V rámci výčtu hlavních technik odhadu reliability měřících nástrojů je nutné zmínit významný model *True Score MTMM*, který je v porovnání s ostatními zmíněnými technikami relativně novým. Tato metoda zohledňuje systematické chyby a kombinuje odhad reliability s odhadem validity a efektu metody. Z důvodu této provázanosti dílčích psychometrických ukazatelů kvality měření, je tomuto přístupu věnována pozornost až v rámci teoreticko-analytického rámce práce, kdy již budou vysvětleny ostatní koncepty (validita a efekt metody).

5.5. Mezi konstrukty a měřením - VALIDITA

Zajímáme-li se o validitu našeho měření, pak máme na mysli, do jakého rozsahu nástroj měří skutečně to, co jsme měli v úmyslu měřit. Validitu statisticky vyjádříme jako sílu vztahu mezi proměnnou, kterou chceme měřit a skutečnou hodnotou. I v případě validity měření a jejího dosahování se setkáme, stejně jako u reliability, s více různými přístupy. V rámci této práce uvádím stručně ty nejvýznamnější, kterými je obsahová validita, konstruktová validita či validita kriteriální [Engel, Schutt 2012: 92-93]. Dále je pak stručně v návaznosti na konstruktovou validitu pojednáno o validitě diskriminační a konvergentní, jakožto významných v prostředí MTMM designu.

5.5.1. Typy validit a jejich dosahování

O *obsahové validitě* hovoříme, pokud chceme vidět, do jaké míry měřící nástroj pokrývá všechny významy a dimenze měřeného konceptu [Engel, Schutt 2012: 92]. Dosahování této formy validity je zcela nestatistické. Není vyčíslitelná a je tak do značné míry subjektivní, jelikož mezi prostředky, kterými je jí dosahováno patří rešerše literatury spojené se zkoumaným konceptem či expertní názory, které mají docílit toho, aby byly vybrány nejvhodnější indikátory, které daný koncept měří [Groves 2004]. Následující formy validity jsou již považovány za silnější, jelikož jsou více empiricky založené. Prvním z nich je *validita kriteriální*. Základem tohoto typu validity jsou určitá kritéria, pod kterými je možné si představit vnější zdroj (objektivní, důvěryhodný, již validován a nezávislý na respondentech), se kterým jsou námi získaná data od respondentů následně porovnána [Groves 2004]. Metoda je známa také jako „record check study" [Groves 2004]. Dalším typem validity je *konstruktová validita,* která spočívá na vztahu mezi měřením a ostatními měřeními, která jsou konkretizovaná v teorii. Jednodušeji řečeno, dosahování tohoto typu validity probíhá a závisí na použití

deduktivní teorie s hypotézami ohledně vztahů mezi konstrukty [Koeska 1994 in Schutt, Engel 2012: 93]. Z této definice tak vyplývá, že je velmi důležité, aby daný konstrukt byl správně a přesně formulován a především vztahy mezi jednotlivými proměnnými musí být teoreticky opodstatnitelné, což znamená, že nástroj měření je validní pokud zjištěné vztahy nalezneme také v teorii.

Další typy validity, o kterých bude na následujících řádcích pojednáno, již nepatří mezi základní běžně uvažované typy validity. Avšak s rozvojem MTMM studií je těmto typům validity věnována větší pozornost, jelikož právě v rámci nich, kdy je zkoumáno více znaků více metodami, hrají významnou roli a jsou proto podrobněji zkoumány. Jedná se o *validitu konvergentní a diskriminační*, jež jsou ve vzájemné kombinaci nezbytné pro dosažení konstruktové validity. Diskriminační validity je dosahováno prostřednictvím porovnání skórů měření, které má být validováno a skórů jiného měření stejné proměnné a také skórů proměnných, které měří jiné koncepty, ale jsou související [Campbell, Fiske 1959]. Tato validita je naplněna, pokud validované měření nejsilněji souvisí s porovnávacím měřením a méně s měřením jiných konceptů/konstruktů. Konvergenční validita poté značí, že lze prezentovat vztah mezi dvěma měřeními stejného konstruktu, které byly provedeny různými metodami [Campbell, Fiske 1959]. Jinými slovy konvergentní validita značí, do jaké míry se rozdílné metody shodují na naměřených hodnotách, zatímco diskriminační validita značí, do jaké míry jsou rozdílné znaky rozlišitelné, přestože jsou odhadovány skrze stejnou metodu.

5.6. Efekt metody

Efekt metody byl stručně zmíněn již v úvodu celé této kapitoly pro účel vysvětlení systematické chyby, kde bylo současně graficky znázorněno ovlivnění respondentovi odpovědi danou metodou. Nyní se tomuto vlivu metody věnuji o trochu podrobněji. Campbell a Fiske konstatují, že validita, reliabilita a efekt metody mohou být odhadovány, pouze pokud je použito více metod [Saris, Gallhofer 2007: 174]. Tento jev je obecně chápán jako nechtěný „hluk“ v rámci našich dat, či rozptyl našich dat, který je přivozen metodou, kterou byla data měřena, a ne rozptylem toho, co jsme skutečně chtěli měřit (nejčastěji některý znak osoby) [Saris, Gallhofer 2007: 174-175]. Největším přínosem použití více metod a tak následného identifikování efektu konkrétních metod je, že ačkoli každá z použitých metod pro změření určitého atributu dané osoby, může být nějakým způsobem chybná, je chybná jiným způsobem a proto vypovídající hodnota o datech o dané osobě je více vypovídající, než kdybychom použili pouze jednu metodu [Saris, Gallhofer 2007: 174].

5.7. Statistický odhad celkové kvality měření

Statistický odhad celkové chyby měření je popsán skrze model, který zohledňuje systematickou chybu způsobenou efektem metody a je tudíž modelem více realistickým, než jednoduchý model vycházející z KTT, který zohledňuje pouze chybu náhodou. Vysvětlena již byla podstata a funkce reliability, validity i efektu metody, které tvoří zásadní komponenty, či psychometrické ukazatele, kvality měření. V této kapitole se pozornost přesouvá ke statistickému odhadu kvality měření, který probíhá za podmínek vzetí v potaz popsaných komponentů. Pro tento odhad kvality měření daných otázek v šetření je základním bodem a stavebním kamenem model skutečné hodnoty měření v podobě, ve které byl navržen Sarisem a Andrewsem, pro dva koncepty (znaky) měřené stejnou metodou [Saris, Andrews 1991]. Tento model je ilustrován graficky schématem č. 8, kde jsou naznačeny i popsány jednotlivé komponenty.

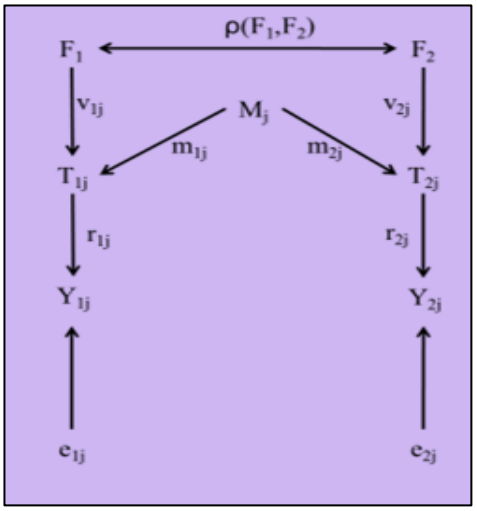


Schéma č. 8: Model měření pro dva znaky měřené jednou metodou

Model umožňuje odhalit proporce kvality měření testovaných otázek. Podstata modelu s jeho hlavními komponenty je následující. Rozdíl mezi naměřenou odpovědí (Yij) a skutečnou hodnotou (Tij) odpovídá náhodné chybě měření (eij) [Saris, Gallhofer 2007: 187]. Koeficienty reliability (Rij) reprezentují efekty skutečných hodnot na měřené hodnoty – tento efekt je tedy menší v případě, že náhodná chyba je větší [Saris, Gallhofer 2007: 188]. Druhá mocnina těchto koeficientů tvoří hodnotu reliability (rij2), která je silou vztahu mezi pozorovanou hodnotou

a skutečnou hodnotou. Koeficienty validity (vij) představují efekt proměnných, které mají být měřeny na skutečné hodnoty proměnných, které jsou reálně měřeny. Druhá mocnina těchto koeficientů je analogicky označován jako validita, která je silou vztahu mezi proměnnými, které mají být měřeny a hodnotou skutečnou [Saris, Gallhofer 2007: 188].

Výše uvedené statistické vztahy v rámci modelu jsou především připomenutím a vyjasněním cesty vedoucí k získání hodnot reliability a validity. Díky takto získaným informacím je možné následně odvodit výpočet kvality měření dané otázky (také označované jako *celková kvalita měření*). Ta je blíže definovaná jako síla vztahu mezi pozorovanou proměnnou a proměnnou, která má být měřena [Saris, Gallhofer 2007: 188]. Celková kvalita měření (qij^2) je tedy produktem reliability a validity a lze ji vypočítat jako:

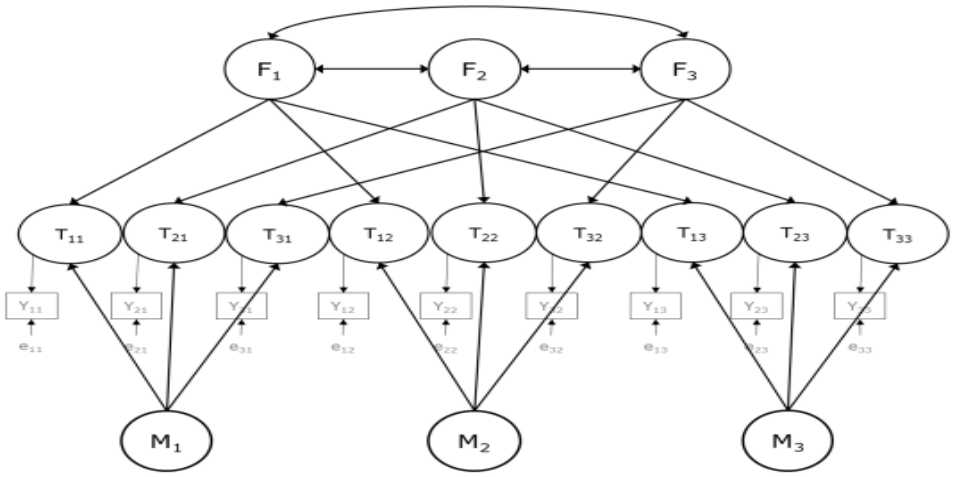
$$qij^2 = rij^2 * vij^2$$

Nyní se ještě jednou podívejme na schéma č. 8. V rámci modelu se potýkáme se dvěma koeficienty reliability, dvěma koeficienty validity, dvěma efekty metody a jednou korelací mezí dvěma latentními znaky, což znamená, že máme 7 neznámých parametrů, ale z dat může být zjištěna pouze jedna korelace. Je nemožné odhadnout oněch 7 parametrů na základě pouze jedné korelace [Saris, Gallhofer 2007: 207]. Z tohoto důvodu byl vyvinut komplexnější přístup, který tyto odhady umožňuje. Jedná se o MTMM design, o kterém blíže pojednává samostatná kapitola v rámci následujícího analyticko-teoretického rámce této práce.

6. MTMM design a odhad kvality měření

V závěru předchozí kapitoly jsem hovořila o problému s odhadem jednotlivých parametrů na základě pouze jedné korelace. Pro řešení tohoto problému musel být vyvinut sofistikovanější přístup, který by takových odhadů byl schopen. Tímto přístupem se stala Multitrait-MultiMethod (metoda či model MTMM), kterou vyvinuli Campbell a Fiske, především pro odhadování konstruktové validity a jejích jednotlivých komponentů viz závěr podkapitoly o validitě [Campbell, Fiske 1959]. V prostředí strukturálního modelování byl model poprvé zahrnut Andrewsem [Andrews 1984]. Metoda je využívána kromě odhadu validity pro současný odhad reliability a také odhad efektu metody. Podstata tohoto typu studií spočívá v měření X znaků prostřednictvím více metod (minimálně a nejčastěji však tří znaků prostřednictvím tří metod), což vede ke vzniku matice, která vypovídá o kvalitě měření zamýšleného konstruktu [Campbell, Fiske 1959].

Abychom byli schopni odhadnout koeficienty vypovídající o kvalitě měření, je třeba dané otázky (měřící různé koncepty/znaky), jež jsou předmětem testování, položit vícekrát (to se děje prostřednictvím aplikace více metod). MTMM design, který je rozšířením modelu uvedeného v předchozí kapitole (odhad celkové kvality měření), je ilustrován graficky viz obrázek č. 2, kde je možné přehledně pozorovat vztahy mezi proměnnými, které mají být měřeny a skutečnými hodnotami [Saris Gallhofer 2007: 208]. Z grafu je zřejmé, že mezi znaky (koncepty) existuje korelace, ale mezi jednotlivými metodami, které je měří nikoli.



Obrázek č. 2: vztahy mezi proměnnými, které mají být měřeny a skutečnými hodnotami (zdroj: Saris, Gallhofer 2007)

6.1. Survey Quality Predictor

Saris, Gallhofer a van der Veld MTMM design podrobně analyzovali včetně poznatků, ke kterým se v rámci jejich použití došlo [Saris, Van der Veld, Gallhofer 2004]. Na základě této praxe Van der Veld vyvinul algoritmus, který je schopen odhadovat kvalitu otázek v duchu designu MTMM [Saris, Van der Veld, Gallhofer 2000]. V rámci metodologie šetření se jedná o významný počin, který má obrovský přínos v oblasti výzkumu a zdokonalování kvality měření. Zmíněný algoritmus je základem nástroje SQP, jehož název je zkratkou pro *Survey Question Predictor*. Do češtiny toto spojení můžeme přeložit jako software pro předpověď kvality otázek. Jeho největším přínosem je schopnost namísto provázení dalších MTMM experimentů odhadnout jednotlivé koeficienty vypovídající o kvalitě, které je následně možné

využít také pro korekci daného měření, aniž bychom sbírali nová data. V rámci mezinárodního porovnání často dojdeme k lišícím se hodnotám u kvality dat jednotlivých zemí. SQP pomáhá odhadnout zátěž dané proměnné chybou také v konkrétní zemi. Saris a Gallhofer shrnují funkce SQP do následujících čtyř bodů [Saris, Gallhofer 2007: 257].

1. čtení a automatické kódování položek šetření v dotazníku
2. předpověď kvality navržených otázek
3. poskytnutí informace o efektu jiných možností
4. poskytnutí návrhu na úpravu položek

V praxi je fungování tohoto nástroje založeno na identifikaci vstupních charakteristik otázek a na základě koeficientů validity a reliability, které pocházejí z existujících MTMM experimentů. Tyto koeficienty tak pomáhají odhadnout koeficienty validity a reliability v podstatě jakékoli otázky, a to na základě zmíněných jednotlivých charakteristik, kterými může být: jazyk, téma otázky, škála měření, a jiné [Chylíková 2015].

7. Koncept ekvivalence v metodologii sociálních věd

Hlavním tématem této práce je kvalita měření, respektive psychometrické aspekty měřících nástrojů. Vysvětleny byly již pojmy jako reliabilita čili spolehlivost těchto měření, validita čili přesnost toho měření a vysvětlen byl také dopad a využití efektu metody na měření. Také je již zřejmé, jak za pomoci těchto kazatelů odhadovat celkovou kvalitu daného měření. Nicméně s ohledem na kvalitu měření a kvalitu dat je nutné zohlednit další významný faktor. Šetření v sociálních vědách nejsou téměř nikdy izolovanými procesy, ale jsou zasazeny do určitých kontextů. V rámci výzkumů porovnáváme různé skupiny, ať už se jedná o věk, národnost a podobně. Především je však toto téma spjato s mezinárodními výzkumy. A přestože sociologie již od svých počátků staví na empirickém srovnávání společenských systémů, což je patrné z četných děl a výzkumů již nejvýznamnějších klasiků sociologie, a srovnávání je podstatou samotné definice sociologie, srovnatelnosti či ekvivalenci, která je vedle reliability či validity stejně důležitým ukazatelem kvality, není věnována stejná pozornost, naopak je v rámci analýz často upozaděna především v perspektivě tuzemských výzkumů. V zahraničí se tomuto problému v rámci aplikovaného výzkumu věnují výzkumníci intenzivněji [Byrne & Campbell 1999; Chen, Sousa, & West 2005; Hendriks et al. 2003; Kwan, Bond, & Singelis 1997; Little 1997; Rhee, Uleman, & Lee 1996; Steenkamp & Baumgartner 1998].

