

**UNIVERZITA KARLOVA V PRAZE**

**FAKULTA SOCIÁLNÍCH VĚD**

Institut sociologických studií

**Vendula Bláhová**

**Vliv délky škály na měření postojů**

*Diplomová práce*

Praha 2014

Autorka práce: **Bc. Vendula Bláhová**

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

Oponent práce:

Datum obhajoby: **2014**

Hodnocení:

## **Bibliografický záznam**

Bláhová, Vendula. 2014. *Vliv délky škály na měření postojů*. Diplomová práce (Mgr.) Univerzita Karlova, Fakulta sociálních věd, Institut sociologických studií. Katedra sociologie. Vedoucí diplomové práce Ing. Mgr. Jiří Remr, Ph.D., MBA.

### **Abstrakt**

Diplomová práce se věnuje hodnotícím škálám, které jsou široce využívány při měření postojů, konkrétně vlivu jejich délky (počtu bodů) na měření, a to z hlediska kognitivních procesů probíhajících v myslích respondentů i dopadu na získaná data. Hypotézy, formulované na základě studia literatury a testované pomocí dat získaných ze split ballot experimentu a kognitivních rozhovorů, se týkají vzájemné srovnatelnosti výsledků, rozložení odpovědí a náročnosti úlohy. Pro práci je zvolen aplikační příklad z prostředí výzkumu trhu, tzv. Net Promoter Score, vyjadřující míru ochoty doporučit využívanou službu či společnost. Pro ten je originálně využívána jedenáctibodová škála od 0 do 10, která je zde pomocí různých rekódů a přeškálování porovnávána s pětibodovou škálou od 1 do 5. Práce ukazuje, že odpovídání na jedenáctibodové a pětibodové škále znamená pro respondenty dvě odlišné kognitivní úlohy, přičemž odpovídání na delší škále je o něco náročnější, ale ve vnímání respondentů přesnější, a že získaná data se od sebe liší z hlediska rozložení odpovědí – na krátké škále je více zastoupena střední a maximální pozitivní hodnota, zatímco na dlouhé respondenti častěji volí méně silné pozitivní hodnoty.

### **Abstract**

This diploma thesis deals with the rating scales, which are widely used in the attitude measurement, namely with the impact of their length (number of scale points) on measurement, both in terms of cognitive processes in the minds of respondents and impact on the data. Hypotheses, formulated on the basis of the literature and tested

using data obtained from a split ballot experiment and cognitive interviews, are related to the comparability of results, distribution of responses and task demands. The Net Promoter Score from the environment of market research, expressing respondent's willingness to recommend a service or company used, is used as an application example. It originally uses 0 to 10 eleven-point scale which is here compared to five-point scale from 1 to 5 using various recoding and rescaling techniques. Analysis shows that rating on eleven-point and five-point scales means two different cognitive tasks for respondents (eleven-point scale is a bit more demanding but more accurate in respondents perception) and that the data differ from each other in terms of distribution of responses – on short range scale, midpoint and extreme positive category are selected more often, while on eleven-point scale, milder positive points are used more often.

## **Klíčová slova**

Hodnotící škály, počet bodů na škále, měření postojů, split ballot experiment, kognitivní procesy, kognitivní rozhovory

## **Keywords**

Rating scales, number of scale points, attitude measurement, split ballot experiment, cognitive processes, cognitive interviews

## **Prohlášení**

1. Prohlašuji, že jsem předkládanou práci v celkovém rozsahu 146 857 znaků zpracovala samostatně, použila jen uvedené prameny a literaturu a že práce nebyla využita k získání jiného nebo stejného titulu.
2. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne 16. 5. 2014

---

Vendula Bláhová

## **Poděkování**

Na tomto místě bych ráda poděkovala především vedoucímu práce, Ing. Mgr. Jiřímu Remrovi, Ph.D., MBA, za průběžné konzultace zamýšlených řešení, velmi podnětné poznámky a celkové směřování mě k úspěšnému dokončení práce. Dále bych chtěla poděkovat respondentům, kteří mi ochotně věnovali čas a dovolili mi nahlédnout do svých procesů odpovídání při kognitivních rozhovorech. Zároveň patří poděkování mým kolegům a nadřízeným, kteří mi vycházeli časově vstříc v náročném období při psaní práce a kteří mi umožnili realizaci split ballot experimentu, zejména potom Mgr. Radkovi Jalůvkovi za tuto možnost a PhDr. Evě Veisové, Ph.D. za další nápady. V neposlední řadě chci poděkovat svému partnerovi, Bc. Pavlu Sklenaříkovi, za celkovou podporu a jazykovou kontrolu textu.

# PROJEKT DIPLOMOVÉ PRÁCE

**Jméno studující:** Bc. Vendula Bláhová

Institut sociologických studií FSV UK: obor Sociologie, zaměření Aplikovaný sociologický výzkum a jeho metodologie

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

**Předpokládaný název práce v češtině:** Vliv délky škály na kvalitu měření postojů

**Předpokládaný název práce v angličtině:** Effect of number of scale points on quality of attitude measurement

**Klíčová slova v češtině:** měření postojů, délka škály, reliabilita, proces odpovídání, Net Promoter Score (NPS)

**Klíčová slova v angličtině:** attitude measurement, number of scale points, reliability, response process, Net Promoter Score (NPS)

## **Námět práce:**

Ve své práci se zaměřím na metodologii měření postojů, konkrétně na délku užitých odpovědních škál v dotazníkových šetřeních. Přestože jsou škály široce využívané, neexistuje (alespoň prozatím) shoda o tom, jaký počet kategorií nabídnout. Také doporučení na základě provedených experimentů se liší.

Při úvahách o optimální délce škály proti sobě stojí informační teorie, hovořící ve prospěch delších, více rozlišujících škál, a úvahy kognitivních vědců na druhé straně ukazující limity škál s mnoha body. Zastánci informační teorie říkají, že čím více kategorií, tím lépe, protože lze získat více informací. Existují však také praktické limity omezující délku škály – vzájemně rozlišit velký počet kategorií může být pro respondenta velmi nesnadné, kromě toho může také ztratit motivaci a namísto optimalizace odpovědi může dojít k tzv. satisficingu. Otázkou tedy je, zda existuje optimální délka škály nebo alespoň hranice, za kterou již s přibývajícím body škály nedochází ke zvyšování kvality měření.

Při měření postojů je důležité zachytit 1) směr – pozitivní, negativní, 2) intenzitu a 3) neutralitu. Nejkratší možná škála, která všechny tři prvky pokrývá, je pětibodová (střední neutrální kategorie, dvě různě intenzivní pozitivní a negativní kategorie). Ve

své práci se tedy zaměřím na pětibodové a delší škály, a to konkrétně na příkladu doporučováním. Ve výzkumu trhu je aktuálně hojně využíván tzv. Net Promoter Score (NPS), index vyjadřující míru doporučení určitého objektu (například organizace poskytující službu, samotné služby apod.), založený na jedenáctibodové škále od 0 do 10. Pětibodovou škálu proto porovnam s touto jedenáctibodovou a dále s určitým kompromisem mezi oběma délkami – sedmibodovou škálou.

Mými základními výzkumnými otázkami jsou:

- Zvyšuje se reliabilita měření s vyšším počtem kategorií?
- Jak funguje proces odpovídání při různých délkách škál? Jaké jsou strategie respondentů, vzorce odpovídání a jejich schopnost rozlišit mezi jednotlivými body na škále?

### **Předpokládané metody zpracování:**

Ve své práci budu při hledání odpovědí na výzkumné otázky vycházet jak z dostupných teoretických zdrojů, tak z vlastního výzkumu.

Za účelem otestování reliability třech různých délek škál provedu klasický split-ballot experiment, kdy v rámci dotazníkového šetření výběrový soubor náhodně rozdělím do třech pod-skupin, z nichž každá bude odpovídat na otázky pomocí jiné délky škály (5, 7, 11 bodů).

K rozkrytí procesů, strategií a vzorců odpovídání dále provedu sérii polo-strukturovaných rozhovorů.

### **Předpokládaná struktura práce:**

Úvod

#### 1. Teoretická část

- Postoje a jejich měření
  - Doporučováním, NPS
- Délka škály a kvalita dat
  - Různé přístupy
  - Příklady dosud provedených výzkumů



## 2. Metodologie

- Popis výzkumu
  - Split-ballot experiment
  - Polo-strukturované rozhovory

## 3. Analýza dat

Závěr

### **Orientační seznam literatury:**

Alwin, Duane F. 1989. "Problems in the Estimation and Interpretation of the Reliability of Survey Data." *Quality & Quantity* 23: 277 – 331.

Alwin, Duane F. 1992. „Information Transmission in the Survey Interview: Number of Response Categories and the Reliability of Attitude Measurement.“ *Sociological Methodology* 22: 83 – 118.

Alwin, Duane F. 1997. „Feeling Thermometers Versus 7-Point Scales: Which Are Better?“ *Sociological Methods and Research* 25 (3): 318 – 340.

Andrews, Frank M. 1984. „Construct validity and error components of survey measures: A structural modeling approach.“ *Public Opinion Quarterly* 48 (2): 409 – 442.

Dawes, John. 2008. „Do data characteristics change according to the number of scale points used ? An experiment using 5 point, 7 point and 10 point scales.“ *International Journal of Market Research* 50 (1): 1 – 19.

Groves, Robert M. 2004. *Survey Error and Survey Costs*. New Jersey: John Wiley & Sons, Inc.

Groves, Rober M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, Roger Tourangeau. 2009. *Survey Methodology*. New Jersey: John Wiley & Sons, Inc.

Churchill, Gilbert A., Paul J. Peter. 1984. "Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis." *Journal of Marketing Research* 21 (4): 360 - 375.

- Preston, Carolyn C., Andrew M. Colman. (2000). "Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power, and Respondent Preferences." *Acta Psychologica* 104: 1 - 15.
- Reichheld, Frederik F. 2003. „The One Number You Need To Grow.“ *Harvard Business Review*.
- Reichheld, Frederik F. 2006. *The Ultimate Question*. Boston: Harvard Business School Press.
- Revilla, Melanie, Williem E. Saris, Jon A. Krosnick. 2009. "Choosing the Number of Categories in Agree-Disagree Scales." *RECSM Working Paper Number 5*. Barcelona: Universitat Pompeu Fabra - Research and Expertise Centre for Survey Methodology.
- Řehák, Jan. 1998. "Kvalita dat II.: Přístupy ohodnocování výzkumných instrumentů založené na modelování kovariančních struktur." *Sociologický časopis* 34 (2): 195 – 204.
- Tourangeau, Roger. 1984. „Cognitive Sciences and Survey Methods.“ Pp. 73 - 100 in Jabine, Thomas B. et al. (eds.). *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington, DC: National Academy Press.
- Tourangeau, Roger, Lance J. Rips, Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge University Press.
- Weng, Li-Jen. 2004. "Impact of the Number of Response Categories and Anchor Labels on Coefficient Alpha and Test-Retest Reliability." *Educational and Psychological Measurement* 64 (6): 956 – 972.