Vnímání každého jednoho člověka na tomto světě je unikátní a stejně tak vnímání otázek, samotného konceptu či položek a jednotlivých možností odpovědí se liší v závislosti na poměrech v dané zemi (politických, ekonomických atd.), hodnotách respondenta, jazyku či obecně kultuře vážící se k dané společnosti, jíž je členem. Zmíněné faktory, mají v takovém mezinárodním komparativním výzkumu zásadní roli, jelikož jsou oproti monokulturním výzkumům mnohonásobné, komplexnější, a proto i riziko systematické chyby je přirozeně větší. Diferenciaci v rámci vnímání lze příhodně pro tuto práci ilustrovat za pomoci problému řešeného v práci Davidova a Schmidta, který se týká odlišného vnímání slova imigrant [Davidov, Schmidt 2007]. Autoři ukazují, že je zásadní, jaká kulturní situace v dané zemi panuje, jelikož například v rámci Evropy je imigrant často viděn jako přistěhovalec, se kterým se pojí problémy, zatímco například v Izraeli je imigrantem myšlen Žid s veškerými právy a povinnostmi v dané zemi [Davidov, Schmidt 2007]. Riziko však nekončí na úrovni pochopení, ale i v ideálním případě, kdy všichni respondenti pochopí otázku a jednotlivé koncepty stejným způsobem, stále existuje riziko selhání na úrovni měření, které nevyhnutelně nastane, pokud v některé zemi existuje například tendence (kulturně podmíněná) uchylovat se ke krajním hodnotám. Takové tendence prokazatelně existují. Na příkladech je demonstrují Smith s Fischerem [Smith, Fischer 2008].

Vzhledem k nedostatečné pozornosti, která je konceptu ekvivalence v našem prostředí a samotnému procesu jejího dosahování věnována, a také proto, že hlavní zdroje dat o (anti)imigračních postojích, která jsou v ČR využívána, jsou mezinárodního charakteru, objasním v této kapitole pojem ekvivalence, popíši jeho dimenze a klasifikaci a v neposlední řadě se zaměřím na metody, kterými je ekvivalence dosahováno či jak je posuzována.

7.1. Definice ekvivalence

Než bude uvedena samotná definice či popis konceptu ekvivalence, považuji za vhodné alespoň velmi stručně zmínit okolnosti jejího vzniku, respektive vzniku počáteční pozornosti, která začala být ekvivalenci věnována a odstartovala rozvoj v této oblasti (aplikovaného) výzkumu. Již bylo zmíněno, že sociologie je komparativní vědou a proto se (mimo jiné) srovnávání věnuje již od svých počátků. Nicméně až v polovině minulého století se začala budovat a etablovat mezinárodní komparativní sociologie jako jedna z disciplín sociologie a vznikla tak mezinárodní sociologická asociace (ISA) [Marsh 1967; Payne 1973]. S tím se neodmyslitelně pojí průkopnický výzkum Almonda a Verby, kteří demonstrují odlišnosti jejich výzkumu zohledňujícího mnohonásobnost, v porovnání s monokulturními výzkumy [Almond,

Verba 1963]. Přirozeně byl tento prvotní výzkum terčem četné kritiky, což postupně vedlo k rozvoji oboru, který se začal kompletně věnovat metodologii šetření, která jsou komplexnější.

Jednotná definice ekvivalence, stejně jako mnohých dalších konceptů v sociologii, neexistuje. Pokusím se zde ale stručně pojem vysvětlit na základě hlavních rysů a příkladů definic některých z autorů intenzivně se věnujících tomuto tématu. Nejčastěji je tato srovnatelnost i samotný koncept ekvivalence spojován s mezinárodními studiemi a porovnáváním odlišných kultur, respektive společností (či populací) – tj. zda zkoumaný fenomén (konstrukt) je vůbec možné ve všech zkoumaných společnostech měřit a následně porovnávat [Anýžová 2014]. Zásadní komparace však probíhá například i na úrovni časové a v kontextu mnoha jiných skupin. Z výše zmíněného vyplývá, že v rámci komparativních studií, hovoříme-li o ekvivalenci, je žádoucí, aby daný sociální fenomén, měřící položky a škály měření byly pochopeny, interpretovány i vnímány stejně napříč všemi respondenty pocházejícími z různých společností a kulturních prostředí, dotazovaných v různých časových bodech. V převážné většině případů padá významná část pozornosti na jazykovou a významovou stránku věci. Nicméně ekvivalenci je třeba se věnovat v mnoha dalších oblastech. Jednotlivé dílčí oblasti lze identifikovat z definice ekvivalence McArdleho a Horna, kteří shledávají ekvivalentní měření jako takové, které díky řádné operacionalizaci vede ke změření totožných znaků i v rozdílných podmínkách studia daného fenoménu, přičemž těmito rozdílnými podmínkami je: nejednota populací, různé metody sběru dat, organizace řídící tento sběr či lišící se čas sběru dat [Horn, McArdle 1992]. Shrneme-li všechny uvedené informace, ekvivalence zjišťuje, zda daný měřící nástroj měří stejným způsobem, stejný koncept, ve všech podskupinách dotazovaného vzorku respondentů.

7.2. Kategorizace ekvivalence

Z předchozí podkapitoly a definice ekvivalence vyplývá, že existuje více jejích druhů. Dokonce Johnson ve své práci uvádí až na 50 druhů ekvivalence, které byly v rámci odborné literatury na toto téma rozpracovány [Johnson 1998: 6-7]. Nicméně přestože ekvivalence podléhá značné kategorizaci, a ačkoli existuje více způsobů jejího dělení a výsledných klasifikací, základním a společným dělením je rozvětvení ekvivalence na *úrovni měření* a na *úrovni teorie* [Jöreskog, 1971]. Pro účely této práce je stěžejní koncept ekvivalence na úrovni měření, nicméně v rámci následujících kapitol objasňuji i druhou zmíněnou úroveň ekvivalence.

7.2.1. Ekvivalence na teoretické úrovni

Na úrovni teorie se teorie a praxe ekvivalence zabývá jednotlivými koncepty, samotným konstruktem či následně vytvořenými indikátory a jednotlivými položkami, které mají být ideálně srovnatelné. S ohledem na tyto dílčí koncepty se ekvivalence na teoretické úrovni může rozvětvit na ekvivalenci *konceptuální* (jejím cílem je, že konceptu či danému sociálnímu jevu všichni respondenti rozumí stejně a připisují mu stejný význam), *operacionalizační* (indikátory, které měří daný konstrukt, jsou všemi respondenty stejně interpretovány) a *položkovou* (položky měřícího nástroje všichni respondenti chápou stejně, nehledě na kulturu) [Hui, Triandis 1985].

7.2.2. Ekvivalence na úrovni měření

Přesuneme-li se na úroveň ekvivalence měření, musíme přesunout také pozornost od konceptů, konstruktů a indikátorů na škály měřícího nástroje a jednotlivé kategorie odpovědí. V tomto případě je dosaženo ekvivalence čili srovnatelnosti, pokud všichni respondenti vnímají škály stejným způsobem, stejně tak jednotlivé kategorie odpovědí, a využívají je zcela shodným způsobem [Anýžová 2014]. Bohužel tohoto požadavku není vůbec lehké dostát, jelikož jak například Saris uvádí, škály měření jsou zatíženy systematickou chybou měření, a výběr odpovědí respondentů se vždy kulturně liší [Saris 275 in Harknes 2003]. Jinými slovy lze ekvivalenci měření popsat jako proceduru, která zjišťuje, zda má daný nástroj stejné psychometrické vlastnosti napříč různorodými skupinami [Chen 2007].

V literatuře, která je převážně zahraničního charakteru, se často setkáme s analogickým označením „measurement equivalence“ nebo „measurement invariance“. Todd Little rozlišuje dvě skupiny této srovnatelnosti na úrovni měření. V rámci první skupiny je hodnocena neměnnost, která souvisí s psychometrickými aspekty měřících škál, druhá skupina pak úzce souvisí s mezi skupinovými rozdíly u průměrů, odchylek a kovariancí, přičemž první zmíněná skupina je jakousi prerekvizitou skupiny druhé, a až na jejím základě je možné interpretovat rozdíly v rámci právě druhé skupiny, které jsou pro výzkum zpravidla zásadní [Little 1997]. S ohledem na otázku neměnnosti měření lze rozlišit tři hlavní podkategorie na základě úrovně, které může měření dosáhnout. Tyto úrovně vycházejí z testování modelů měření konfirmační faktorovou analýzou, které navrhl Jöreskog a jsou nezbytné, chceme-li odpovědět na korespondující dílčí výzkumné otázky[Joreskog 1971]. Těmito úrovněmi jsou: skalární ekvivalence položek (scalar invariance/measurement intercepts), která má nejmenší

systematickou chybu měření, metrická ekvivalence (metric invariance/measurement weights) a konfigurální ekvivalence (configural invariance/unconstrained measurement) položek, která má systematickou chybu měření největší [Joreskog 1971]. Konfigurální ekvivalence je pomyslně nejnižší úrovní ekvivalence měření, v rámci které se setkáme s největší systematickou chybou měření. Jedná se o případ, kdy latentní proměnné (konstrukty) nejsou identické a srovnatelnost je tedy možná pouze na úrovni zkoumaného fenoménu [Chen 2007]. Je vyžadováno, aby se shodovaly napříč skupinami pouhé struktury faktorů. V praxi tato podmínka znamená, že daný koncept může být smysluplně diskutován v rámci všech zemí. Pokud je tato úroveň ekvivalence naplněna je možné se v rámci analýzy přesunout k vyšší úrovni, na které je odhadována metrická ekvivalence. V rámci této úrovně jsou testovány faktorové zátěže a pro naplnění této úrovně ekvivalence je vyžadováno, aby faktorové zátěže mezi pozorovanými položkami a latentní proměnnou byly shodné napříč skupinami [Davidov 2014]. Opět jinými slovy lze říci, že tato ekvivalence zajišťuje shodnost intervalů škály, která měří daný konstrukt, tedy že význam toho, když dojde ke zvýšení o jednotku na měřící škále, je stejný v jednotlivých skupinách [Bollen 1989; Davidov 2014]. Předmětem zájmu této úrovně ekvivalence je (ne)jednota ve vnímání kategorií a velikostí škál napříč jednotlivými kulturami (skupinami) [van de Vijver, Leung 1997]. Konečně skalárně ekvivalentní měřící nástroje předpokládají škály, které jsou si nejpodobnější. Použité škály napříč společnostmi/skupinami mají stejnou velikost či kategorie a jsou respondenty stejně vnímány a chápany, a proto nehrozí takové systematické vychýlení odpovědí.

7.3. Dosahování/testování ekvivalence – úroveň teoretická

V této části výkladu se zaměřím na stručný popis dosahování ekvivalence na úrovni teoretické, jelikož jak vyplývá již z výše zmíněného, bez dosažení této úrovně srovnatelnosti, není přirozeně možné dosahovat ekvivalence na úrovni měření. Následující řádky proto popisují praktiky a přístupy, které se váží k procesu, který bezprostředně předchází fázi měření. Pozornost je zde kladena na koncepty, na jednotné pochopení jednotlivých pojmů všemi respondenty, na překlady, zkrátka kompletní vyvinutí otázek, které budou ve výsledku v ideálním případě ekvivalentní. Do metod dosahování ekvivalence na teoretické úrovni řadíme především: otázku poměru etických a emických položek, překlad dotazníku a aplikaci kognitivního přístupu při testování jednotlivých otázek.

7.3.1. Etické x emické položky

Zásadní otázkou, kterou si metodologové a analytici musí položit je, zda jsou použité koncepty emické nebo etické podstaty. Stejná otázka se naskýtá pro jednotlivé položky. Rozdíl mezi těmito dvěma typy je následující. Etické položky respektive koncepty jsou takové, které řeší aspekty lidského života jako názory, chování apod., které sdílí všichni lidé bez ohledu na jejich příslušnost dané kultuře [Davidson et al. 1976; Harris 1976; Ægisdóttir et al. 2008]. Naopak emické koncepty jsou ty, které zastávají vzorce myšlení a chování specifické pro danou kulturu, které je respondent příslušníkem (to je například i postoj k imigraci a imigrantům). Ze zmíněného vyplývá, že je více, než užitečné v rámci mezinárodních komparativních studií využít etických konceptů a položek v části dotazování, která bude všem respondentům společná avšak současně část dotazníku vyčlenit pro položky a koncepty emické, které daná země vyvine a uzpůsobí sama pro své vlastní využití. Zásadní chybou a ohrožením následného měření by proto bylo, pokud by byl emický koncept, jako jsou například anti (imigrační) postoje, měřen etickými položkami. To znamená, byl by měřen ve všech kulturních prostředích zcela stejně bez zohlednění specifičnosti tohoto prostředí, což by přirozeně vedlo k systematickému vychýlení. Z tohoto důvodu, abychom předešli umělému vytváření kulturních odlišností, je doporučeno pro měření emických položek, jako jsou (anti)imigrační postoje, použití výhradně emických konceptů i presto, že bude snížena úroveň ekvivalence [Anýžová 2014]. Etické koncepty je možné měřit jak emickými tak etickými otázkami. Propojení této teoretické části metodologie s částí zkoumající srovnatelnost na úrovni měření je zřejmé. Etický koncept, který je měřen emickými otázkami, které zní v každém jazyce jinak, není tolik srovnatelný mezi kulturami, respektive nejsou srovnatelné jednotlivé položky. Nicméně jak uvádí ve své práci Anýžová, teoretické koncepty srovnatelné jsou a nízká míra ekvivalence se tak může objevit pouze v rámci rozhodnutí o konkrétní statistické analýze [Anýžová 2014]. V případě ekvivalence skalární je již nutné měřit etické koncepty etickými otázkami, které jsou doslovně přeložené [Anýžová 2014].

7.3.2. Překlad dotazníku

S rozlišením etických a emických položek se úzce pojí překlad dotazníku. Jak vysvětlují van de Vijver a Leung, je otázkou, jak by se k překladu mělo přistupovat, s ohledem na odlišné skupiny konceptů viz předchozí odstavec. Autoři navrhují, že doslovně by měly být překládány otázky a koncepty etické, které jsou použity u zemí, které mají podobné jazykové základny či kulturní prostředí, a naopak v případě zemí, které se kulturně či jazykově značně liší, jsou

doporučovány modifikace v překladu [van de Vijver 1997: 265]. Z předchozího jasně vyplývá, že výraznou výhodou etického přístupu a překladu je minimalizace nákladů díky univerzálnosti jeho použití a v případě stejného vnímání je jeho devizou také vysoká srovnatelnost dat a tudíž možnost podrobné statistické analýzy. Oproti tomu emický přístup statistickou analýzu znesnadňuje díky nízké ekvivalenci na úrovni měření, avšak jeho přidanou hodnotou je hlubší porozumění kulturně specifickým oblastem [van de Vijver 1997; Anýžová 2014] Vzhledem k povaze obou přístupů je vhodné adekvátně kombinovat oba dva přístupy a to s ohledem jak na přítomnost etických a emických konceptů tak na povahu zkoumané populace.

Technické provedení překladu, který je zásadní v rámci mezinárodních šetření skýtá několik možností. Nejpoužívanější technikou je metoda zpětného překladu. Postup této metody rozpracovává například Harkness a Schoua-Glusberg. Metoda spočívá v přeložení originální verze do rodného jazyka respondenta dané země, tomu následuje opětovné přeložení do původní verze - v ideálním případě by každý z překladů měl provádět jiný člen výzkumného týmu - a následně je porovnán originál s opětovným překladem do původní verze, přičemž cílem je, aby tyto verze byly co nejpodobnější (v opačném případě se pokračuje v překladech dál) [Harkness, Schoua-Glusberg 1998: 97; Harkness 2003: 41]. Vylepšením metody zpětného překladu je metoda tzv. TRAPD (Translation, Review, Adjudication, Pretesting, Documentation), která zakládá na dvou nezávisle provedených překladech, poté jejich revizi/korekci a následném posouzení [Dorer 2003]. Celý průběh této techniky je současně důkladně dokumentován. V případě provedení překladu se objevuje otázka, zda byl tento překlad skutečně úspěšný. Z tohoto důvodu se v rámci pilotáže tato kvalita překladu testuje, a to opět mnohými způsoby. Mezi přístupy takového testování mohou být zařazeny posudky bilingvních osob či přístupy statistické, které využívají porovnávání skórů a t-testů. V neposlední řadě jsou využívány také kognitivní rozhovory, které se nepojí pouze s překladem, ale i dalšími aspekty kvality měřícího nástroje a respondentovým porozumění otázkám a „uživatelské přívětivosti“ nástroje. Těmto rozhovorům se velmi stručně věnuji v následující podkapitole.

7.3.3. Kognitivní metody

Kognitivních metod, které se dnes používají je široké spektrum. Patří mezi ně: debriefing, expertní posudky, kódování chování, měření reakčního času a jiné [Vinopal 2008: 50-57]. Nicméně pro typ šetření, kterému se věnuji v analytické části práce (mezinárodní komparativní studie), je využíváno převážně metod kognitivních rozhovorů, případně debriefing. V závěru

předchozí kapitoly bylo naznačeno, že kognitivní rozhovory nejsou technikou využívanou pouze v rámci kvality překladů, ale obecně zhodnocují otázky daného instrumentu, čili zjišťují, jak respondenti na dané části dotazování reagují. Proto můžeme říci, že zásadním rozdílem běžně uvažovaných a kognitivních rozhovorů je, že v prvním případě je naším cílem získat věcnou odpověď na danou otázku, zatímco v druhém případě nás zajímá proces, který vedl ke konkrétní odpovědi [Vinopal 2008: 51]. Definic tohoto přístupu existuje mnoho a panuje mezi nimi značná nejednota i z důvodu četnosti technik, které jsou v rámci něj využívány. To vede metodology k užívání kognitivního rozhovoru spíše jako přístupu, než konkrétní techniky. Jiří Vinopal kognitivní přístup v testování kvality nástroje popisuje jako přístup, který se zabývá chybami měření (pozorování), které mohou vznikat ve spojitosti s použitým měřícím nástrojem, tazatelem, metodou sběru dat a respondentem [Vinopal 2009: 399].