# Obsah

<b>OBSAH</b> .....	<b>1</b>
<b>ÚVOD</b> .....	<b>3</b>
<b>1. TEORETICKÉ ZAKOTVENÍ PRÁCE</b> .....	<b>4</b>
1.1 <i>Postoje</i> .....	4
1.1.1 <i>Definice postojů</i> .....	5
1.1.2 <i>Měření postojů</i> .....	7
1.1.3 <i>Délka hodnotící škály: teoretické předpoklady</i> .....	9
1.1.4 <i>Délka hodnotící škály: provedené výzkumy</i> .....	10
1.2 <i>Kognitivní model procesu odpovídání ve výzkumných šetřeních</i> .....	13
1.2.1 <i>Porozumění</i> .....	14
1.2.2 <i>Vyvolání informací z paměti</i> .....	17
1.2.3 <i>Úsudek</i> .....	18
1.2.4 <i>Editace odpovědi</i> .....	19
1.2.5 <i>Shrnutí a aplikace modelu</i> .....	21
<b>2. METODIKA VÝZKUMU</b> .....	<b>22</b>
2.1 <i>Kognitivní interview</i> .....	22
2.1.1 <i>Thinking-aloud</i> .....	23
2.1.2 <i>Verbal probing techniques</i> .....	24
2.1.3 <i>Kombinace technik</i> .....	25
2.1.4 <i>Získaná data</i> .....	26
2.1.5 <i>Analýza dat</i> .....	27
2.2 <i>Split ballot experiment</i> .....	28
2.2.1 <i>Výběrový soubor</i> .....	29
2.2.2 <i>Porovnání dat</i> .....	30
2.3 <i>Diskuze použitých metod</i> .....	30
2.3.1 <i>Validita a reliabilita</i> .....	31
2.4 <i>Příklad postoje: Net Promoter Score</i> .....	32
2.5 <i>Research design</i> .....	34
2.5.1 <i>První fáze: split ballot experiment</i> .....	34
2.5.1.1 <i>Cílová populace a výběr</i> .....	35
2.5.1.2 <i>Randomizace</i> .....	36
2.5.1.3 <i>Sběr dat</i> .....	36
2.5.2 <i>Druhá fáze: kognitivní rozhovory</i> .....	36
2.5.2.1 <i>Cílová populace a výběr</i> .....	36
2.5.2.2 <i>Sběr dat</i> .....	37
2.5.2.3 <i>Scénář rozhovoru a další využívané materiály</i> .....	38
2.6 <i>Etické souvislosti výzkumu</i> .....	38
<b>3. ANALÝZA DAT</b> .....	<b>39</b>
3.1 <i>Porozumění</i> .....	40
3.2 <i>Provedení úlohy: vyvolání informací a úsudek</i> .....	41
3.2.1 <i>Množství vyvolaných informací</i> .....	42
3.2.2 <i>Strategie respondentů při provádění úlohy</i> .....	43
3.3 <i>Editace odpovědi</i> .....	44
3.3.1 <i>Vlastnosti proměnných</i> .....	44
3.3.2 <i>Rozložení odpovědi</i> .....	46
3.3.3 <i>Rozložení odpovědi v podskupinách</i> .....	54
3.3.4 <i>Schopnost škál diskriminovat</i> .....	57

3.3.5	<i>Vnímání bodů na škále</i> .....	59
3.3.5.1	<i>Krajní a střední hodnoty</i> .....	62
3.3.6	<i>Náročnost a preference respondentů</i> .....	64
<b>ZÁVĚR</b> .....		<b>65</b>
<b>SUMMARY</b> .....		<b>70</b>
<b>POUŽITÁ LITERATURA</b> .....		<b>73</b>
<b>SEZNAM PŘÍLOH</b> .....		<b>79</b>
<b>PŘÍLOHY</b> .....		<b>80</b>

## Úvod

Tato diplomová práce se zabývá metodikou měření postojů, konkrétně se zaměřuje na jeden z parametrů výzkumného nástroje - délku užitých hodnotících škál. Ty jsou ve výběrových šetřeních široce využívány, mezi výzkumníky ale neexistuje shoda o optimálním počtu bodů na škále. [Alwin 1997: 318] V práci je testován vliv délky škály na měření postojů při aplikaci na konkrétní příklad, tzv. Net Promoter Score (NPS), vyvinutý v roce 2003 Frederikem F. Reichheldem a využívaný v prostředí výzkumu trhu. NPS je jednoduchý ukazatel založený na jedné otázce zjišťující ochotu doporučit určitou službu či jejího poskytovatele přátelům/známým. Tato ochota je měřena na jedenáctibodové škále od 0 (určitě nedoporučil/a) do 10 (určitě doporučil/a). [Reichheld 2003]

Je ale jedenáct bodů skutečně ten nejvhodnější počet? Délka škály má vliv na obtížnost zodpovězení otázky pro respondenta, stejně jako na získaná data: „S příliš málo kategoriemi mohou hodnotící škály selhat v diskriminaci mezi respondenty s odlišnými úsudky; s příliš mnoho kategoriemi mohou selhat respondenti ve spolehlivém rozlišení mezi sousedními kategoriemi.“ [Tourangeau, Rips, Rasinski 2000: 249; překlad autorka] V této práci je jedenáctibodová škála porovnávána s pětibodovou; pět kategorií je nejmenší možný počet pro rozlišení 1) směru - pozitivní nebo negativní, 2) intenzity a 3) neutrality, tedy aspektů nezbytných pro měření postojů. [Alwin 1997: 321] Dlouhá, pro NPS originální jedenáctibodová škála je tedy porovnána s nejkratší možnou škálou zachycující všechny tři uvedené aspekty; to umožňuje zaměřit se pouze na jeden parametr, kterým je právě délka škály (nikoliv například sudý či lichý počet kategorií), a zároveň porovnat velmi kontrastní škály s velkým a malým počtem bodů.

Práce si klade za cíl odhalit, zda a jaký vliv má délka škály na měření, a to jak na kognitivní procesy probíhající v mysli respondenta při hodnocení na škále, tedy na proces získávání dat, tak na samotná data. Odráží se volba délky škály v kognitivní náročnosti úlohy pro respondenty a jejich schopnosti rozlišit jednotlivé body? Jaké strategie respondenti při odpovídání volí? Liší se výsledky výzkumu v závislosti na zvolené délce škály? Pokud ano, jak? Pro zodpovězení těchto otázek bylo nutné zvolit komplexní řešení; práce je tak založena na výzkumu, jehož research design je dvoufázový – split ballot experiment variující délku hodnotící škály byl následován sérií kognitivních rozhovorů nad oběma verzemi výzkumného nástroje. Výzkum se opírá

o dosavadní znalosti a studie provedené v této oblasti; v českém prostředí nebylo dosud této problematice věnováno příliš pozornosti, proto se jedná především o zahraniční poznatky. Práce tak přispívá ke zkoumání dané problematiky v našem prostředí.

Jelikož jsou hodnotící škály významnou součástí měření postojů, je v práci nejdříve představena tato širší oblast – různá pojetí postojů a přístupy k jejich měření. Následně je již pozornost věnována konkrétně hodnotícím škálám a jejich zde zkoumanému aspektu – počtu bodů. Vzhledem k tomu, že pro měření postojů jsou zásadní také kognitivní procesy probíhající v myslích respondentů, je zde představena také tato problematika, konkrétně čtyř-složkový model trojice amerických vědců Rogera Tourangeau, Lance Ripse a Kennetha Rasinskiho. [Tourangeau, Rips, Rasinski 2000] Následuje nutné představení samotného výzkumu, a to jak použitých metod jako takových (split ballot experimentu a kognitivních rozhovorů), tak research designu. Těžiště práce potom leží v její analytické části, kde autorka představuje nejdříve testované hypotézy a následně výsledky výzkumu. Závěrem jednotlivá zjištění propojuje a diskutuje.

## 1. Teoretické zakotvení práce

Výzkum provedený v této práci je zasazen do kontextu měření postojů, které proniká do řady vědních oborů; postoje měří (sociální) psychologové, sociologové, ale i politologové či ekonomové. [Krosnick, Judd, Wittenbrink 2005: 21] Je také běžnou součástí komerčního prostředí výzkumu trhu, kde jsou zjišťovány postoje zákazníků k (využívaným) společnostem, produktům či službám, jako je tomu také v případě doporučování a Net Promoter Scoru, který v této práci posloužil jako konkrétní aplikační příklad.

Ale co vlastně postoje jsou? Jak je v současné době vědci chápou a ke kterému pojetí se přiklání autorka této práce?

### 1.1 Postoje

Od té doby, co Thomas a Znaniecki v roce 1918 v *Polském sedlákovi v Evropě a Americe* definovali sociální psychologii jako studium postojů [Albarracin, Johnson, Zanna 2005: vii] a uvedli, že postoje by měly být základním pilířem sociálních věd,

pronikl pojem velmi rychle i do sociologie. [Vávra 2006: 9] Následující kapitola představuje různá chápání tohoto pojmu.

### 1.1.1 Definice postojů

Vymezení pojmu postoj je stejně jako v případě mnoha jiných termínů nejasné, definic a přístupů existuje celá řada. V prvním desetiletí 21. století bylo publikováno mnoho prací o automatických, implicitních aspektech zpracovávání postojů, což osvěžilo diskuzi o nejlepší konceptualizaci postojů. Většina vědců se shodne alespoň na základních aspektech postojů, tedy že jde o hodnocení určitého objektu, který člověk udržuje v mysli; tímto objektem může být cokoliv všedního i abstraktního - věci, lidé, skupiny, myšlenky. Zastřešující definici představila dvojice Alice H. Eagly a Shelly Chaiken; podle nich je postoj „psychologická tendence, vyjádřená hodnocením konkrétního objektu s určitou mírou přízně či nepřízně.“ [Eagly, Chaiken 2007: 582; překlad autorka] Detailnější definice se ale liší v tom, zda vidí postoje jako stabilní entity uložené v paměti nebo naopak jako dočasné úsudky konstruované v danou chvíli z dostupných informací. [Bohner, Dickel 2011: 392] Příkladem definice postoje konceptualizovaného jako stabilní entita je vymezení Rusella H. Fazio [Fazio 2007], který vidí postoje jako hodnotící asociace objektů uchované v paměti. Reprezentace objektu je dle něj v mysli spojena s jeho globálním hodnocením; pokud tedy osoba přijde do styku s daným objektem, pomocí tohoto asociativního spojení jí přijde na mysl jeho hodnocení. Tato teorie je tedy podobná tradičnějšímu pojetí postojů, které předpokládá, že postoje jsou pre-existující hodnocení určitého objektu, v paměti uložená „složka“, do které stačí respondentovi jen nahlédnout. [Tourangeau, Rips, Rasinski 2000: 166 - 167] Naopak na konstruktivistickém konci spektra definic vidí Norbert Schwarz [Schwarz 2007] postoje ne jako trvalé osobní předpoklady, ale jako hodnotící úsudky zkonstruované v dané situaci (např. dotazovací, pozn. autorky) na základě aktuálně dostupných informací (i když jejich součástí mohou být i v paměti uložená hodnocení).

Pro účely této práce je stěžejní definice postojů dle Rogera Tourangeau, Lance Ripse a Kennetha Rasinskiho, kteří je vidí jako „druh struktury paměti obsahující existující hodnocení, vágní dojmy, obecné hodnoty a relevantní pocity a názory. Při každé příležitosti, když přemýšlíme o nějakém problému, přijde nám na mysl nějaká podmnožina těchto obsahů. V závislosti na tom, které úvahy si vybavíme, a na přesných požadavcích aktuální úlohy, můžeme jednoduše zopakovat již existující hodnocení,

aktualizovat ho nebo jej rozšířit k pokrytí nového aspektu problému; nebo můžeme udělat zcela nový úsudek.“ [Tourangeau, Rips, Rasinski 2000: 194; překlad autorka] Tato definice tedy připouští existenci předem zformovaného hodnocení v hlavě respondenta, zároveň ale pracuje s dalšími obsahy a procesy. Oba konce spektra definic postojů (stabilní entita vs. úsudek zkonstruovaný na místě) mají své silné i slabé stránky, nejužitečnější se tedy zdá být kombinace obou přístupů uvažující stabilní i situačně variabilní aspekty postojů. [Bohner, Dickel 2011: 394] Tourangeau s kolegy přemýšlí o jakémisi kontinuu odpovídajícímu míře zformování respondentova postoje; na jednom konci má respondent předem vytvořený názor pouze čekající na to, aby mohl být nabídnut tazateli/výzkumníkovi; na opačném konci nemá respondent vůbec žádný názor. Mezi těmito dvěma extrémy existuje celá řada situací; respondent může při formování názoru využít volně související set myšlenek nebo může zaujímat určité stanovisko, od kterého se při odpovídání na otázku odrazí. [Tourangeau, Rips, Rasinski 2000: 12]