7.4. Testování/ odhad ekvivalence – úroveň měření

Ekvivalence na teoretické úrovni, o které bylo hovořeno v předchozích odstavcích, nelze testovat, respektive nelze ji zpětně a statisticky testovat, a je tedy otázkou pouze přípravné fáze. Naopak v případě vzniklých indikátorů je potřeba danou ekvivalenci otestovat po jejich použití a před použitím dalším, jelikož je žádoucí zjistit, zda jsou rozdíly v odpovědích způsobeny kulturní odlišností nebo nejednotou interpretací položek. Toto testování probíhá zpravidla statistickými metodami. Existuje dlouhý výčet technik takového testování, proto zájemce o detailní informace, odkazuji na Brauna a Johnsona, kteří poskytují skutečně vyčerpávající výčet těchto technik včetně faktorů hrajících zásadní roli při jejich volbě [Braun, Johnson 2010: 377; Johnson 1998: 25-29]. Tato práce uvádí pouze stručný výčet a popis hlavních technik, které jsou nejvyužívanější.

Mezi nejzákladnější či nejjednodušší techniky spadá souhrnně označena *analýza prvního a druhého stupně*. V rámci využití této techniky je možné srovnat data s ohledem především na zkoumaný problém, posoudit, zda jsou tato data srovnatelná, avšak problematická část nastává v případě hlubší metodologické analýzy, pro kterou již tato cesta dostačující není [Braun, Johnson 2010: 375]. Kromě nepřehlednosti, která nastane v případě srovnávání většího počtu zemí, se setkáme i s dalšími úskalími této techniky jako je komplikace v případě chybějících hodnot aj. [Braun, Johnson 2010: 378-381]. Pro jasnější představu, do těchto technik řadíme například t-testy či analýzy rozptylu, případně také korelace.

Druhá skupina zahrnující o mnoho komplexnější a náročnější analýzy, je hromadně označena jako *mnohorozměrná statistická analýza*. V rámci jednotlivých analýz je možné použít velké množství proměnných, skupin (například zemí), které srovnáváme, a které můžeme do analýzy zařadit. Mezi techniky mnohorozměrné statistické analýzy patří například analýza latentních tříd (LCA), teorie odpovědi na otázku (IRT), mnohorozměrné škálování. Pro potřeby testování ekvivalence se nejčastěji využívají techniky tzv. strukturního modelování (SEM) v rámci kterého je provedena více skupinová konfirmační faktorová analýza (MGCFA). Z tohoto důvodu a také z důvodu vlastní analýzy s využitím strukturního modelování v rámci analytické části práce, se zde zabývám pouze popisem práce v rámci prostředí SEM, samotným modelováním a konkrétní statistickou metodou běžně využívanou pro testování ekvivalence měření, kterou je MGCFA.

7.4.1. Strukturní modelování

Jelikož je strukturní modelování v dnešní době nejvíce využívanou metodou (pro testování ekvivalence na úrovni měření), a aplikuji ho v rámci své analýzy této práce, věnuji mu pozornost na úkor ostatních technik, které v rámci této práce z důvodu omezeného rozsahu podrobně popsány nejsou. Strukturní modelování je často označované zkratkou SEM z anglického překladu „Structural Equation Modeling“. Označení strukturní modelování popisuje statistickou praxi, která využívá statistických modelů, které jsou sestaveny z více strukturálních rovnic. Jinými slovy lze říci, že se jedná o modely, které jsou tvořeny jak matematicky, tak graficky na základě určité teorie vypovídající o vztazích proměnných a skutečnostech skrývajících se za daty a kauzálními vztahy [Matějů 1989: 402]. Tato technika či spíše soubor statistických technik je založen na regresní analýze a analýze faktorové, jelikož je možná práce jak s přímo pozorovanými tedy manifestními proměnnými, tak proměnnými latentními, které přímo pozorované sice nejsou, ale v rámci modelu je na základě manifestních proměnných definujeme [Matějů 1989]. Je důležité se v rámci strukturního modelování oprostit od běžně uvažovaných závislých a nezávislých proměnných, se kterými pracujeme v rámci zmíněné dílčí regresní či faktorové analýzy, jelikož v prostředí strukturních modelů se můžeme setkat s proměnnými, které jsou současně proměnnou závislou i nezávislou. Proto si při práci se strukturními modely musíme osvojit jinou terminologii proměnných a jejich dělení na endogenní a exogenní. To znamená na ty, které jsou ovlivňovány alespoň některou proměnnou v našem modelu (endogenní) a na ty, které nejsou ovlivňovány žádnou (exogenní).

V rámci strukturního modelování je možné testování od jednoduchých vztahů a korelací či lineárních regresí až po komplikované modely kombinujících více technik. Nejjednodušším případem je model, kdy máme pouze dvě proměnné (X a Y), předpokládající lineární vztah a komponent u, který představuje všechny proměnné, které ovlivňují Y, ale nejsou explicitně zmíněny [Saris, Gallhofer 2007: 196]. Jedná se o jednoduchou lineární regresi.

Mezi komplikovanější modely se poté řadí testování použité v rámci mezinárodních šetření, která jsou provázena v různých kulturních prostředích a současně v periodických opakováních. Pro analýzy tohoto rozměru je nejčastěji využíváno konfirmační faktorové analýzy více skupin. Možnost testování v softwarech umožňujících konstruovat komplexní modely je proto velkou výhodou. Proces testování ekvivalence měření je součástí analytické části mé práce, kde provádím více skupinovou konfirmační analýzu (MGCFA) na datech z Evropského šetření European Social Survey mapujících (anti)imigrační postoje obyvatel jednotlivých států.

7.4.2. Více skupinová konfirmační faktorová analýza

Základním přístupem, který volím pro analýzu srovnatelnosti dat ESS je, jak již bylo zmíněno výše, využití rozšířené konfirmační faktorové analýzy, jelikož ekvivalenci měření zde stavím do pozice hypotézy, kterou je třeba testovat. Již výše jsem zmínila, že existuje několik technik testování ekvivalence. Více skupinová konfirmační faktorová analýza je však nejvyužívanějším způsobem.

V souvislosti s více skupinovou konfirmační faktorovou analýzou, která je v analytické části prostředkem pro posouzení ekvivalence měření je důležité porozumět či alespoň stručně se seznámit s tzv. indexy vhodnosti modelu, které jsou významnými výstupy analýzy. Jsou důležité, jelikož statistiky chi-kvadrát testu jsou příliš citlivé na velikost vzorku respondentů a nedodržení předpokladu o normalitě [Davidov 2014]. Jedná se o indexy, které porovnávají vhodnost základního a uvažovaného modelu. Vzhledem k zaměření a omezenému rozsahu této práce zde představím pouze indexy základní, které jsou nejčastěji interpretované, a které interpretuji také já v rámci své analýzy. Prvním z nich je RMSEA (Root mean square error of approximation). RMSEA je absolutním indexem, který není citlivý na velikost dotazovaného vzorku, ale je citlivý na komplexitu modelu [Chen 2007, Van de Schoot, Lugting, Hox 2012]. Ideálně by hodnota tohoto indexu měla být vyšší, než 0,8. V případě CFI, které je také relativně

nezávislé na velikosti vzorku respondentů, je vhodnost modelu považována za adekvátní, pokud hodnoty překračují 0,9 [Van de Schoot, Lugtig, Hox 2012].

Na závěr této kapitoly zde ještě uvádím tabulku, kterou přehledným způsobem zkonstruovala na základě předlohy Brauna a Johnsona pro potřeby své dizertační práce Petra Anýžová. Je možné v ní pozorovat množství existujících technik testování ekvivalence na úrovni měření i s příklady jejich užití v konkrétních studiích, na které se může čtenář s větším zájmem o ně obrátit.

Tabulka 1 Vybrané mnohorozměrné statistické techniky pro testování ekvivalence postojových škál po sběru dat

Metoda	Vhodný typ proměnné	Empirické studie – aplikace metody na datech	Typ ekvivalence
Explorační faktorová analýza	Kardinální data Ordinální data* Nominální data (CATPCA)	Welkenhuysen-Gybels, van de Vijver 2001	Konstruktová
Testování reliability škál	Kardinální data Ordinální data*	Devins et al. 1997	Překladová
Mnohorozměrná korespondenční analýza	Kategoriální data (grafická metoda)	Blasius, Thiessen 2006	Konstruktová
Mnohorozměrné škálování	Ordinální data (grafická metoda)	Braun, Scott 1998; Fontaine 2003	Konstruktová
Víceúrovňové modelování	Kardinální data Ordinální data*	van de Vijver, van Hemert, Poortinga (eds.) 2008	Konceptuální
Strukturní modelování	Kardinální data Ordinální data* Nominální data (za podmínky vypočítání matice jiných než Pearsonovských korelací, např. polychorických, PRELIS)		Konfigurální Metrická Skalární
a) strukturní modelování		Mullen 1995; Billiet 2003; Byrne, Watkins 2003	
b) víceúrovňové strukturní modelování		van Hemert et al. 2002; van de Vijver, Poortinga 2002; Davidov et al. 2012	
c) konfirmační faktorová analýza		Hsueh et al. 2005; Lievens et al. 2007; Davidov 2010	
d) MTMM design		Saris 2003; Saris, Gallhofer 2007; Revilla, Saris 2011	Konceptuální
Analýza latentních tříd	Kategorická i kardinální data	Kankaras, Moors 2009	Konfigurální Metrická Skalární
Teorie odpovědi na položku	Dichotomická data Ordinální proměnné	Wang, Russel 2005; May 2006; Woehr, Meriac 2010	Překladová Skalární

Zdroj: vlastní zpracování podle Brauna a Johnsona [2010: 377]

8. Měření (anti)imigračních postojů

Denně nebo přinejmenším velmi často, se účastníme diskuzí týkajících se migrace, aktérů jednajících v rámci ní či dopadů, které tento jev přináší, nebo jsme alespoň svědky takových diskusí a četných zpráv, které tuto situaci hodnotí. Ať už se jedná o první či druhou variantu a jakýkoli obsah těchto diskuzí na téma migrace a především imigrace, nelze přehlédnout, že toto téma s sebou přináší značné emoce a mnohé diskuze často ústí ve vyhrocené (nejen) ústní souboje. Přirozeně si lidé na takto diskutovaný a zásadní fenomén vytvoří názor a určitý postoj k němu. Z velkého množství výzkumů provedených na téma (anti)imigračních postojů lze vypozorovat určité tendence. Negativní postoje na základě těchto empirických dokladů jsou silnější mezi staršími lidmi, lidmi s nižším vzděláním a nižšími příjmy. V návaznosti na tento stav převládající v mnoha společnostech se vědci pokoušeli analyzovat, na základě čeho tento postoj vzniká, jak se liší napříč kulturami a co jsou hlavní faktory ovlivňující tyto postoje ať už na makro či mikroskopické úrovni. Není v úmyslech této práce se svým výzkumem přidat do řad těchto studií. Ačkoli je inklinace české populace k určitým postojům naznačena při analýze proměn postojů v čase, hlavním předmětem následující analýzy je měřící nástroj, který postoje jedinců k imigrantům a imigraci měří a jeho kvalita.

Česká republika včetně dalších postkomunistických zemí, je svými postoji k imigraci a imigrantům jedním z výrazně specifických prostředí. Specifické je především z důvodu vysoce negativních postojů k imigraci a imigrantům navzdory malému podílu na kontaktu s nimi či zkušenostmi s negativními jevy s imigranty spojovanými (násilí, útoky aj.). O této specifičnosti zde mluvím, protože zdroje dat o postojích k imigrantům a imigraci jsou ve velké většině mezinárodního charakteru komparativních studií a je otázkou, zda v takto specifickém prostředí jsou získaná data porovnatelná s daty, která jsou získaná v prostředí, které je specifické například zcela opačným způsobem (pozitivní vliv k imigrantům i přes bohaté zkušenosti s negativními jevy s nimi spojovanými atp.) či zcela jiným kulturním prostředím a vnímáním tohoto fenoménu.

V teoretickém ukotvení celé této práce bylo vysvětleno, jak narůstá komplexnost potenciálních chyb v prostředí mezinárodních studií a jak se současně násobí riziko ohrožení měření chybami. Zásadní roli vedle reliability a validity či efektu metody, hraje v takových výzkumech koncept ekvivalence. Ta je předmětem analýzy této práce. Konkrétně je analyzováno měření (anti)imigračních postojů v rámci mezinárodního šetření ESS a současně je provedena analýza ekvivalence měření vybraných položek, které jsou již od roku 2002 velmi

často interpretovány a adoptovány v rámci sekundárních analýz jiných výzkumů napříč celým světem.

Analytická část diplomové práce je rozdělena na dvě hlavní části. V rámci první popisuji projekt ESS především s ohledem na jeho metodologii a konkrétní otázky věnující se tématu imigrace. V rámci této části je také popsán odhad kvality měření vybraných otázek měřících postoje k imigraci a imigrantům. Druhou částí analytické části je již samotná analýza. V rámci té je, jak již bylo řečeno, otestována srovnatelnost (ekvivalence) na úrovni měření a popsán vývoj (anti)imigračních postojů v čase.

8.1. European Social Survey

Zdroj dat, respektive jedna z jeho škál, který bude podroben bližší analýze s ohledem na jeho metodologii při dosahování kvality šetření je **European Social Survey** (dále ESS). Jedná se o rozsáhlé mezinárodní kontinuální šetření v oblasti sociálních véd, které probíhá již od roku 2001. Tento výzkumný projekt se zabývá studiem vzorců a proměn chování či postojů obyvatel všech dotazovaných (evropských) států. Výzkumná šetření probíhají pravidelně v dvouletých intervalech a jsou organizována centrálním týmem „Core Scientific Team" (CST). Tato skupina vědců řízená ředitelem Rory Fitzgeraldem, ve spojení s národním koordinátorem, usiluje o co nejvyšší možnou kvalitu dat. Jelikož je ESS projektem pod akademickou záštitou, splňuje veškeré metodologické standardy. Zakládá si na transparentnosti, apeluje na standardizaci a staví na dodržování přesných postupů.

Téma imigrace, respektive modul věnující se tomuto tématu, nepatří mezi ta, která se opakují v každé vlně šetření. Téma se objevilo poprvé v rámci první a dále až sedmé, zatím předposlední, vlny šetření. Česká republika se účastnila obou těchto vln. Existují však otázky, které mapují postoje k imigraci a opakují se s každým kolem šetření. Analyzovaná škála obsahuje otázky právě tohoto charakteru.

8.2. Metodologie ESS

Šetření ESS, respektive dotazník, který je v jednotlivých zemích předkládán, lze rozdělit na dvě části. První z nich můžeme označit jako hlavní část a druhou pojmenujme jako část rotující. Hlavní část dotazníku obsahuje proměnné a otázky měřící v oblastech nejrůznějších témat a mapuje také sociodemografické pozadí respondentů. Jedná se o otázky, které jsou každé dva roky pokládány respondentům, a jejich měřením je docíleno jednoho z nejzákladnějších

cílů ESS, kterým je mezinárodní komparace v čase. Rotující části jsou pak přirozeně ty, které byly k hlavní sekci přiřazeny relativně nahodile, jelikož neexistuje žádná pravidelnost v jejich výskytu v rámci jednotlivých kol šetření. Mezi takové patří i modul imigrace. Samozřejmě i v rámci druhého zmíněného příkladu dotazníku je nejvyšším cílem dosáhnout srovnatelnosti napříč státy, proto je v průběhu designu výzkumu (především dotazníku) použito jak kvalitativních, tak kvantitativních přístupů, které o srovnatelnost usilují (viz. kapitola o dosahování ekvivalence). Design výzkumu trvá celkem 20 měsíců. Mezi nejvýznamnější metody v rámci jednotlivých přístupů patří využívání kódování otázek s využitím SQP pro zjištění kvality otázek, kognitivní rozhovory, znalecké posudky a diskuze, komplexní překlad či jednotlivé techniky kvantitativního testování. Od roku 2008, kdy byla provedena 4. vlna šetření, je dokumentován důkladně celý tento proces s výsledným zněním otázek, kterému předcházely všechny uvedené změny a kde jsou prezentovány diskuze ohledně jednotlivých změn spolu s autory připomínek a argumentů (question design template).

8.2.1. Statistický odhad kvality měření v rámci analýz ESS

Nejvyšší možné kvality dat v rámci šetření se v ESS dosahuje prostřednictvím odhadů kvality měření, komparability daných měřících nástrojů a posouzením sociodemografického složení vzorku respondentů s ohledem na vnější nezávislá kritéria. V předchozí podkapitole bylo pojednáno o tvorbě otázek a jejich měření, které následně probíhá, a jehož kvalita je testována. Samotné měření přirozeně úzce s touto tvorbou a jednotlivými kroky či výslednými rozhodnutími v rámci designu dotazníku souvisí. Celá tato problematika kvality měření je detailněji popsána v teoretické části této práce. Zde se přímo zaměřuji pouze na techniky, kterými se kvality měření dosahuje v ESS a následně uvádím dílčí výsledky statistického testování otázek měřících (anti)imigrační postoje, které odhaduje kvalitu měření skrze MTMM design.