Postoje bývají spojovány také s chováním či jednáním<sup>1</sup>; touto problematikou se dlouhodobě zabývá například Icek Ajzen. Jeho *teorie plánovaného chování* je modelem vztahu mezi postoji a chováním. Centrálním prvkem modelu jsou lidské intence chovat se určitým způsobem, motivační faktory ovlivňující chování (jak moc je člověk ochoten se snažit, kolik úsilí vynaložit). Obecně platí, že čím silnější je tento záměr, tím pravděpodobnější je i výskyt chování (musí však jít o situaci, kdy má člověk na výběr, zda se takto bude nebo nebude chovat – některé typy chování totiž podléhají skutečné kontrole, jako je dostupnost času, peněz, schopností, spolupráce ostatních atd.). Intence k určitému chování jsou pak ovlivňovány 1) postoji, 2) subjektivními normami a 3) vnímanou kontrolou nad chováním (tato kontrola může přímo ovlivňovat i chování samotné). [Ajzen 1991: 181 - 184] Postoje jsou tedy v tomto komplexním modelu chápány jako jedna z proměnných (nepřímo, skrze intence) ovlivňujících výsledné chování. Nejsou to však samotné postoje, které organizují jednání k hodnoceným objektům, reálné jednání je ovlivněno dalšími proměnnými, ať již osobními (další postoje, sociální schopnosti apod.) nebo sociálními (normy, sociální kontrola apod.). Kromě toho je ve vztahu mezi postoji a jednáním problematická kauzalita – postoje mohou ovlivňovat jednání, mohou být ale také jednáním ovlivněny. [Vávra 2006: 9]

---

<sup>1</sup> Rozlišení pojmů chování a jednání je zde ponecháno stranou, postoje bývají spojovány jak konkrétně s jednáním, tak obecněji s chováním.



### 1.1.2 Měření postojů

K empirickému měření postojů výzkumníci již dlouho využívají tzv. *self-report scales*, kterými respondenta explicitně žádají, aby ohodnotil určitý objekt na hodnotící škále pomocí jedné nebo více položek (otázek, výroků). [Bohner, Dickel 2011: 394] Jeho úkolem je tedy popsat své vlastní postoje. [Krosnick, Judd, Wittenbrink 2005: 31] Jedná se o nejjednodušší přístup, kdy respondentovy odpovědi o tom, co si myslí o určitém objektu, jaký k němu má vztah, považujeme za pravdivé. [Vávra 2006: 9] Za využitím této metody zjišťování postojů stojí předpoklad, že lidé jsou jak ochotni, tak schopni přesně vyjádřit své postoje.<sup>2</sup> Dalším předpokladem je, že postoje existují na kontinuu od pozitivních po negativní; takto bývají chápány i měřeny. [Alwin, Krosnick 1991: 144] Respondentovi je položena otázka týkající se jeho postoje k určitému objektu a jeho úkolem je vybrat odpověď na škále od maximálně pozitivní po maximálně negativní (nebo naopak).

Nejdříve byly k měření postojů jeho průkopníky vyvinuty složité self-report techniky, využívající velkou sadu položek k měření jednoho postoje – technika zdánlivě stejných intervalů Louise Thurstona (1928), technika sumovaných odhadů (Likertova škála) Rensise Likerta (1932) a sémantický diferenciál Charlese Osgooda (1957). Tyto techniky jsou náročné na přípravu (výběr vhodných položek – výroků) a díky počtu hodnocených položek je i odpovídání časově náročné a klade na respondenta velké nároky. Ač jsou tedy tyto i další škálovací techniky využívající více položek k měření jednoho postoje oceňovány 1) pro svou zjevnou validitu, 2) fakt, že celkové skóre je zatíženo menší náhodnou chybou měření než jeden přímý dotaz a 3) že jednotlivé položky jsou vybírány na základě empirických důkazů o blízkosti jejich interpretací napříč různými respondenty, časový tlak typický pro dnešní výzkumy obvykle nedovolí designovat složité škály k měření jediného postoje. Nehledě na to, že pro výzkumníky může být těžké ospravedlnit vynaložení potřebných zdrojů k jejich přípravě i sběru dat. [Krosnick, Judd, Wittenbrink 2005: 32 - 33] Tlak na čas a náklady je velmi patrný v komerčním prostředí výzkumu trhu, ze kterého pochází i NPS – realizátor výzkumu

---

<sup>2</sup> V souvislosti se skepticismem k tomuto předpokladu a za účelem překonání problémů jako je například snaha respondentů prezentovat se pozitivně, byly v posledních desetiletích představeny různé implicitní techniky k měření postojů, například implicitní asociační test [Bohner, Dickel 2011: 394 - 395] nebo pozorování zjevného chování. [Krosnick, Judd, Wittenbrink 2005: 53] Ve vztahu k této práci se však již jedná o jinou oblast metodologie měření postojů, která je proto ponechána stranou.

musí zadavateli obhájit navržené postupy a výhody náročnějších řešení oproti těm méně náročným.

Většina výzkumníků se tak dnes spoléhá na měření postojů pomocí malého počtu otázek, které nebyly vybrány na základě rozsáhlé fáze vývoje a pretestování. Tato praxe tedy klade velké nároky na těch několik málo otázek, aby přinesly maximálně validní a reliabilní hodnocení. Výzkumníci proto musí udělat několik rozhodnutí ohledně struktury a znění otázek i odpovědí. Zvolené postupy se však různí jak mezi třemi výše uvedenými průkopníky, tak případy dnešních výzkumů. [Krosnick, Judd, Wittenbrink 2005: 34]

Důležitým aspektem, o kterém je třeba při měření postojů (ale i behaviorálních či faktických otázek) rozhodnout, je forma odpovědí. V dotazníkových šetřeních existuje těchto forem celá řada; různé formy odpovědí slouží různým účelům. Jedním z prvotních rozhodnutí je, zda bude otázka otevřená (umožňující respondentovi odpovědět v jakékoliv formě si přeje, vlastními slovy) nebo uzavřená, s omezeným počtem předdefinovaných odpovědí. Obecně dávají výzkumníci přednost spíše uzavřeným otázkám – jejich zpracování je jednodušší a navíc usnadňují respondentovi pochopení otázky i určení požadovaného typu odpovědi. [Tourangeau, Rips, Rasinski 2000: 230 - 231] I v případě uzavřených otázek, konkrétně hodnotících (self-report) škál široce využívaných pro měření postojů, je ale třeba rozhodnout o některých aspektech nabízeného setu odpovědí:

- Jaký počet bodů zvolit (délka škály)
- Jak jednotlivé body označit – zda je označit verbálně a pokud ano, zda označit všechny nebo jen některé (například krajní)
- V jakém pořadí hodnoty prezentovat
- Zda nabídnout také možnost „nevím“ [Krosnick, Judd, Wittenbrink 2005: 34]
- Vizuální prezentace škál – vertikální či horizontální, ale i zda je využita nějaká grafika (teploměry, smajlíci apod.)

Každý z těchto aspektů může data určitým způsobem ovlivnit. Aby bylo možné tento vliv sledovat, je třeba se vždy zaměřit pouze na jeden aspekt, aby nedošlo ke kombinaci více vlivů, které by nebylo možné rozlišit. V této práci je pozornost věnována prvnímu aspektu – délce škály. Jsou proto porovnány dvě různé škály, které mají rozdílný počet bodů (11 vs. 5), ale ostatní aspekty mají shodné – obě nabízí střední hodnotu, ale již ne odpověď „nevím“, odpovědi mají seřazené od negativních po

pozitivní, stejně vizuálně prezentované a v obou případech jsou slovně označeny pouze krajní hodnoty (určitě doporučil/a, určitě nedoporučil/a), body mezi nimi jsou označeny jen číselně. Běžně bývají v dotazníkových šetřeních kratší pětibodové škály verbálně olabelovány celé. Některé studie naznačují, že reliabilita měření je vyšší, pokud jsou slovně označeny všechny hodnoty a že respondenti jsou zároveň spokojenější, mohou-li odpovídat na takové škále. [Krosnick, Judd, Wittenbrink 2005: 40] Po vzoru Johna Dawese [Dawes 2008], který prováděl podobný experiment, se však autorka rozhodla u pětibodové škály označit pouze krajní hodnoty (jako je tomu v originální metodologii NPS), aby byl přístup v případě obou škál naprosto stejný.

### 1.1.3 Délka hodnotící škály: teoretické předpoklady

Aby mohl fungovat proces odpovídání na hodnotící škále, kdy respondent nejdříve vnitřně ohodnotí svůj postoj a následně najde bod na hodnotící škále, který mu nejlépe odpovídá (viz kognitivní model dále), musí být splněny některé obecné předpoklady. 1) Body na škále musí pokrývat celé kontinuum, nevynechávat žádnou jeho část. 2) Hodnotící škála musí být ordinální, kdy body postupují od jednoho konce kontinua ke druhému. Význam sousedních bodů se přitom nesmí překrývat. 3) Každý respondent musí významu jednotlivých bodů na škále rozumět; toto porozumění by mělo být relativně přesné a stabilní. 4) Porozumění či interpretace jednotlivých bodů by měla být u většiny nebo nejlépe všech respondentů stejná a musí být známá výzkumníkovi. Pokud některá z těchto podmínek není splněna, trpí tím kvalita získaných dat. Délka škály přitom může ovlivnit každou z nich. [Krosnick, Judd, Wittenbrink 2005: 35 - 37]

V teoretické rovině uvažování o délce škály proti sobě stojí dvě stanoviska. Tradiční *informační teorie* říká, že škály s větším počtem bodů zprostředkují více informací o měřené proměnné a poskytují tedy respondentovi efektivnější rámec pro vyjádření postojů, díky kterému může přesněji vyjádřit své interní stavy; reliabilita měření je tedy vyšší než u krátkých škál. Rozlišení množství zprostředkované informace je nejlépe vidět u velmi krátkých škál - zatímco dva body (například souhlasím vs. nesouhlasím) přenesou pouze jednu informaci, směr postoje (pozitivní vs. negativní), tříbodová škála umožňuje navíc vyjádření neutrality. Zlepšení je vidět také u čtyřbodové škály, která kromě směru přenáší také informaci o intenzitě postoje. Všechny tři tyto typy informací pak spojuje pětibodová škála (směr, síla, střední hodnota). [Alwin 1997: 321 - 322] „Pokud tedy předpokládáme, že postoje existují na kontinuu, které má směr,

intenzitu a oblast neutrality, pak minimálně 5 bodů je nezbytných ke komunikaci postojů.“ [Alwin 1997: 321; překlad autorka] Informační teorie však může být rozšířena nad rámec pěti odpovědních kategorií – škály s větším počtem bodů rozlišují více stupňů intenzity. Na druhou stranu nekritické neomezené rozšiřování této teorie na více kategorií nelze odůvodnit – velmi dlouhé škály mohou mít své praktické limity. [Alwin 1997: 321 - 322] Existuje však nějaká hranice, za níž už rozšiřování škály nepřináší další zlepšení?

Kognitivní teoretici naznačují, že tento horní limit pravděpodobně existuje – počet kategorií, které ještě respondent zvládne. Při odpovídání na příliš dlouhé škály mají lidé kognitivní potíže s rozlišením jednotlivých bodů. Lze tedy očekávat, že kvalita měření se bude zvyšovat do určitého bodu, za nímž již bude informace ve skutečnosti ztracená, protože jednotlivé kategorie odpovědí mají tendenci znamenat méně. [Alwin 1997: 322] Kromě toho může dojít také k problémům spojeným s motivací. Dlouhé škály jsou pro respondenta kognitivně náročnější a ten nemusí být dostatečně motivován, aby se se složitým úkolem během dotazování potýkal, tedy aby vyvinul úsilí ke smysluplnému rozlišení jednotlivých bodů na škále. Místo optimalizace své odpovědi se tak může uchýlit k *satisficingu*. [Alwin 1997: 323] Otázkou tedy zůstává, kde se nachází hranice, za níž již rozšiřování škály nepomáhá, ale spíše škodí. Je jedenáctibodová škála využívaná při měření NPS již za touto hranicí? Nebo přináší více informací než krátká pětibodová škála při zachování únosnosti kognitivní úlohy?