Nejzásadnějším prvkem je bez pochyb, jako zásadní proces při odhadování kvality otázek (měření) v ESS, MTMM design. Jednotlivé experimenty, které byly provedeny v rámci ESS, a souvisí s otázkami z modulu imigrace, jsou uvedeny v samostatné (následující) kapitole, kde jsou výsledky také interpretovány. Zmíním zde pouze, že každé z dosud publikovaných testování mělo podobu dvou-skupinového Split Ballot MTMM experimentu, pouze v 8. vlně šetření byl proveden experiment třískupinový. Ten nicméně v této práci není zohledňován, jelikož otázky z modulu imigrace v rámci něj testovány nebyly. Druhým významným komponentem při dosahování kvality měření v prostředí ESS je nástroj SQP, který se s MTMM

experimenty úzce pojí, jelikož je založen na informacích získaných z těchto experimentů prostřednictvím meta analýzy, jak bylo vysvětleno již v samostatné kapitole věnující se MTMM designu a SQP.

8.2.2. Odhad kvality měření otázek z modulu imigrace v rámci všech kol šetření ESS skrze MTMM experimenty

Dotazování v rámci první vlny probíhalo od konce listopadu 2002 do první poloviny března roku 2003. Zaměříme-li se na kvalitu měření, můžeme v rámci dokumentace metodologie ESS pozorovat jaké kroky a především techniky byly aplikovány pro dosažení co nekvalitnějších měřících nástrojů, dat a současně dat, která budou srovnatelná napříč státy. Významný díl pozornosti v rámci kvality měření spadá na MTMM experimenty a analýzu jejich výsledků. Nebylo možné testovat všechny otázky hlavní části dotazníku najednou (především z důvodů časových a přílišného zatížení respondenta). Připomínám, že odhady kvality jsou ovlivněny také faktory, jako je pozice otázek v dotazníku, délkou otázek a dalšími. Pro odhady kvality měření se zohledněním všech aspektů, včetně oněch zmíněných výše, musela být provedena meta-analýza MTMM experimentů. Výsledné odhady pak byly založeny na více, než 1000 otázkách z šetření.

Testované otázky v rámci témat testovaných MTMM experimenty v první vlně šetření nezahrnovaly výslovně otázky na postoje k imigrantům ani imigraci obecně, nicméně testování v rámci první vlny je zde dobré uvést z důvodu prvotních výstupů a obecných doporučení pro další testování v budoucnu (týkající se všech instrumentů). Všechna testovaná témata odkazují na velmi dobrou kvalitu otázek, ale výsledky se znatelně liší s ohledem na jednotlivá témata a zemi, ve které dotazování proběhlo. Jinak řečeno, byla prokázána kvalita dat jako velmi dobrá, ale měla by být studována více napříč jednotlivými sekcemi (zeměmi, potažmo jazyky v daných zemích). Ačkoli tedy nebyly v rámci první vlny testovány otázky na postoje k imigrantům, analytici došli k závěrům pojícím se k vhodnosti použití dané metody, které jsou aplikovatelné v následujících vlnách šetření. Základními pravidly, kterými by se měli metodologové v rámci budoucí tvorby a úpravy měřících nástrojů řídit, jsou následující:

1. Použít co nejvíce kategorií, pokud je to smysluplné. V případě volby sedmi kategorií by měly být uvedeny fixní body
2. U výroků uvedených v bateriích by se mělo vyhýbat formátu odpovědí „souhlasí/nesouhlasí“

3. Použití karet by mělo být přítomno za jakýchkoli podmínek
4. Možnost „nevím“ by měla být nabídnuta u otázek v hlavní části dotazníku
5. Pokud je to možné, měly by být použity bipolární škály a použití unipolárních by mělo být minimální

Je třeba uvést, že odhad kvality pro otázky v tomto kole byl pouze pro otázky v angličtině, holandštině a němčině. Proto je nezbytné pro další výzkum získat hodnoty pro další jazyky, aby bylo možné významné mezikulturní srovnání.

Poprvé byla v rámci MTMM testování odhadována kvalita otázek se zaměřením na imigraci v roce 2006 v rámci 3. kola ESS. Je zřejmé, že při dotazování na imigraci je výrazně vhodnější využití tzv. „trait specific“ otázek, než použití škál „souhlasí/nesouhlasí“ a to i přes malý počet kategorií. „Trait specific“ nebo také „item specific“ otázky (dále TS nebo IS) se konkrétně ptají, jak by respondenti ohodnotili konkrétně určitý stav (například, jaké je jejich zdraví – odpovědi jsou: výborné, velmi dobré, dobré, …), naopak škály míry souhlasu umožňují vyjádřit míru souhlasu s položkou spojenou s konkrétní situací (například, do jaké míry souhlasí či nesouhlasí s tím, že je jejich zdraví výborné). Toto zjištění je poměrně zajímavé vzhledem k testování v rámci první vlny, která prosazovala využití co největšího počtu kategorií. Stejná zjištění se objevují pro otázky v oblasti následků, které přináší nárůst imigrace. Na základě tohoto experimentu můžeme pozorovat, že co se počtu kategorií týče, nárůst počtu kategorií automaticky neznamená lepší měření a lepší data. Mnohem důležitější se ukazuje použití vhodného typu škály, kterým je v tomto TS. Tento typ škály se projevuje jako vhodnější pravděpodobně z důvodu usnadnění vnímání kategorií a respondenti proto nemusí vyvinout takové úsilí při jejich rozlišování.

	Main Questionnaire	Scales used in the supplementary questionnaires of the different groups		
	all	group 1	group2	group 3
Questions				
Immigration				
B35-B37	4 TS	5 A/D	4 TS	7 D/A
Consequences				
B38-B40	11 TS	5 A/D	11 D/A	7 D/A
Satisf. work				
E26 –E28	5 A/D	5 A/D	11 TS	11 D/A
Satisfy contacts				
E40, E43,E45	5 A/D	5 A/D	5 A/D	7 D/A

TS = trait specific question A/D = Agree /disagree scale D/A = Disagree/agree scale

In this design each group gets 12 extra questions and we can evaluate:
1 the quality of the measure of immigration
2 the quality of the measure of the consequences of more immigration
3 The effect of the number of answer categories
4. the quality of the trait specific scales can be compared with the quality of different agree/disagree scales
5. the effect of the similarity of scales after 2 sets and 3 sets of the same scales.
6. the cross cultural comparability of the 11, 7 and 5 point scale can be compared

Tabulka č. 1: škály použité v hlavním a alternativních dotaznících (zdroj: ESS, MTMM test 3. vlna šetření)

	all	group 1	group2	group 3
Questions				
Immigration				
B35-B37	.75 (4 TS)	.59 (5 A/D)	.91 (4 TS)	.59 (7 D/A)
Consequences				
B38-B40	.68 (11 TS)	.51 (5 A/D)	.32 (11 D/A)	.36 (7 D/A)
Satisfaction with work				
E26 –E28	.56 (5 A/D)	.75 (5 A/D)	.65 (11 TS)	.62 (11 D/A)
Satisfaction with contacts				
E40, E43,E45	.65 (5 A/D)	.61 (5 A/D)	.56 (5 A/D)	.45 (7 D/A)

TS = trait specific question A/D = Agree /disagree scale D/A = Disagree/agree scale

Tabulka č. 2: hodnoty kvality vybraných škál hlavního a doplňkových dotazníků (zdroj: ESS, MTMM test 3. vlna šetření)

Tabulka č. 3 ilustruje, že žádná z použitých a testovaných škál nevykazuje známky větší vhodnosti v rámci daných států, respektive lepší srovnatelnost napříč kulturami. Pouze sedmibodová škála má hodnoty kvality nízké, a tedy naznačuje nevhodnost jejího použití, jak bylo naznačeno již výše (nicméně data pro Českou republiku v tom případě nebyla zahrnuta).

Country	quality
Spain	.78
Bulgaria	.72
Germany	.71
UK	.69
Portugal	.68
Estonia	.67
Denmark	.67
Poland	.66
Belgium	.65
Switzerland	.64
France	.61
Slovenia	.60
Slowakia	.59
Russia	.58

Tabulka č. 3: kvalita škály v dané zemi (zdroj: ESS, MTMM test 3. vlna šetření)

Otázky na postoje k imigraci byly dále testovány v rámci 6. kola šetření ESS v roce 2012. Opět v tabulce vidíme, které z proměnných byly testovány. Jedná se o tři výroky z baterie mapující postoje k imigraci v odlišných dimenzích, které byly měřeny třemi různými metodami (11 -ti bodová IS škála, 7 -mi bodová IS škála a 5 -ti bodová IS škála).

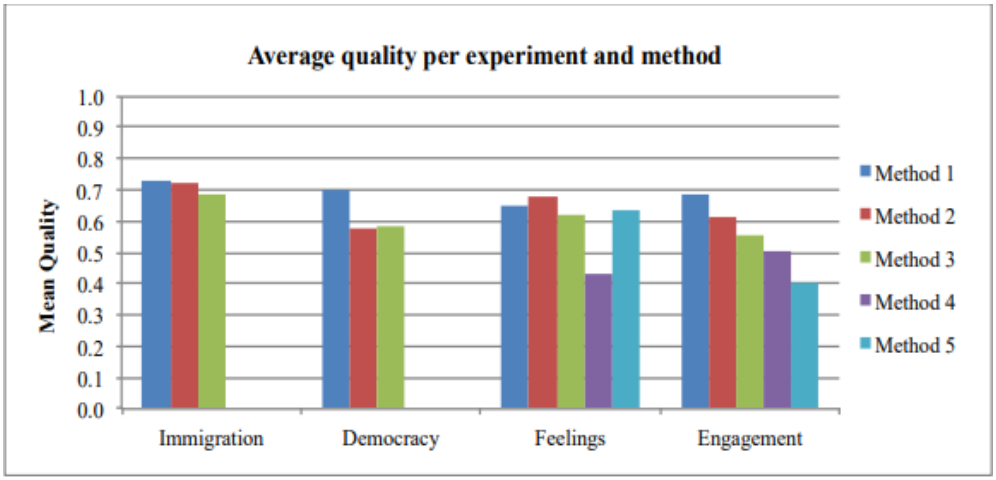
Experiment	ID	Trait	Wording of the questions
Immigration	B32	Economy	- Bad or good for [country]'s economy that people come to live here from other countries
	B33	Culture	- [Country]'s cultural life is undermined or enriched by people coming to live here from other countries
	B34	Place	- [Country]'s made a worse or a better place to live by people coming to live here from other countries

Tabulka č. 4: testované proměnné v rámci SB-MTMM testu 6. vlny šetření (zdroj: ESS)

I v tomto případě byl proveden dvou-skupinový design experimentu, kde každá ze skupin respondentů odpovídala na hlavní část dotazníku, která byla pro obě skupiny v metodě shodná, a následně na další část dotazníku, která se metodou lišila pro každou skupinu. Jednotlivé skupiny jsou ilustrovány následující tabulkou.

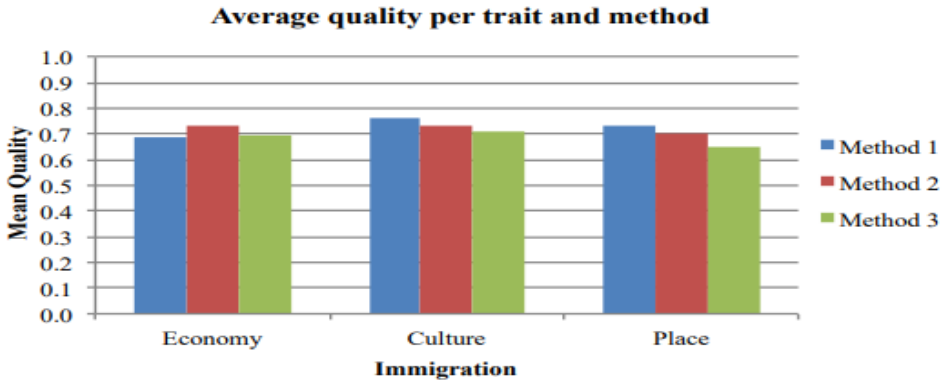
	Hlavní dotazník	Doplňkový dotazník
Skupina 1	Metoda 1	Metoda 2
Skupina 2	Metoda 1	Metoda 3

V prostředí níže uvedeného grafu je možné pozorovat agregovaná data (nikoli rozdíly mezi znaky a mezi zeměmi). Jak bylo již řečeno, v rámci tohoto experimentu jsou pozorovány rozdíly mezi škálami s rozdílným počtem kategorií v případě IS škál. Výsledky v rámci modulu imigrace nejsou markantní, avšak lze pozorovat nepatrně převažující kvalitu metody číslo 1.



Graf č. 1: průměr kvality otázek s ohledem na použitou metodu (zdroj: ESS, MTMM test 6. vlny šetření)

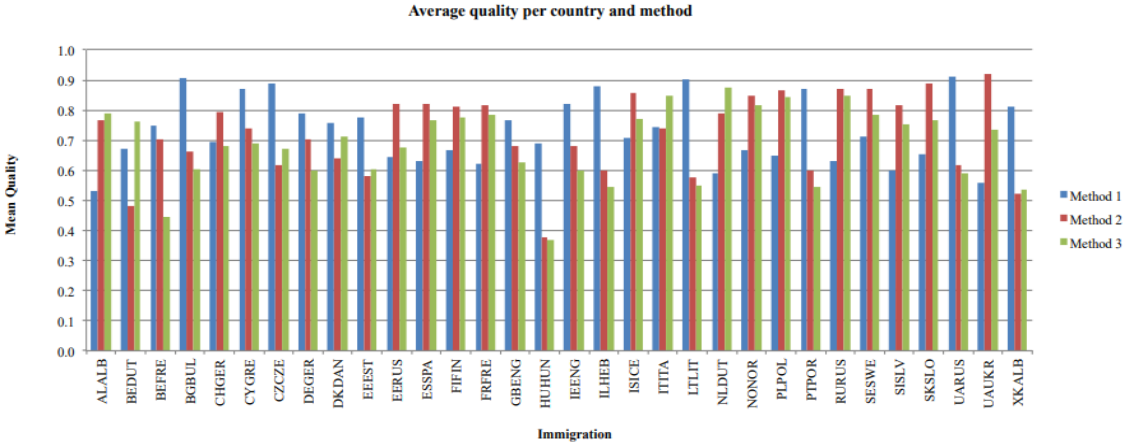
Na základě grafu č. 2 můžeme konstatovat, že pro znaky *kultura* a *místo* se průměrná kvalita měření snižuje se snižujícím se počtem kategorií a obecně lze říci, že škály s menším počtem kategorií mají menší průměrné hodnoty kvality. V případě výroku s *ekonomickým zaměřením*, má největší hodnotu kvality škála s kategoriemi sedmi. Nicméně rozdíly mezi jednotlivými metodami jsou velmi malé.



Graf č. 2: průměrná kvalita otázek pro daný znak a použitou metodu (zdroj: ESS, MTMM experiment 6. vlny šetření)

Zjištění, která jsou uvedena výše, sice odporují původním pravidlům vycházejícím z prvního testování skrze MTMM design, ale naopak podporují zjištění vycházející z druhého zde zmíněného testování postojových otázek.

Poslední graf prezentuje rozdíly v měření jednotlivých metod napříč státy. Je zřejmé, že mezi metodami v rámci jednotlivých států existuje značná nejednota. Pro shrnutí lze na základě výstupů konstatovat, že je vhodnější využití škály s více kategoriemi.



Graf č. 3: kvalita otázek pro jednotlivé země se zohledněním metod (zdroj: ESS, MTMM experiment 6. vlny šetření)

Další v řadě a zatím poslední vlna šetření, která se věnovala v rámci testování kvality měření otázkám měřícím postoje k imigraci, byla vlna sedmá (probíhající v roce 2014).

Testování využitím SB MTMM experimentů bylo opět realizováno v podobě dvou-skupinového rozdělení respondentů jako tomu bylo v předchozím případě. Tabulka č. 5 ilustruje otázky z modulu imigrace, které byly testovány v tomto případě.

Experiment	Trait[2]	ID[3]	Variable[4]	Questions' request for an answer
Immigration	Ability to speak language	D2 IF1 IF10	qfimlng testf1 testf10	How important do you think being able to speak the country's language should be in deciding whether someone born, brought up and living outside should be able to come and live here.
	To be white	D4 IF2 IF11	qfimwht testf2 testf11	How important you think being white should be in deciding whether someone born, brought up and living outside should be able to come and live here?
	Commitment way of life	D6 IF3 IF12	qfimcmt testf3 testf12	How important you think being committed to the way of life should be in deciding whether someone born, brought up and living outside should be able to come and live here?