#### **1.1.4 Délka hodnotící škály: provedené výzkumy**

Stejně jako existují různé teoretické pohledy, také výzkumníci se ve svých závěrech na základě provedených analýz liší. Dvojice Carolyn Preston a Andrew Colman [Preston, Colman 2000] podává přehled o výzkumech provedených do 90. let 20. století – někteří výzkumníci uvedli, že reliabilita měření není délkou škály ovlivněna, jiní došli k tomu, že reliabilita stoupá s vyšším počtem kategorií až do sedmi bodů, při dalším rozšiřování škály se ale již nezvyšuje. Nechybí ale ani závěry podporující jiné délky škál. Podobné přehledy přináší i další autoři a tato variabilita je patrná i v novějších studiích.

Přímo Preston a Colman [Preston, Colman 2000] provedli komplexní studii zabývající se reliabilitou a validitou různě dlouhých hodnotících škál, ale i preferencemi samotných respondentů. Zkoumali přitom škály se dvěma až jedenácti body a se 101 body (0 – 100). Podle jejich výzkumu je nejvyšší reliability (test–retest, interní

konzistence) dosaženo u škál se sedmi až deseti body, u delších již test-retest reliabilita mírně klesá. Validita je rovněž nižší u krátkých škál se dvěma až čtyřmi body a vyšší u pětibodových a delších škál. Design výzkumu nebyl stejný jako v této práci, zahrnoval pouze jednu (spíše menší: N = 149) skupinu respondentů, která odpovídala na sérii otázek týkající se kvality služeb poskytovaných obchodem a restaurací.<sup>3</sup> Ti nejdříve na škále od 0 do 100 ohodnotili kvalitu vybraného obchodu či restaurace celkově, pak ji hodnotili detailněji v 5 dimenzích. Každý respondent přitom odpověděl na stejnou sérii pěti otázek jedenáctkrát – pokaždé s jinou délkou škály. Jejich pořadí bylo náhodné. Zhruba za dva týdny vyplnila většina (návratnost 86 %) respondentů dotazník znovu pro otestování test-retest reliability. Z hlediska preferencí respondentů byly škály s 5, 7 a 11 body hodnoceny jako relativně jednoduché, krátké škály se 2 – 4 body jako rychlé na odpovídání, ale nevhodné pro adekvátní vyjádření postoje. Z tohoto hlediska byly podle respondentů naopak nejlepší dlouhé škály s 10, 11 a 101 body.

Duane Alwin [Alwin 1997] rovněž testoval vliv délky škály na kvalitu měření (na validitu a reliabilitu); porovnával přitom sedmibodovou (klasicky zobrazenou) a jedenáctibodovou škálu zobrazenou jako teploměr při hodnocení různých oblastí životní spokojenosti. Výzkum byl stejně jako předchozí studie designován jako multitrait multimethod model (MTMM) – obsahoval různé proměnné (oblasti životní spokojenosti), z nichž každá byla hodnocena na dvou délkách škály. Byl součástí *Quality of Life Survey* (rok 1987). Autor si je však vědom dvou omezení: 1) škály byly respondentům předkládány napříč otázkami i respondenty ve stejném pořadí, 2) slovní označení bodů se lišilo – zatímco na dlouhé škále měly verbální labely pouze krajní hodnoty (spokojen/a, nespokojen/a), u sedmibodové škály byl navíc označen i neutrální střední bod. Třetím omezením je dle autorky této práce porovnání prosté škály s „pocitovým teploměrem“. Alwin nicméně svým výzkumem podpořil informační teorii o tom, že více bodů přenáší informaci reliabilněji, rovněž odhady validity byly u jedenáctibodové škály vyšší.

Tým metodologů sestávající z Melanie Revilla, Willema Sarise a Jona Krosnicka [Revilla, Saris, Krosnick 2014] v kontrastu s Alwinovými závěry doporučuje použít pouze pětibodovou škálu (spíše než sedmi- nebo jedenáctibodovou), jedná se však o škálu vyjadřující souhlas, která se podle nich chová odlišně. Autoři provedli rovněž robustní experiment - v rámci *European Social Survey*. Použili komplexnější přístup,

---

<sup>3</sup> Obchod a restaurace byly vybrány na základě předpokladu, že všichni respondenti v nedávné době navštívili obchod nebo restauraci, což výzkum potvrdil. [Preston, Colman 2000: 4]

vyvinutý Sarisem a kolegy [Saris, Satorra, Coenders 2004], kombinující výhody split ballot experimentu a MTMM (SB-MTMM). Také v jejich případě však nebyly všechny aspekty škály zafixované, aby bylo možné sledovat pouze vliv počtu bodů – pětibodová škála měla slovně označené všechny hodnoty (rozhodně souhlasím, souhlasím, ani souhlasím, ani nesouhlasím, nesouhlasím, rozhodně nesouhlasím) a byla prezentována vertikálně. Škály se 7 a 11 body byly prezentovány jako horizontální řada čísel se slovním označením pouze krajních hodnot (rozhodně souhlasím, rozhodně nesouhlasím). Sami autoři tedy přiznávají, že vyšší kvalita dat získaná pětibodovou škálou může být kombinací vlivů délky škály a kompletního slovního označení.

Výše uvedené studie se zabývaly kvalitou dat, méně pozornosti však bylo věnováno dopadu délky škály na výsledky a popisné charakteristiky, jako je tomu také v této práci. Na tuto oblast se zaměřil John Dawes, zabývající se mj. metodologií výzkumu trhu, motivován otázkou, zda lze data získaná prostřednictvím různě dlouhých škál vzájemně srovnávat, například pokud dojde v longitudiálním výzkumu z nějakého důvodu ke změně délky škály nebo zda některá forma přináší výrazně lepší/horší výsledky. [Dawes 2008: 1] Dawes [Dawes 2002], provedl dva split ballot experimenty testující vliv délky škály (pětibodové, jedenáctibodové) na charakteristiky dat jako je průměr, variabilita, šikmost a špičatost. V prvním z nich šlo o face-to-face dotazování studentů univerzity na kavárnu v kampusu; pokládaná série otázek byla doprovázena buď pětibodovou, nebo jedenáctibodovou škálou vyjadřující míru souhlasu; slovní označení bodů se však mírně lišilo (zatímco u pětibodové byly body od rozhodně nesouhlasím po rozhodně souhlasím, u jedenáctibodové škály znamenala nula žádný souhlas a desítka 100% souhlas). V druhém experimentu probíhalo dotazování telefonicky a týkalo se využití taxislužby; opět byly použity stejné škály. Aby bylo možné porovnání průměrů, Dawes pětibodovou stupnici bodů přeškáloval, aby odpovídala jedenáctibodové škále (1 = 0; 2 = 2,5; 3 = 5; 4 = 7,5; 5 = 10), provedl tedy jednoduchou aritmetickou operaci, kdy krajní hodnoty přiřadil ke krajním a u ostatních bodů zvolil hodnotu na základě stejných intervalů mezi body. Variabilitu sledoval pomocí variačního koeficientu reflektujícího různé úrovně průměrů – tato relativní míra variability se počítá jako podíl směrodatné odchylky a průměru; Dawes ji (zřejmě pro přehlednost) ještě vynásobil 100: směrodatná odchylka / průměr \* 100. [Dawes 2002: 7] Výsledky experimentu neukazují žádný systematický vztah mezi délkou škály a tvarem dat (šikmost, špičatost), nicméně z hlediska průměru a variability jsou určité rozdíly patrné. Jedenáctibodová škála generovala mírně vyšší průměr (průměrně o 0,25 bodu),