Tabulka č. 5: testované položky zahrnuté v hlavním a doplňkových dotaznících v 7. vlně šetření (zdroj: ESS, MTMM experiment 7. vlny šetření)

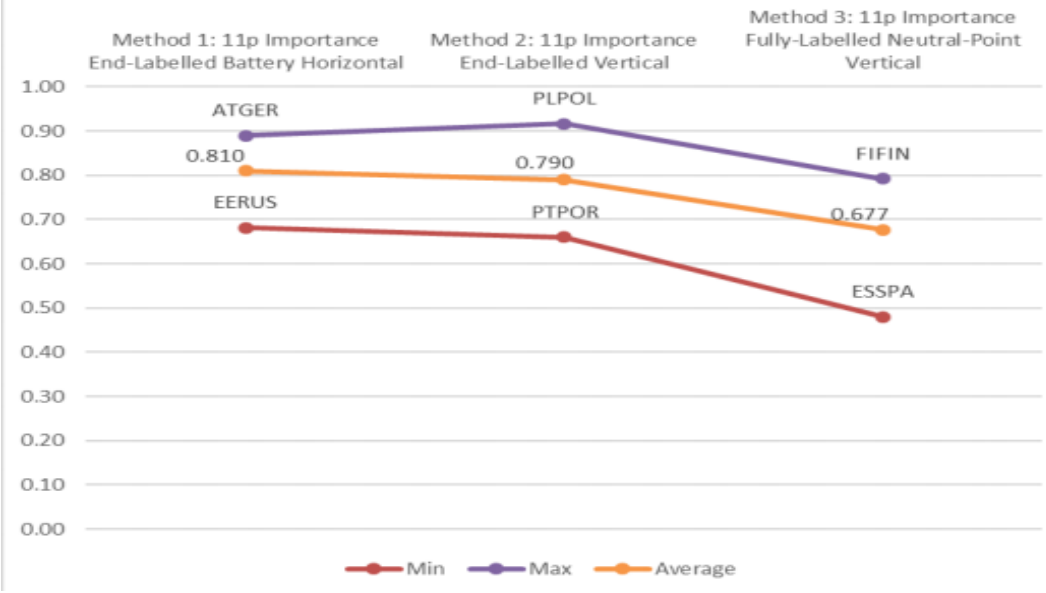
První metoda měření, která je společná všem respondentům je v podobě 11 -ti bodové škály, která má horizontální podobu a pouze krajní body vyznačené. Druhá metoda použitá pro část respondentů se liší ve vertikálním rozložení kategorií. Třetí metoda je pak také svým rozložením bodů vertikálního charakteru, nicméně veškeré body jsou zde vyznačeny svými hodnotami. Tyto jednotlivé metody jsou pro jasnější představu naznačeny v tabulce č. 6.

Experiment	Method 1	Method 2	Method 3
	D2, D4, D6	*IF1, IF2, IF3*	*IF10, IF11, IF12*
Immigration	Extremely unimportant 0 … Extremely important 10	0 Extremely unimportant … 10 Extremely important	0 Extremely unimportant 1 Very unimportant 2 Quite unimportant 3 Rather unimportant 4 A bit unimportant 5 Neither unimportant nor important 6 A bit important 7 Rather important 8 Quite important 9 Very important 10 Extremely important

Tabulka č. 6: formulace škál vybraných položek (zdroj: ESS, MTMM experiment 7. vlny šetření)

Výsledky, které jsou zanesené do grafu č. 4, odkazují na nižší průměrnou kvalitu měření metody 3, zatímco mezi metodou 1 a 2 není významný rozdíl. Tento výsledek podporuje teorii, i přirozený rozum, jelikož metoda 3 představuje metodu, která je se svými kategoriemi nejnáročnější z kognitivního hlediska odpovídání respondenta. Současně můžeme v prostředí

tohoto grafu pozorovat hodnoty kvality měření dané metody v konkrétních vybraných zemích, respektive maxima a minima těchto hodnot. Přesněji se jedná o Rakousko s nejvyšší hodnotou kvality měření prostřednictvím metody 1 a nejnižší hodnotou v případě rusky mluvících respondentů v Estonsku.



Graf č. 4: průměrná kvalita položek napříč zeměmi se zohledněním metody (zdroj: ESS, MTMM experiment 7. vlny šetření)

V dílčích grafech sestavených pro jednotlivé znaky můžeme vidět kromě stability průměrných hodnot kvality prvního a druhého znaku a variability třetího, jak se liší hodnoty napříč zeměmi, což jen potvrzuje důležitost braní zřetele na specifická prostředí daných zemí a jejich výsledky. Znak „schopnost mluvit daným jazykem“ ukazuje největší rozdíly mezi minimem a maximem hodnot kvality měření v případě metody 3. Nejmenší rozdíl mezi těmito hodnotami pozorujeme v případě měření stejného znaku, ale metodou první. To značí, že je pravděpodobné, že metoda 3 je náchylnější k odlišným interpretacím.

Obrázek č. 3: soubor grafů značících kvalitu otázky napříč jednotlivými znaky a zeměmi (zdroj: ESS, MTMM experiment 7. vlny šetření)

9. Statistické testování ekvivalence v SPSS a AMOS

V rámci teoretického ukotvení této práce, bylo řečeno, že pro odhad ekvivalence na úrovni měření, respektive její dosažení, je nezbytné, aby byla naplněna ekvivalence teoretická, což zjednodušeně řečeno znamená, že při dotazování budou všichni respondenti chápat koncepty stejným způsobem a bude zohledněno, zda došlo ke změně vnímání konceptů v čase či ne. Tato část práce, již čistě analytická, se věnuje ekvivalenci na úrovni měření. Konkrétně je zkoumáno, zda jsou vybrané otázky (potažmo získaná data), které měří (anti)imigrační postoje v ESS, srovnatelné mezi jednotlivými státy. Toto statistické testování ekvivalence měření je uskutečněno, jelikož prostředí České republiky se například od prostředí Belgie či států Severní Evropy, ať už svými postoji či historií spojenou s tímto tématem, výrazně liší a tvoří tak prostředí značně specifické. Výsledky testování odhalí, zda daný měřící nástroj dosahuje vyšší kvality v některých z na výzkumu participujících zemích a v jiných menší, či zda jsou ekvivalentní. Současně je možné v rámci analýzy pozorovat změnu a vývoj postojů respondentů jednotlivých zemí v čase. Stěžejními jsou však především výsledky pojící se s Českou republikou, kvalitou měření v rámci ní a postoji jejích obyvatel.

Analýza je provedena v prostředí statisticko-analytického softwaru IBM SPSS (Statistical Package for Social Sciences) a současně, a převážně, v programu AMOS, který je komerčním produktem SPSS. Název je zkratkou výrazu „Analysis of MOment Structure“ a jedná se o program, který poskytuje grafické prostředí, kde lze modelovat kauzální vztahy prostřednictvím SEM.

Konkrétně jsou v rámci analýzy zahrnuty tři proměnné/otázky.

1. To what extent do you think [country] should allow people of the same race or ethnic group as most [country] people to come and live here? (**imsmetn**)
2. How about people of a different race or ethnic group from most [country] people? (**imdfetn**)
3. How about people from the poorer countries outside Europe? (**impcntr**)

Jedná se o položky, které jsou obsaženy v hlavní části dotazníku šetření ESS a jsou proto dotazovány v rámci každé vlny šetření. Každá z těchto položek by měla měřit postoje k imigraci. První a druhá položka se týkají míry, do jaké by dle respondenta měla daná země povolovat lidem stejného nebo jiného etnického původu či rasy vstoupit do země a žít v ní. Třetí otázka/položka zjišťuje, do jaké míry má být tento přístup a usazení v zemi povolen lidem z chudších mimoevropských zemí. Když se podíváme na škálu měření těchto položek, uvidíme, že jsou proměnné měřeny na čtyřstupňové stupnici bez přítomnosti střední hodnoty (od 1 = povolit vstup všem až po 4 = nepovolit vstup žádnému). Tyto proměnné by na základě skórů měly vykazovat míru (anti)imigračního postoje dotazovaných respondentů.

Pro analýzu jsou použita data z prvního (2002), čtvrtého (2008) a osmého (2016) kola šetření. Jsou selektovány odpovědi od respondentů, respektive zemí, které se účastnily všech těchto tří vln šetření. Tato volba měření s šestiletým a osmiletým odstupem by měla být schopna zhodnotit, zda došlo k významným změnám v postojích s ohledem na události, které se od roku 2002 udály. Země, které jsou zahrnuty do analýzy, jsou: Belgie, Švýcarsko, Německo, Česká Republika, Estonsko, Finsko, Francie, Velká Británie, Maďarsko, Izrael, Nizozemsko, Norsko, Polsko, Portugalsko, Slovinsko, Irsko, Švédsko.

Nejzásadnější ze statistických procedur, které jsou aplikovány, je faktorová analýza. Nejdříve je provedena faktorová analýza explorační, která zodpovídá otázku, zda vybrané koncepty měří skutečně jednu dimenzi. Následně je přistoupeno k faktorové analýze konfirmační, respektive mnohoskupinové konfirmační analýze (zkratka MGCFA z anglického překladu Multi Group Confirmatory Factor Analysis), aby mohla být analyzována již samotná srovnatelnost měření na všech třech úrovních (viz teoretická část - kapitola o ekvivalenci

a testování ekvivalence). Celému tomuto procesu však předchází explorační část mapující postoje respondentů české republiky a jejich vývoj od roku 2002 do roku 2016.

9.1. (Anti)imigrační postoje obyvatel ČR

Změříme-li se na postoje respondentů České republiky, které jsou měřeny ve třech časových období, lze pozorovat, že přestože je znatelný evidentní trend ve stálosti převažujícího postoje, docházelo také k určitým změnám. Veškeré informace uvedené v této podkapitole vycházejí z výstupů uvedených v příloze č. 1 – 6. Zmíněný trend, který představuje poměrnou stálost v postojích, spočívá v inklinaci respondentů odpovídat prostřednictvím dvou prostředních hodnot (jinými slovy neextrémními hodnotami). V roce 2002 bylo celkové procentuální zastoupení prostředních možností v podobě 83,6 % pro proměnou Imsmetn, 79,3 % pro proměnnou Imdfetn a 79,8 % v případě proměnné Impcntr, přičemž rozdělení odpovědí mezi tyto dvě možnosti je poměrně vyvážené. Rok 2008 s procenty 78,9 %, 73,5 % a 73,2 % značil s ohledem na poměrové rozložení všech odpovědí podobnou situaci, avšak lze pozorovat, že se číslo snižuje. V roce 2016 poté pozoruje zcela totožný jev s ještě větším poklesem procentuálního zastoupení prostředních možností odpovědí (73,7 %, 62 %, 66 %). Tento pokles se přirozeně projevil na nárůstu v jiné oblasti odpovědí, ale jak lze pozorovat ne ve všech zbylých odpovědích. Zaměřím se proto nyní konkrétněji na změny, ke kterým docházelo postupně v průběhu jednotlivých vln šetření.

V případě položky Imsmetn, která se ptá, do jaké míry by měl stát povolovat vstup a žití v daném státě mnoha lidem stejného etnika a rasy, můžeme vidět, že od roku 2002 do roku 2016 dochází k poklesu a část respondentů prosazujících tento vstup se zmenšuje (v roce 2002 by povolilo vstup mnoha imigrantům popsané podoby 9,3 %, v roce 2008 8,9 % a v roce 2016 5,1 %). Analogicky množství respondentů zastávajících názor, že by do země neměli být pouštěni žádní imigranti stejného etnika a rasy se zvětšuje (v roce 2002 to bylo 7,1 % respondentů, v roce 2008 12,2 % a v roce 2016 21,2 %). Položka Imdfetn, zjišťující obdobně názor respondentů na povolení země vstupovat imigrantům do České republiky a žít zde, konkrétně řeší imigranty jiného etnika a jiné rasy. V tomto případě docházelo k ještě většímu poklesu přívětivosti respondentů ke vstupu mnoha imigrantů, který se propadl až na 1 %. Zde je odmítavý postoj ještě o něco znatelnější, než tomu bylo v případě předchozí položky. Stejně tak se poměrně výrazně zvyšovalo v průběhu let dotazování procento respondentů považujících za vhodné nepovolit vstup žádným imigrantům jiné rasy a jiného etnika, které se zvyšovalo z původních 14 % v roce 2002, přes 21 % v roce 2008 až na 37 % v roce 2016. V případě postojů

respondentů k povolení vstupu imigrantů z chudších mimoevropských zemí je také evidentní pokles přívětivosti v případě povolení vstupu mnoha takovýmto imigrantům (rok 2002 představoval 7 % takových respondentů, v roce 2008 to bylo 5 % a v roce 2016 2 %). Naopak zakázat vstup všem imigrantům z chudších mimoevropských zemí, by v roce 2002 podpořilo 13 % respondentů, v roce 2008 21,5 % a v roce 2016 celých 32,2 %. Je tedy evidentní, že odmítavý postoj ke vstupu imigrantů z chudších mimoevropských zemí a jejich usazení se v této zemi, se zvyšuje.

Když si pro porovnání prohlédneme data z prostředí Belgie, lze pozorovat výstupy, které jsou zásadně odlišné, respektive vesměs opačné. Tabulky ukazují, že od roku 2002 se procento respondentů, kteří by povolili vstup imigrantům stejného etnika či rasy, zvýšilo z původních 14 % na 26 % v roce 2016. Naopak procento těch kteří by této skupině imigrantů zakázali vstup kleslo z 8, 6 % v roce 2002 na necelých 5 % v roce 2016. Stejně tak přívětivost vůči příchozím odlišného etnika a rasy se v tamním prostředí zvýšila. Zatímco v roce 2002 by povolilo mnoha takovýmto příchozím vstup 8 %, v roce 2016 by to povolilo 17 %. Analogicky míra respondentů, kteří by všem příchozím jiného etnika a rasy zakázali vstup klesla ze 14, 5 % v roce 2002 na 9,4 % v roce 2016. A konečně postoj respondentů k imigrantům z mimoevropských chudších zemí. I v tomto případě je belgické obyvatelstvo přívětivější, jelikož z původních 9, 6 % respondentů v roce 2002, kteří by povolili vstup mnoha imigrantů z popsaného prostředí, vzrostla hodnota na 18 % v roce 2016. A současně z 13 % představujících respondenty, kteří v roce 2002 uvedli, že by zakázali vstup všem imigrantům z chudších mimoevropských zemí, se stalo procent 10 v roce 2016.

Na základě analýzy dat a výše popsaného vývoje postojů obyvatel ČR, které jsme porovnali s daty z belgického prostředí, lze potvrdit, že názory a obecně (anti)imigrační postoje jsou skutečně napříč jednotlivými státy nekonzistentní. Jak již bylo mnohokrát v této práci zmíněno, nebude zde věnována větší pozornost faktorům, které tyto postoje mohou ovlivňovat a řešení toho problému. Naopak vzhledem k rozličným výsledkům a vzhledem k tomu, že se jedná o mezinárodní dotazování, bude věnována pozornost samotnému měření a tomu, zda je ekvivalentní, jinými slovy, zda jsou získaná data porovnatelná a nejsou výsledné rozdíly ovlivněné spíše charakterem daného nástroje a jeho vývojem, než skutečnými odlišnostmi populací.

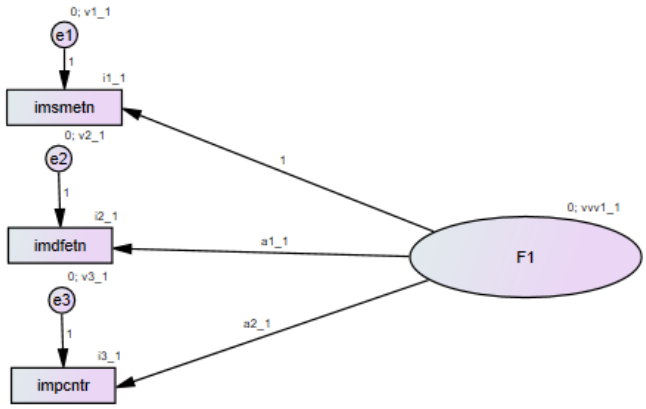
9.2. Explorační faktorová analýza

Po prvotním zhlédnutí dat, pro informaci o rozložení hodnot v datech a představu o daných postojích respondentů, lze říci, že větším problémem jsou pro respondenty imigranti, kteří jsou jiného etnika nebo pochází z chudých zemí ležících mimo Evropu. Nyní přistoupím již k samotné analýze ekvivalence měření, která je provedena prostřednictvím více skupinové konfirmační faktorové analýzy. Nicméně ještě před tím, než bude tento samotný krok proveden a následně interpretován, je třeba provést faktorovou analýzu explorační. Ta je provedena, jelikož, jak vyplývá i z teoretického ukotvení této práce o dosahování vyšší reliability měření v podobě škály, je potřeba otestovat, zda jednotlivé položky skutečně měří daný konstrukt, zda jsou dostatečně silně korelovány.

Postupně lze z výstupů faktorové analýzy (příloha č. 7) vyčíst, že faktorová analýza je vhodným nástrojem a je vhodné ji použít (hodnota KMO je vyšší, než 0,5 a nulovou hypotézu o tom, že proměnné nejsou statisticky významně provázané a FA tudíž nemá smysl, můžeme na základě nulové hodnoty signifikance zamítnout). Stejně tak v matici Anti-image se přesvědčujeme, že není třeba žádnou z proměnných vyřadit, jelikož hodnoty jsou dostatečně vysoké. V tabulce komunalit vidíme, že část proměnné, kterou lze vysvětlit daným faktorem je také poměrně velká. Následující výstupy včetně „hokejkového“ grafu již demonstrují doporučený počet faktorů, který nás utvrzuje v tom, že je pouze jeden, tudíž vybrané položky skutečně měří jeden konstrukt a lze s nimi pracovat jako se škálou.

9.3. Více-skupinová konfirmační faktorová analýza

Samotná analýza ekvivalence měření je provedena prostřednictvím více skupinové konfirmační faktorové analýzy v prostředí programu AMOS umožňujícího strukturální modelování. Na obrázku níže je možné vidět schéma ilustrující konstrukt a jednotlivé položky, které ho měří. Před samotným popisem a interpretací výstupů je dobré objasnit, že v případě testování ekvivalence měření, lze hovořit o ekvivalenci dílčí (partial) a celkové (full). Celková ekvivalence měření značí, že parametry pro všechny položky v modelu měření jsou si rovny napříč skupinami [Byrne 1989]. Já v této práci analyzuji (anti)imigrační postoje respondentů 17 zemí ve tří časových okamžicích, což znamená porovnání průměrů latentní proměnné u 51 skupin. Z tohoto důvodu (porovnání) musí být otestováno, zda škála vykazuje znaky skalární dílčí ekvivalence, což znamená, že alespoň dvě ze tří položek, faktorové zátěže jsou rovnocenné napříč skupinami.



Obrázek č. 4: schéma ilustrující konstrukt a jednotlivé položky, které ho měří (program AMOS)

Jinými slovy lze tento problém vysvětlit s pomocí teoretického ukotvení uvedeného výše, které pojednává o třech úrovních ekvivalence měření. Nejdříve je testován model, který má stejnou faktorovou strukturu napříč skupinami (ale žádná omezení/fixování parametrů), to znamená, že faktorové zátěže a konstanty se mohou lišit napříč zeměmi (konfigurální ekvivalence). Druhý model obsahuje fixaci faktorových zátěží tak, aby byly stejné ve všech skupinách (metrická ekvivalence). A konečně třetí model má kromě zátěží stejné také konstanty napříč skupinami (skalární ekvivalence). Výsledná ekvivalence měření pak může být na základě vhodnosti těchto modelů odhadována.

Nyní se již zaměřme na vhodnost jednotlivých modelů provedené analýzy. Prvním krokem posuzování této vhodnosti je ukazatel/tabulka CMIN, která obsahuje hodnoty chi-kvadrát testů a počet stupňů volnosti jednotlivých modelů. „Unconstrained model“, který odpovídá konfigurální ekvivalenci má hodnotu něco málo přes 1 a 0 stupňů volnosti. Model „measurement weights“, který odpovídá ekvivalenci metrické má hodnotu chi-kvadrát testu 2025 na 100 stupňů volnosti, a model „measurement intercept“, který značí ekvivalenci skalární, má hodnotu chi-kvadrát testu 22 058 na 250 stupňů volnosti. Čím větší hodnota chi-kvadrát testu je, tím je větší nesoulad mezi modelem a pozorovanými daty. Hodnota signifikance je ve všech případech nulová, což indikuje značnou nepravděpodobnost, že model pozorovaná data vhodně popisuje. Jinak řečeno, jak metrická, tak skalární ekvivalence je zde na základě těchto hodnot zamítnuta. Nicméně, než se unáhlím s negativními závěry, je třeba vzít v potaz, že chi-kvadrát testy jsou známy svou citlivostí na velikost zkoumaných vzorků a

porušení normality rozložení, které způsobují vychýlení hodnot a následné odmítnutí celého modelu [(Bollen, 1989; Tucker & Lewis, 1973, Chen 2007].

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Unconstrained	459	1,013	0		
Measurement weights	359	2025,136	100	,000	20,251
Measurement intercepts	209	22058,372	250	,000	88,233
Saturated model	459	,000	0		
Independence model	306	151283,716	153	,000	988,782

Tabulka č. 7: hodnoty chi-kvadrát testu a stupně volnosti (výstup z programu AMOS)

Alternativou pro posouzení vhodnosti modelu jsou hodnoty Root Mean Squared Error of Approximation (RMSEA) and Comparative Fit Index (CFI) viz teoretický rámec práce. V případě CFI značí přijatelné modely hodnoty větší, než 0,9. To je v případě našich dat splněno u konfigurální a metrické ekvivalence, ale nikoli u skalární.

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Unconstrained	1,000		1,000		1,000
Measurement weights	,987	,980	,987	,981	,987
Measurement intercepts	,854	,911	,856	,912	,856
Saturated model	1,000		1,000		1,000
Independence model	,000	,000	,000	,000	,000

Tabulka č. 8: hodnoty informačního kritéria CFI (výstup z programu AMOS)

Analogicky hodnoty RMSEA, které jsou menší, než 0,05 by měly indikovat vhodný model. Tohoto požadavku je zde docíleno v případě jak ekvivalence metrické, tak v případě ekvivalence skalární.

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Measurement weights	,015	,015	,016	1,000
Measurement intercepts	,033	,032	,033	1,000
Independence model	,110	,110	,111	,000

Tabulka č. 9: hodnoty informačního kritéria CFI (výstup z programu AMOS)

Na základě uvedených informací nelze jednoznačně konstatovat, že je škála porovnatelná napříč státy a jednotlivými vlnami měření. Naopak na základě posouzení jednotlivých ukazatelů a jejich hodnot, lze o porovnatelnosti pochybovat. Nicméně jak bylo uvedeno výše, celková ekvivalence vyžaduje, aby alespoň dvě ze tří položek splňovaly požadované podmínky. V případě práce s tak velkým vzorkem jako je ten mnou analyzovaný (81 620) jsou však často hodnoty a posouzení vhodnosti modelu zásadně ovlivněny. Celkový model může být přijat (či odmítnut), ale jeho výsledná (ne)vhodnost může být způsobena výraznější nepřesností v rámci jedné či více skupin či parametrů. Z tohoto důvodu není dobré se spoléhat pouze na výše zmíněná a interpretovaná kritéria. Pro identifikování takových nepřesností jsem se rozhodla zaměřit na modifikační indexy. Ty indikují, jak moc by hodnota chi-kvadrát testu klesla, pokud by parametr nebyl fixovaný (to znamená, jak by se vylepšil model). Vysoké hodnoty v těchto případech odkazují na nevhodnost vázání parametrů. Nicméně, jak již bylo řečeno výše, hodnoty chi-kvadrát testu jsou citlivé na velké zkoumané vzorky, tudíž jako skutečně významné rušivé elementy by měly být brány ty skupiny a parametry, u nichž se vyskytují hodnoty vyšší, než 20 a současně výsledná změna parametru je větší, než 0,10 (v absolutní hodnotě). Jedná se o doporučení, které bylo aplikováno v rámci testování ekvivalence v prostředí ESS.

Skupina a proměnná	M.I.	Parameter change
CH8 IMSMETN ← F1	20,8	-0,129
DE4 IMSMETN ← F1	37,1	-0,109
DE8 IMSMETN ← F1	95,5	-0,198
DE8 IMDFETN ← F1	35,3	0,104

ES1 IMSMTN ← F1	50,8	0,107
FI8 IMSMETN ← F1	25,9	-0,108
IL4 IMSMETN ←F1	133,2	-0,709
IL4 IMDFETN ← F1	33,2	0,234
IL8 IMSMETN ← F1	278,5	-0,778
NO4 IMSMETN ←F1	95,6	-0,269
PT4 IMSMETN ←F1	168,9	0,129
PT4 IMDFETN ←F1	142,13	-0,106

V tabulce jsou zaznamenány modifikační indexy větší, než 20, kde je zároveň změna parametru větší, než 0,10. Interpretace této tabulky je následující. Pro ilustraci si vezměme například poslední uvedenou skupinu PT4, která odpovídá měření v Portugalsku v rámci 4. vlny šetření. Modifikační index 142,13 značí, že pokud nebude rovnoměrné faktorové zatížení, tak se hodnota chi-kvadrát testu sníží o 142,13 jednotek a faktorová zátěž bude o 0,106 nižší, než v současném modelu. Konkrétně nám tato informace říká, že postoje k imigrantům stejné etnické skupiny (imsmetn) v portugalské populaci více souvisí s (anti)imigračními postoji, než je tomu u jiných skupin. Nicméně po prohlédnutí tabulky je zřejmé, že většina změn souvisí právě s jednou konkrétní proměnnou (imsmetn). Tato skutečnost indikuje, že faktorové zátěže pro tuto položku se mezi skupinami poměrně významně liší, zatímco zbylé dvě položky jsou v tomto ohledu mnohem stabilnější. Modifikační indexy ve výstupech značí, že analyzovaná škála by podle všech důkazů nemusela být zcela porovnatelná napříč státy. Ekvivalence je zásadně ovlivňována okolnostmi jedné ze tří položek, konkrétně imsmetn. Můžeme tedy říci, že tato položka, respektive její obsah hovořící o postoji respondentů k imigrantům ze stejné etnické skupiny, má jiný význam pro respondenty z různých států.

10. Diskuze a závěr

Vzhledem k zásadní aktuálnosti tématu migrace a především imigrace, respektive imigrantů, a krize či problémů s tímto fenoménem spojovaných, je stále častěji využíváno dat, která tuto situaci mapují. Takových dat existuje velké množství a to ať už bereme v potaz data prvotní či analyzování a práci s již existujícími daty. Velká část výzkumů, ne-li převážná, či alespoň dílčí část této většiny, se zaměřuje na konkrétní postoje. Na to, jaké jsou a jak respondenti daný fenomén vnímají. To znamená, zda je postoj respondentů k imigrantům přívětivý či odmítavý, jaké aspekty migrace jim vadí, apod. Další skupina výzkumů týkajících se tématu migrace se zabývá sociálními charakteristikami respondentů a snaží se odhalit potenciální korelace mezi těmito znaky respondenta a jeho postojem k imigraci. Jinými slovy, zkoumají to, co ovlivňuje respondentovo vnímání imigrace a imigrantů na mikroskopické úrovni. V souvislosti s těmito výzkumy se množství dalších studií zaměřuje s obdobným přístupem na faktory ovlivňující postoje respondenta, avšak na úrovni makroskopické jako je například vyspělost dané země, ve které dotazování probíhá aj. Poslední skupinou studií, která analyzuje toto téma a kterou lze zařadit do sféry sociálních věd, je analýza postojů různých politických stran k fenoménu migrace, potažmo imigrace.

Je nepochybné, že výzkumy v každé této oblasti jsou velmi přínosné a důležité, o to více jsou-li studovány v jejich kombinaci a vědomí všech oblastí a provázanosti, která mezi nimi existuje. Současně se jedná o téma, které je celosvětově rozšířeno a tak je také zkoumáno. Každým okamžikem přibývají studie či diskuze řešící toto téma v perspektivě jednotlivých prolínajících se sfér lidských životů či v perspektivě mezinárodní. A proto, jelikož je toto téma frekventovaným předmětem mezinárodních studií, lze zřetelně pozorovat, že se jedná o fenomén, ke kterému respondenti napříč skupinami (především napříč státy a generacemi) zastávají odlišné postoje. Z důvodů jako je výše zmíněná komplexnost, rozdílnost v postojích napříč skupinami a obecně známé rozličnosti světových kultur či historických vývojů jednotlivých států, je velmi důležité k výzkumům zmíněným v prvním odstavci zařadit také výzkumy, které se zabývají metodologií. Považuji za důležité, aby nebyla metodologie vytržena z kontextu celkového problému a byla vyvíjena v souvislosti s teoretickými znalostmi o zkoumaném tématu. Toto metodologické zaměření ve spojení s problematikou (anti)imigračních postojů je obsahem mé diplomové práce, jejíž přínos spočívá v jejím zaměření na dosahování kvality měření zkoumaného fenoménu a na to, co je možné dělat pro

to, aby byla získaná data co nejlepší, nejkvalitnější a především v kontextu mezinárodních šetření, aby byla měření, respektive získaná data, srovnatelná.

Srovnatelnost měření a posouzení, zda je daný měřící nástroj vůbec použitelný v konkrétní zkoumané skupině respondentů je velmi zásadním bodem v otázce kvality dat/měření, kterému však není věnována taková pozornost jako jiným (psychometrickým) ukazatelům kvality, v prostředí České republiky, a zde publikovaných studií, je pozornost tomuto tématu minimální. Přitom považuji za více, než významné tyto výzkumy provádět. Imigrace a imigranti jsou předmětem diskuzí na denním pořádku. Jsou každodenně využívána data o pohybu obyvatelstva a především o tom, jak se k tomuto pohybu obyvatelstvo staví. Ale jak bylo již několikrát zmíněno, mimo jiné i České prostředí je svým vnímáním a postoji k imigrantům specifické. Stejně jako ostatní postkomunistické země a především jako jedna ze zemí, kde je tento jev nejznatelnější, se populace ČR staví k imigraci, imigrantům a otázkám s nimi spojenými velmi negativně, zatímco například v Belgii či severských zemích Evropy jsou postoje mnohem přívětivější. Nicméně jedná-li se o data z mezinárodních kontinuálních šetření, která jsou stěžejním a nejčastějším zdrojem dat o těchto postojích, naskýtá se mnoho otázek, která souvisí s důležitostí právě srovnatelnosti daného měření. Ačkoli sesbíraná data ilustrují, jaké postoje respondentů jsou, co je možným faktorem, který má na tyto postoje vliv a další podobné informace, je významnou otázkou, zda si všichni respondenti napříč dotazovanými státy vyložili otázku stejně, zda je konotace zvolených slov a frází všude stejná, zda nemají respondenti tendenci se uchylovat ke konkrétním hodnotám škály, zda vnímají délku a typ škály stejným způsobem atp. Z těchto důvodů tato práce volí přístup výzkumu, který zkoumá jeden z nejvýznamnějších zdrojů dat o (anti)imigračních postojích, zkoumá, jaké byly podniknuty kroky k dosažení srovnatelného měření a dat a především se snaží přispět svou analýzou aplikovatelosti měření ve specifickém prostředí České republiky.

V této práci byla srovnatelnost konkrétního měření hodnocena prostřednictvím více skupinové konfirmační faktorové analýzy, která je pravděpodobně nejvyužívanější metodou pro odhad ekvivalence měření. Alternativou této metodě může být metoda teorie odpovědi na položku (zkratka IRT z anglického překladu Item Response Theory). Shodnost těchto metod spočívá v jejich ohledu na funkci měnných a neměnných položek, to znamená, že všechny položky, ať už se napříč skupinami mění či ne, jsou užitečné v reprezentaci individuálních vlivů na latentní proměnnou (v rámci každé skupiny) [Reise, Widaman, Pugh 1993]. Co se však rozdílů mezi těmito přístupy týče, objektivně lze říci, že s ohledem na specifikaci modelu pro

naši datovou matici se procedury konfirmační faktorové analýzy zdají být více zdokonalené, nekomplikované a především uživatelsky přívětivé, zatímco ty procedury, které jsou vyvinuty pro přístup IRT jsou spíše opakem [Reise, Widaman, Pugh 1993]. Současně je dobré zmínit, že konfirmační faktorovou analýzu lze provést v mnoha statistických a modelujících programech jako jsou AMOS, LISREL, EQS, SAS, které jsou všechny schopny tuto proceduru provést a jsou podobně uživatelsky přívětivé. V případě IRT je nutné pracovat se speciálním programem Multilog [Thissen 1991]. Rozdíl můžeme pozorovat také v indexech vhodnosti modelu. Pro IRT modely je jedinou standardní mírou vhodnosti hodnota chi-kvadrát testu, odhadovaná způsobem obdobným tomu v případě statistických měr vhodnosti u konfirmační faktorové analýzy, nicméně praktické indexy vhodnosti, které byly interpretovány v této práci (RMSEA, CFI, atp.), pro modely IRT nebyly řádně vyvinuty [Reise, Widaman, Pugh 1993].

Výsledky analýzy, ke kterým jsem se dopracovala v této práci já, uvedenými metodami, potvrdily poměrně významný rozdíl ve vnímání imigrace, respektive rozdíl v postojích k imigrantům panující v Evropě a jeho vývoj v letech 2002 až 2016. Ve srovnání s ostatními zeměmi (především západního a severního bloku, ale i množství dalších zemí) je ČR jednou s nejsilnějšími antiimigračními postoji. Co je důležité, je, že české obyvatelstvo se od roku 2002 stává vůči imigrantům stále odmítavější. Na otázku zda jsou výsledná data kvalitní, generalizovatelná a skutečně je měření ekvivalentní (jelikož se jedná o mezinárodní šetření) jsem získala odpověď analýzou ekvivalence měření v programu AMOS prostřednictvím více skupinové konfirmační faktorové analýzy. Po zohlednění všech zásadních informací pro interpretaci výstupů se měření projevilo jako ekvivalentní. Pouze jedna položka se zdála být po prohlédnutí modifikačních indexů jako mírně problematická, jelikož se v jejích případě vyskytovaly problematické hodnoty u indexů a její hodnoty ovlivňovaly nejvíce skalární ekvivalenci. Konkrétně se jednalo o položku Imsmetn, která zjišťuje postoje respondentů k imigrantům stejného etnika či rasy, respektive jejich náklonnost k tomu, aby k nám přicházeli a usadili se zde. Zmíněný jev by naznačoval nejednotné vnímání této položky respondenty napříč jednotlivými státy. Podobné testování ekvivalence měření provedl tým ESS. Jejich výsledky analýzy jsou vesměs shodující s těmi mými [Meuleman 2009]. Možným následujícím řešením je provést znovu zhodnocení modelu, avšak s úpravou omezení/fixování v případě dané proměnné, která podle výsledků způsobovala problémy (imsmetn). To znamená, že by byla testována dílčí ekvivalence, kde by si konstanty nebyly rovné. Pouze u dvou položek by si rovné byly, třetí lišící by však bylo nutné zvolit jinou, než právě položku imsmetn, která by neměla být použita jako „marker item“. Tuto prohlubující analýzu provedli Meuleman s kolegy

na datech velmi se podobajících datům mým, proto čtenáře se zájmem o průběh této procedury odkazuji právě na ně.

Výsledky analýzy srovnatelnosti měření konkrétní škály, měřící (anti)imigrační postoje, v mé práci se shodují s výsledky podobného výzkumu. Přestože je provedená analýza spíše jakousi základnou pro hlubší výzkum a následné řešení je pouze navrženo, shledávám její přínos v ověření dané srovnatelnosti škály měření, v zaměření aplikovatelnosti daného nástroje v českém prostředí či v samotném důrazu, který práce klade na pozornost, která by měla být věnována kvalitě dat, především s ohledem na srovnatelnost, která je mezi českými výzkumy studována ve velmi malém rozsahu, přitom data o pohybu obyvatelstva a především postojích obyvatel k příchozím, jsou využívána stále častěji. Je proto důležité se řádně zabývat kvalitou daných dat, v perspektivě mezinárodních šetření pak především ekvivalencí konkrétního měření.

11. Literatura

ÆGISDÓTTIR, Stefanía; GERSTEIN, Lawrence H.; ÇINARBAŞ, Deniz Canel. Methodological issues in cross-cultural counseling research: Equivalence, bias, and translations. The counseling psychologist, 2008, 36.2: 188-219.

ALBA, Richard; NEE, Victor. Rethinking assimilation theory for a new era of immigration. International migration review, 1997, 31.4: 826-874.

ALMOND, Gabriel Abraham; VERBA, Sidney. The civic culture: Political attitudes and democracy in five nations. Princeton university press, 2015.

ALWIN, Duane F. Margins of error: A study of reliability in survey measurement. John Wiley & Sons, 2007.

ANDERSON, Christopher J. Economics, politics, and foreigners: Populist party support in Denmark and Norway. Electoral Studies, 1996, 15.4: 497-511.

ANDREWS, Frank M. Construct validity and error components of survey measures: A structural modeling approach. Public opinion quarterly, 1984, 48.2: 409-442.

ANÝŽOVÁ, Petra. Mezinárodní srovnatelnost postojových škál sociálního kapitálu a politického odcizení v Evropském sociálním výzkumu. 2014.

BABBIE, Earl R. The basics of social research. Cengage learning, 2013.

BERKA, KAREL. K pojetí měření v československé sociologii. Sociologický Časopis/Czech Sociological Review, 1971, 545-558.

BOLLEN, Kenneth A. A new incremental fit index for general structural equation models. Sociological Methods & Research, 1989, 17.3: 303-316.

Bollen, K. A. 1989. Structural Equations with Latent Variables. New York: Wiley.

BRAUN, Michael; JOHNSON, Timothy P. An illustrative review of techniques for detecting inequivalences. Survey methods in multinational, multiregional, and multicultural contexts, 2010, 373-393.

BYRNE, Barbara M.; CAMPBELL, T. Leanne. Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology*, 1999, 30.5: 555-574.

CAMPBELL, Donald T.; FISKE, Donald W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological bulletin, 1959, 56.2: 81.

COENDERS, Germà, et al. Stability of three-wave simplex estimates of reliability. Structural Equation Modeling: A Multidisciplinary Journal, 1999, 6.2: 135-157.

Davidov, E., P. Schmidt. 2007. „Values and attitudes towards immigrants: A comparison between Germany and Israel.“ Prezentace, Workshop: Immigration, public opinion and governmental policies, Pompeu Fabra University, Barcelona, March 2007.

DAVIDOV, Eldad, et al. Measurement equivalence in cross-national research. *Annual review of sociology*, 2014, 40: 55-75.

DAVIDSON, Andrew R., et al. Cross-cultural model testing: toward a solution of the etic-emic dilemma. International Journal of Psychology, 1976, 11.1: 1-13.

DEMING, W. Edwards. On errors in surveys. American Sociological Review, 1944, 9.4: 359-369.

DORER, Brita. ESS round 6 translation guidelines. Mannheim: European Social Survey, GESIS, 2012.

ENGEL, Rafael J.; SCHUTT, Russell K. The practice of research in social work. Sage Publications, 2012.

FISHBEIN, Martin; AJZEN, Icek. Belief, attitude, and behavior: An introduction to theory and research. Reading, Mass.: Addison Wessley, 1975.

GROVES, Robert M. Survey costs and survey errors. New York, 1989.

GROVES, Robert M.; LYBERG, Lars. Total survey error: Past, present, and future. Public opinion quarterly, 2010, 74.5: 849-879.

GROVES, Robert M., et al. Survey methodology. John Wiley & Sons, 2011.

GUTMANN, Amy. Multiculturalism. 1994.

HARKNESS, Janet; SCHOUA-GLUSBERG, Alicia. Questionnaires in translation. 1998.

HARKNESS, Janet A., et al. (ed.). Cross-cultural survey methods. Hoboken, NJ: Wiley-Interscience, 2003.

HARRIS, Marvin. History and significance of the emic/etic distinction. Annual review of anthropology, 1976, 5.1: 329-350.

HENDRIKS, A. A., et al. The five-factor personality inventory: cross-cultural generalizability across 13 countries. *European journal of personality*, 2003, 17.5: 347-373.

HOOGHE, Marc, et al. Migration to European countries: A structural explanation of patterns, 1980–2004. International Migration Review, 2008, 42.2: 476-504.

HORN, John L.; MCARDLE, J. Jack. A practical and theoretical guide to measurement invariance in aging research. Experimental aging research, 1992, 18.3: 117-144.

HUI, C. Harry; TRIANDIS, Harry C. Measurement in cross-cultural psychology: A review and comparison of strategies. Journal of cross-cultural psychology, 1985, 16.2: 131-152.

CHEN, Fang Fang; SOUSA, Karen H.; WEST, Stephen G. Teacher's corner: Testing measurement invariance of second-order factor models. *Structural equation modeling*, 2005, 12.3: 471-492.

CHEN, Fang Fang. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 2007, 14.3: 464-504.

CHYLÍKOVÁ, Johana. Měření validity a reliability otázek v šetření European Social Survey a jeho využití. Data & Research: SDA Info/Data a Vyzkum: SDA Info, 2015, 9.1.

CHYLÍKOVÁ, Johana. Analytické metody odhadu chyb měření v datech z výběrových šetření. 2017.

ILLNER, MICHAL. Poznámka k pojetí měření v sociologii. Sociologický Časopis/Czech Sociological Review, 1972, 515-526.

JOHNSON, Timothy P. Approaches to equivalence in cross-cultural and cross-national survey research. 1998.

JÖRESKOG, Karl G. Simultaneous factor analysis in several populations. Psychometrika, 1971, 36.4: 409-426.

JUNG, Hosung. A test for autocorrelation in dynamic panel data models. 2005.

KAPLAN, Abraham. The conduct of inquiry: Methodology for behavioural science. Routledge, 2017.

KROSNICK, Jon A.; ALWIN, Duane F. An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opinion Quarterly, 1987, 51.2: 201-219.

KWAN, Virginia SY; BOND, Michael Harris; SINGELIS, Theodore M. Pancultural explanations for life satisfaction: Adding relationship harmony to self-esteem. *Journal of personality and social psychology*, 1997, 73.5: 1038.

LAZARSFELD, Paul Felix; HENRY, Neil W. Latent structure analysis. Houghton Mifflin Co., 1968.

LAZARSFELD, Paul Felix; MERTON, Robert King (ed.). Continuities in Social Research: Studies in the Scope and Method of" The American Soldier". Free Press, 1950.

LITTLE, Todd D. Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. Multivariate behavioral research, 1997, 32.1: 53-76.

LUBBERS, Marcel; GIJSBERTS, Mérove; SCHEEPERS, Peer. Extreme right-wing voting in Western Europe. European Journal of Political Research, 2002, 41.3: 345-378.MARSH, Roberto M. Comparative sociology: a codification of cross-societal analysis. 1967.

MARSH, Roberto M. Comparative sociology: a codification of cross-societal analysis. 1967.

MATĚJŮ, Petr. Metoda strukturního modelování: Přehled základních problémů. Sociologický Časopis/Czech Sociological Review, 1989, 399-417.

MEULEMAN, Bart; DAVIDOV, Eldad; BILLIET, Jaak. Changing attitudes toward immigration in Europe, 2002–2007: A dynamic group conflict theory approach. *Social science research*, 2009, 38.2: 352-365.

MODOOD, Tariq. Multiculturalism. The Blackwell Encyclopedia of Sociology, 2007, 1-4.

NOVICK, Melvin R. The axioms and principal results of classical test theory. Journal of mathematical psychology, 1966, 3.1: 1-18.

PAYNE, Geoff. Comparative sociology: Some problems of theory and method. The British Journal of Sociology, 1973, 24.1: 13-29.

REISE, Steven P.; WIDAMAN, Keith F.; PUGH, Robin H. Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychological bulletin*, 1993, 114.3: 552.

RHEE, Eun; ULEMAN, James S.; LEE, Hoon Koo. Variations in collectivism and individualism by ingroup and culture: Confirmatory factor analysis. *Journal of Personality and Social Psychology*, 1996, 71.5: 1037.

ŘEHÁK, JAN. Definice měření ve společenských vědách. Sociologický Časopis/Czech Sociological Review, 1971, 638-648.

ŘEHÁK, Jan. „Kvalita dat II. Přístupy ohodnocování výzkumných instrumentů založené na modelování kovariančních struktur.“Sociologický časopis, 1998, 34.2: 195-204.

ŘEHÁK, Jan. „Kvalita dat I. Klasický model měření reliability a jeho praktický aplikační význam.“Sociologický časopis, 1998, 34.1: 51-60.

SARIS, Willem E.; VAN DER VELD, W. M.; GALLHOFER, Irmtraud. A Program for prediction of the Quality of Survey Measurement. In: Paper to be presented in October 2000 at the methodology conference in Köln, Germany. 2000.

SARIS, Willem E.; VAN DER VELD, William; GALLHOFER, Irmtraud. Development and improvement of questionnaires using predictions of reliability and validity. Methods for testing and evaluating survey questionnaires, 2004, 275-297.

SARIS, Willem E.; GALLHOFER, Irmtraud N. Design, evaluation, and analysis of questionnaires for survey research (Vol. 548). Hoboken, NY: Wiley Series in Survey Methodology, 2007.

SCHUMAN, H.; PRESSER, S. The assessment of no opinion. Questions and Answers in Attitude Surveys. 1981.

SMITH, Peter B.; FISCHER, Ronald. Acquiescence, extreme response bias and culture: A multilevel analysis. Multilevel analysis of individuals and cultures, 2008, 285-314.

SMITH, Tom W. In Search of House Effects: A comparison of responses to various questions byfferent survey organizationsTOM. Public Opinion Quarterly, 1978, 42.4: 443-463.

SMITH, Tom W. House effects and the reproducibility of survey measurements: A comparison of the 1980 GSS and the 1980 American national election study. Public Opinion Quarterly, 1982, 46.1: 54-68.

SMITH, Tom W. Developing comparable questions in cross-national surveys. Cross-cultural survey methods, 2003, 69-92.

SMITH, Tom W. Developing and evaluating cross-national survey instruments. Methods for testing and evaluating survey questionnaires, 2004, 431-452.

SMITH, Tom W. Total survey error. 2005.

SMITH, Tom W. Formulating the laws of studying societal change. In: FCSM Research Conference, Arlington, VA. 2007.

SMITH, Tom W. Surveying across nations and cultures. Handbook of survey research, 2010, 2: 733-764.

SMITH, Tom W. Refining the total survey error perspective. International Journal of Public Opinion Research, 2011, 23.4: 464-484.

SMITH, Tom W. 2006-2008 General Social Survey Panel Validation. GSS Methodological Report, 2009.

SPEARMAN, Charles. Correlation calculated from faulty data. British Journal of Psychology, 1904-1920, 1910, 3.3: 271-295.

SPEARMAN, Charles. " General Intelligence," objectively determined and measured.

SPEARMAN, Charles. " General Intelligence," objectively determined and measured. The American Journal of Psychology, 1904, 15.2: 201-292.

STEENKAMP, Jan-Benedict EM; BAUMGARTNER, Hans. Assessing measurement invariance in cross-national consumer research. *Journal of consumer research*, 1998, 25.1: 78-90.

THISSEN, David. *MULTILOG user's guide: Multiple, categorical item analysis and test scoring using item response theory*. Scientific Software International, 1991.

TOURANGEAU, Roger; RIPS, Lance J.; RASINSKI, Kenneth. The psychology of survey response. Cambridge University Press, 2000.TOURANGEAU, Roger; RASINSKI, Kenneth A. Cognitive processes underlying context effects in attitude measurement. Psychological bulletin, 1988, 103.3: 299.

VAN DE SCHOOT, Rens; LUGTIG, Peter; HOX, Joop. A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 2012, 9.4: 486-492.

VAN DE VIJVER, Fons; LEUNG, Kwok. Methods and data analysis of comparative research. Allyn & Bacon, 1997.

VERTOVEC, Steven. Super-diversity and its implications. Ethnic and racial studies, 2007, 30.6: 1024-1054.

VINOPAL, Jiří. Kognitivní přístupy v metodologii výzkumných šetření: metoda okamžité validizace. 2008.

12. Přílohy

Seznam příloh

Příloha č. 1: Frekvenční tabulky proměnných (Imsmetn, Imdfetn, Impcntr) popisující rozložení v rámci první vlny šetření v ČR – výstupy z programu SPSS (tabulka)

Příloha č. 2: Frekvenční tabulky proměnných (Imsmetn, Imdfetn, Impcntr) popisující rozložení v rámci čtvrté vlny šetření v ČR – výstupy z programu SPSS (tabulka)

Příloha č. 3: Frekvenční tabulky proměnných (Imsmetn, Imdfetn, Impcntr) popisující rozložení v rámci osmé vlny šetření v ČR – výstupy z programu SPSS (tabulka)

Příloha č. 4: Frekvenční tabulky proměnných (Imsmetn, Imdfetn, Impcntr) popisující rozložení v rámci první vlny šetření v Belgii – výstupy z programu SPSS (tabulka)

Příloha č. 5: Frekvenční tabulky proměnných (Imsmetn, Imdfetn, Impcntr) popisující rozložení v rámci čtvrté vlny šetření v Belgii – výstupy z programu SPSS (tabulka)

Příloha č. 6: Frekvenční tabulky proměnných (Imsmetn, Imdfetn, Impcntr) popisující rozložení v rámci osmé vlny šetření v Belgii – výstupy z programu SPSS (tabulka)

Příloha č. 7: Výstupy explorační faktorové analýzy provedené v programu SPSS (tabulky)

Příloha č. 8: Teze diplomové práce

Příloha č. 1

imsmetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	108	9,3	9,3	9,3
	2 Allow some	539	46,3	46,3	55,6
	3 Allow a few	434	37,3	37,3	92,9
	4 Allow none	83	7,1	7,1	100,0
	Total	1164	100,0	100,0	

imdfetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	81	7,0	7,0	7,0
	2 Allow some	461	39,6	39,6	46,6
	3 Allow a few	462	39,7	39,7	86,3
	4 Allow none	160	13,7	13,7	100,0
	Total	1164	100,0	100,0	

impcntr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	83	7,1	7,1	7,1
	2 Allow some	501	43,0	43,0	50,2
	3 Allow a few	428	36,8	36,8	86,9
	4 Allow none	152	13,1	13,1	100,0
	Total	1164	100,0	100,0	

Příloha č. 2

imsmetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	161	8,9	8,9	8,9
	2 Allow some	662	36,7	36,7	45,6
	3 Allow a few	763	42,2	42,2	87,8
	4 Allow none	220	12,2	12,2	100,0
	Total	1806	100,0	100,0	

imdfetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	95	5,3	5,3	5,3
	2 Allow some	555	30,7	30,7	36,0
	3 Allow a few	773	42,8	42,8	78,8
	4 Allow none	383	21,2	21,2	100,0
	Total	1806	100,0	100,0	

impcntr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	95	5,3	5,3	5,3
	2 Allow some	533	29,5	29,5	34,8
	3 Allow a few	789	43,7	43,7	78,5
	4 Allow none	389	21,5	21,5	100,0
	Total	1806	100,0	100,0	

Příloha č. 3:

imsmetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	108	5,1	5,1	5,1
	2 Allow some	712	33,4	33,4	38,5
	3 Allow a few	858	40,3	40,3	78,8
	4 Allow none	452	21,2	21,2	100,0
	Total	2130	100,0	100,0	

imdfetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	22	1,0	1,0	1,0
	2 Allow some	435	20,4	20,4	21,5
	3 Allow a few	887	41,6	41,6	63,1
	4 Allow none	786	36,9	36,9	100,0
	Total	2130	100,0	100,0	

impcntr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	39	1,8	1,8	1,8
	2 Allow some	442	20,8	20,8	22,6
	3 Allow a few	963	45,2	45,2	67,8
	4 Allow none	686	32,2	32,2	100,0
	Total	2130	100,0	100,0	

Příloha č. 4

imsmetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	229	14,1	14,1	14,1
	2 Allow some	868	53,5	53,5	67,7
	3 Allow a few	385	23,8	23,8	91,4
	4 Allow none	139	8,6	8,6	100,0
	Total	1621	100,0	100,0	

imdfetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	130	8,0	8,0	8,0
	2 Allow some	754	46,5	46,5	54,5
	3 Allow a few	503	31,0	31,0	85,6
	4 Allow none	234	14,4	14,4	100,0
	Total	1621	100,0	100,0	

impcntr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	156	9,6	9,6	9,6
	2 Allow some	747	46,1	46,1	55,7
	3 Allow a few	508	31,3	31,3	87,0
	4 Allow none	210	13,0	13,0	100,0
	Total	1621	100,0	100,0	

Příloha č. 5

imsmetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	299	19,5	19,5	19,5
	2 Allow some	847	55,3	55,3	74,8
	3 Allow a few	282	18,4	18,4	93,2
	4 Allow none	104	6,8	6,8	100,0
	Total	1532	100,0	100,0	

imdfetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	180	11,7	11,7	11,7
	2 Allow some	710	46,3	46,3	58,1
	3 Allow a few	432	28,2	28,2	86,3
	4 Allow none	210	13,7	13,7	100,0
	Total	1532	100,0	100,0	

impcntr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	198	12,9	12,9	12,9
	2 Allow some	696	45,4	45,4	58,4
	3 Allow a few	446	29,1	29,1	87,5
	4 Allow none	192	12,5	12,5	100,0
	Total	1532	100,0	100,0	

Příloha č. 6

imsmetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	385	26,1	26,1	26,1
	2 Allow some	767	52,0	52,0	78,1
	3 Allow a few	251	17,0	17,0	95,1
	4 Allow none	72	4,9	4,9	100,0
	Total	1475	100,0	100,0	

imdfetn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	247	16,7	16,7	16,7
	2 Allow some	691	46,8	46,8	63,6
	3 Allow a few	398	27,0	27,0	90,6
	4 Allow none	139	9,4	9,4	100,0
	Total	1475	100,0	100,0	

impcntr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Allow many to come and live here	261	17,7	17,7	17,7
	2 Allow some	694	47,1	47,1	64,7
	3 Allow a few	376	25,5	25,5	90,2
	4 Allow none	144	9,8	9,8	100,0
	Total	1475	100,0	100,0	

Příloha č. 7

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,713
Bartlett's Test of Sphericity	Approx. Chi-Square	141980,572
	df	3
	Sig.	,000

Anti-image Matrices

		imsmetn	imdfetn	impcntr
Anti-image Covariance	imsmetn	,491	-,159	-,071
	imdfetn	-,159	,295	-,207
	impcntr	-,071	-,207	,347
Anti-image Correlation	imsmetn	,815[a]	-,417	-,173
	imdfetn	-,417	,658[a]	-,645
	impcntr	-,173	-,645	,701[a]

a. Measures of Sampling Adequacy(MSA)

Communalities

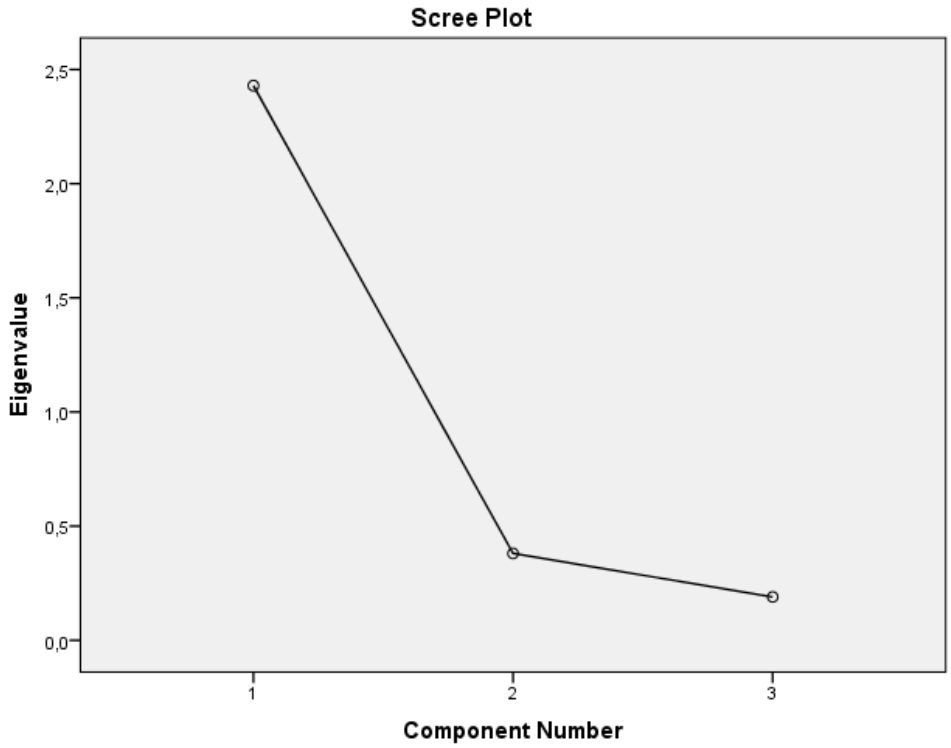
	Initial	Extraction
imsmetn	1,000	,742
imdfetn	1,000	,867
impcntr	1,000	,820

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,429	80,983	80,983	2,429	80,983	80,983
2	,381	12,688	93,671			
3	,190	6,329	100,000			

Extraction Method: Principal Component Analysis.



Příloha č. 8

Předpokládaný název práce:

Koncept chyby měření v kontextu sociálních věd, analytické metody jejího odhadu a jejich aplikace při tvorbě výzkumného nástroje měřícího vnímání strachu a nebezpečí s ohledem na teroristické útoky

Preliminary title of the thesis:

Concept of measurement error in social sciences, analytical methods of its estimation and its application in research tool-creating process measuring the perception of fear and danger with regard to terrorist attacks

Klíčová slova: systémová chyba, chyba měření, reliabilita, validita, kvalita dat, metodologie, výběrové šetření, celková výzkumná chyba

Key words: Bias, measurement error, reliability, validity, data quality, methodology, survey, TSE

Konzultant: Mgr. Jiří Remr, Ph.D.

Úvodní představení tématu

Je předpokladem, že bez ohledu na zaměření, nástroje, použité metody, přístupy a další, je všem výzkumníkům společné jedno a to skutečnost, že se snaží získat data, která jsou kvalitní. Snaha v rámci každého výzkumu, kde pracujeme s daty, docílit co nejvyšší přesnosti dat by měla být na prvním místě. Nicméně ačkoli se možnosti zdokonalující přesnost dat vyvíjejí, mnoho výzkumníků a výzkumů jim stále nevěnuje dostatečnou pozornost a tak riskují značné zkreslení svých výsledků a následných interpretací. Rozhodneme-li se dělat výzkumné šetření, které obsahuje sběr kvantitativních dat, měli bychom již automaticky počítat s tím, že se budeme potýkat s chybami, respektive odchylkami. Některé z nich nemáme šanci ovlivnit, jiné jsou naopak zaviněny například přímo výzkumníky. Ve většině případů výzkumná šetření nezahrnují do dotazování celou populaci, ale pouze její vzorek, a následně je výsledek vztahován na populaci celou. Už jen toto je jeden z důvodů, který vede k odchylkám ve výzkumech. Jedná se o chyby nepozorování, či chybějícího pozorování, jelikož jsme nepozorovali každého jednoho člena celé populace, kterou zkoumáme. Na druhou stranu musíme brát v potaz také chyby, které se vyskytují v našem souboru dat, který jsme napozorovali, tj. chybami pozorování, současně označovanými jako chyby měření. Toto rozlišení chyb je jeden z příkladů, jak lze odchylky v šetření dělit. Rozlišit chyby můžeme ale

také na základě toho, zda je jejich výskyt možné ovlivnit (systémové chyby) či není (variační/náhodné chyby). Součet všech zmíněných odchylek, které se v šetření mohou vyskytnout, spolu s jejich jednotlivými podsložkami, tvoří celek zvaný celková chyba šetření.

Jak již bylo naznačeno, těmto chybám, především pak chybám měření není věnována taková pozornost, jaká by mohla být. Chyby v šetřeních budou vždy, ale je velmi důležité, aby byla tato skutečnost brána v potaz a výzkumníci měli v rámci svých projektů na paměti testy reliability a validity svých nástrojů, především dnes, kdy existuje stále více a více možností, jak testovat používané výzkumné nástroje, jak docílit co nejpřesnějšího měření či jak odhadovat velikost odchylky, která v datech vznikla.

Zaměření mé práce

V mé diplomové práci bych se především chtěla zaměřit na onu problematiku chyb měření. Práce bude obsahovat část teoretickou, která bude ve formě kompilace, ale souběžně bude obsahem práce také část praktická, kde se budu věnovat vlastnímu výzkumu. V teoretické části práce rozeberu podrobně problematiku chyb v rámci výběrových šetření na poli sociálně-vědního výzkumu. Pozornost budu věnovat konkrétním typům chyb, ale především se zaměřím na chyby měření. V souvislosti s touto problematikou následně rozeberu metody odhadování chyby a procesy, kterými lze docílit vyšší přesnosti měření. Současně bych také ráda věnovala alespoň stručnou pasáž problematice rostoucího (?) strachu z cizího, migrantů, násilných či teroristických útoků aj., jelikož právě tato problematika je tématem mého výzkumu, který je obsahem praktické části práce.

V druhé, praktické, části mé práce pak tedy provedu vlastní výzkum. Jeho hlavním cílem bude vytvoření co nejpřesnějšího nástroje měření vnímání pocitu bezpečí či strachu s ohledem na současné události, terorismus, migrační situaci a koncepty s tím spojené. Tento nástroj budu testovat za pomoci vybraných testů, o kterých jsem mluvila v teoretické části a při jeho tvorbě se řídit jednotlivými kroky vedoucími k takovému měřícímu nástroji, který zajistí co nejmenší systémovou chybu. Kromě toho, že zjistím postoje respondentů, budu také testovat sílu question wording či question ordering právě v otázce vnímané bezpečnosti a strachu s ohledem na současné dění ve světě či tuzemsku.

Metodologie

Výzkum, který v rámci mé diplomové práce provedu, bude kvantitativní. Vytvořím takový výzkumný nástroj, který bude měřit, jak lidé vnímají své bezpečí, migraci, světové události, jestli mají strach a jakou roli hrají teroristické útoky, jejich postoj k „cizímu“ apod. Dotazováni budou respondenti v Belgii a České republice, přičemž toto dotazování proběhne samo vyplňovacím dotazníkem. Tyto dotazníky budou ve dvou verzích. Pro dvě verze jsem se rozhodla z důvodu testování síly znění otázek či řazení otázek, avšak obě dvě sady budou obsahovat i totožné otázky. Současně bude tento dotazník rozdělen na dvě části, kde první bude věnována zmíněnému tématu vnímání bezpečí, migrace a podobně, zatímco druhá část dotazníku by obsahovala otázky zjišťující kognitivní procesy respondentů při odpovídání na otázky výzkumu.

Dosavadní poznání v oblasti

Co se dostupné literatury k této diplomové práci týče, není velký problém nalézt dostatečné informace a to k jakékoli dílčí části. Literatura věnující se chybám (nejen) výběrových šetření je skutečně bohatá a stále se rozrůstající. Nalezneme velké množství knih a prací, které se obecně tomuto tématu věnují (De Leeuw, Hox, Dillman 2012; Groves 2004; Weisberg 2009) či se zaměřují konkrétněji například na měření reliability a validity (Alwin 2007; Chylíková 2015) či jiné dílčí problémy, které se s chybami ve výběrovém šetření pojí. Literatury, která se zaměřuje na vnímání bezpečí, imigrantské situace, terorismu apod. je už podstatně méně, jelikož se nejedná o tak snadno uchopitelné pojmy, ale i přesto nalezneme dostačující a stále narůstající množství prací i na toto téma, jelikož je velmi aktuální. Setkáme se například s pracemi, které řeší strach z terorismu jakožto jeden z nejsoučasnějších strachů co se pronásledování týče, kde mnoho z těchto prací odkazuje na skutečnost, že ženy jsou více bázlivé, než muži (Ferraro, 1995; Fisher & Sloan, 2003; Madriz, 1997; Warr, 1984, 2000). V jiných pracích pak nalezneme naopak zaměření na problematiku imigrantů (Paxton, Mughan 2006). Jako hlavní předlohou a cenným zdrojem informací je pro mne práce Handmera a Jamese, kteří věnují pozornost společnosti jakožto měnícímu se prvku, kde se současně mění a tvoří nové druhy risk/nebezpečí. V souvislosti s tím se zabývají i terorismem jakožto formou nebezpečí, “která sdružuje jak ztělesněné praktiky, tak abstraktní procesy - násilné činy lidí, kteří jsou ochotni ohrozit svůj život, aby rozšířili možnosti technologií, které mohou nebo nemusejí být navrženy k zabíjení, ale mohou být zasazeny do mašinerie každodenního života”

[Handmer, James 2007]. Totiž jak nás tito autoři informují, modernismus s sebou přináší více abstraktní a neztělesněná rizika a rizika nově vyvstávající do popředí jsou spíše trvalá, než dočasná, spíše globální než místní, a jsou utvářeny jak prakticky tak i ideologicky [Handmer, James 2007]. Mnoho informací k tomuto tématu "risk society" získáme od Ulricha Becka, Anthonyho Giddense či Mary Douglas.

Předpokládaná použitá literatura:

Alwin, D. F., J. A. Krosnick. (1991). The Reliability of Survey Attitude Measurement. The Influence of Question and Response Attributes. *Sociological Methods and Research* 20 (1), 139 – 181.

Alwin, D. F. (2007). Margins of error: A study of reliability in survey measurement. Hoboken, Wiley

Biemer, P. B., R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, S. Sudman. (2004). *Measurement Errors in Surveys*. New Jersey: Wiley.

De Leeuw, Edith D.; Hox, Joop; Dillman, Don (2012). *International handbook of survey methodology*. Routledge.

Dillman, D. A., L. M. Christian. (2005). Survey Mode as a Source of Instability in Responses across Surveys. *Field Methods* 17 (1), 30 – 52.

Erikson, K. T . (1995). A New Species of Trouble: The Human Experi-ence of Modern Disasters (New York: Norton).

Ferraro, K. (1995). Fear of crime: Interpreting victimization risk. Albany: State University of New York Press

Fisher, B. S., & Sloan, J. J. (2003). Unraveling the fear of victimization among college women: Is the
"shadow of sexual assault" hypothesis supported? *Justice Quarterly, 20*, 633-659

Fischhoff, B. (1995). „Risk Perception and Communication Unplugged: Twenty Years of Progress",Risk Analysis, Vol. 15, No. 2 (1995), pp. 137–145.

Groves, R. M.; Lyberg, L. (2010). Total survey error: Past, present, and future. *Public opinion quarterly*, 74.5: 849-879.

Groves, R. M. (2004). *Survey Error and Survey Costs*. New Jersey: Wiley.

Groves, R. M., F. J. Fowler Jr., M. P. Couper, J. M. Lepkowski, E. Singer, R. Tourangeau. (2004). *Survey Methodology.* Wiley New Jersey

Handmer, John; James, Paul. (2007). Trust us and be scared: The changing nature of contemporary risk. *Global Society*, 21.1: 119-130.

Harkness, J. A., F. J. R. van de Vijver, P. P. Mohler. (2002). *Cross – Cultural Survey Methods*. New Jersey: Wiley

Harzing, A. (2006). Response styles in cross-national survey research: a 26 – country study. *International Journal of Cross Cultural Management* 6 (2), 243 – 266.
Chylíková, J. (2015). Měření validity a reliability otázek v šetření European Social Survey a jeho využití. *Data a výzkum - SDA Info* 9 (1), 5 – 29.

Madriz, E. I. (1997). Images of criminals and victims: A study on women's fear and social control.
Gender & Society, 11, 342-356.

Nellis, A. M. (2009). Gender differences in fear of terrorism. *Journal of Contemporary Criminal Justice*, 25.3: 322-340.

Paxton, P.; Mughan, A. What's to fear from immigrants? Creating an assimilationist threat scale. *Political Psychology*, 2006, 27.4: 549-568.

Saris, W. E., I. Gallhofer. (2007). Estimation of the Effects of Measurement Characteristics on the Quality of Survey Questions. *Survey Research Methods* 1 (1), 29 – 43.

Warr, M. (1994). Public perceptions and reactions to violent offending victimization. In A. J. Reiss, Jr. & J. A. Roth (Eds.), *Understanding and preventing violence: Consequences and control* (Vol. 4, pp.
1-66). Washington, DC: National Research Council.

Warr, M. (2000). Fear of crime in the United States: Avenues for research and policy. Washington, DC: National Institute of Justice

Weisberg, Herbert F (2009). *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press,