část tohoto rozdílu však Dawes připisuje tomu, že jedenáctibodová škála má hodnotu nula, která při procesu přeškálování může uměle mírně snížit průměr pětibodové škály. [Dawes 2008: 4] Navíc je nutné mít na paměti, že nelze vyloučit také vliv různých slovních označení bodů na škále. Kromě mírné odlišnosti průměrů se lišila i variabilita v datech – jedenáctibodová škála přináší větší variabilitu, jelikož jednoduše poskytuje více možností odpovědi. Dawes tedy doporučuje spíše jedenáctibodovou škálu; hodnoty průměrů jsou podle něj porovnatelné, ovšem větší variabilita je vhodnější pro sledování závislosti mezi proměnnými.

Stejný autor realizoval o několik let později další split ballot experiment [Dawes 2008], tentokrát porovnávací škály s 5, 7 a 10 body. Dawes respondentům prostřednictvím telefonického dotazování položil sérii osmi otázek, respektive výroků, které se týkaly hlídání cen, nakupování v akci apod. Respondenti byli vybráni náhodně a rozděleni do tří skupin, z nichž každá získala stejné otázky s jinak dlouhou hodnotící škálou vyjadřující míru souhlasu od 1 rozhodně nesouhlasím do X rozhodně souhlasím, ostatní body nebyly slovně označené. Na rozdíl od první studie tedy byly všechny tři délky škál administrovány stejně. Dawes došel k závěru, že šikmost, špičatost a směrodatná odchylka se dle délky škály neliší. Průměry se (po přeškálování kratších škál na hodnoty od 1 do 10 dle metody výše) neliší mezi pětibodovou a sedmibodovou škálou, desetibodová však generuje o něco nižší průměr (průměrně o 0,3 bodu). Výsledek je tedy odlišný než v případě prvního experimentu, kde delší škála produkovala mírně pozitivnější výsledky; přímé porovnání však není možné, jelikož zde šlo o sudý počet (10) bodů, tedy bez neutrální střední hodnoty. Zastoupení jednotlivých odpovědí autor neuvádí, není tedy možné porovnat „průběh“ škál. Dawes však považuje výsledky získané na jednotlivých škálách za relativně srovnatelné. [Dawes 2002: 2; Dawes 2008: 9 - 10] Nad rámec těchto analýz Dawes dále ukázal, že mají-li respondenti na výběr větší počet odpovědí, tak jich také více využijí (průměrně 2,9 u pětibodové, 3,6 u sedmibodové a 4,0 u desetibodové škály v sérii osmi výroků).

## ***1.2 Kognitivní model procesu odpovídání ve výzkumných šetřeních***

Při úvahách o měření postojů je třeba uvažovat také o probíhajících mentálních procesech respondentů při odpovídání na otázky ve výzkumu; přesnost výzkumu totiž závisí na přesnosti jejich odpovědí. [Tourangeau, Rips, Rasinski 2000: 1 - 2] Tyto

mentální nebo kognitivní aspekty dotazovací situace zohledňují kognitivní přístupy zastřešené názvem *Cognitive Aspects of Survey Methodology* (CASM). Tento metodologický proud se věnuje chybám pozorování, tedy těm, které vznikají na úrovni tazatele, respondenta, výzkumného nástroje a módu sběru dat. Jeho cílem je odhalování možných zdrojů zkreslení vznikajících při procesu odpovídání a zvyšování kvality výzkumných nástrojů. Zdrojem tohoto přístupu jsou klasická dotazníková šetření a kognitivní psychologie. Institucionálnímu zakotvení nového směru v první polovině 80. let 20. století na seminářích/konferencích předcházela vzrůstající spolupráce kognitivních psychologů a výzkumníků realizujících velké výzkumné projekty ve Velké Británii a USA na konci 70. let. Ojedinelé snahy o zkoumání této problematiky lze však sledovat již dříve. [Vinopal 2009: 399]

Kognitivní vědci se zabývají různými *response effects*, tedy odlišnostmi ve výsledcích výzkumu v důsledku zdánlivě nedůležitých procesních atributů jako je pořadí odpovědí nebo délka škály. Právě postojové otázky jsou k těmto efektům náchylné. [Tourangeau, Rips, Rasinski 2000: 1 - 2] Pochopení procesu odpovídání je proto klíčové pro kvalitu výzkumu a minimalizaci rizika získání nepřesných odpovědí. Trojice amerických vědců, Roger Tourangeau, Lance J. Rips a Kenneth Rasinski [Tourangeau, Rips, Rasinski 2000], představila kognitivní model popisující proces odpovídání ve výzkumu, rámec pro pochopení response effects ve výzkumných šetřeních. Model je složen ze čtyř hlavních složek: 1) porozumění otázce/položce, 2) vyvolání informací z paměti, 3) využití těchto informací k požadovanému úsudku a 4) výběr a uvedení odpovědi. Každý z těchto kroků zahrnuje určité mentální procesy, které respondent provádí, v každém z nich může zároveň dojít k response effects – respondent může například nesprávně pochopit otázku, zapomenout klíčovou informaci, udělat chybný závěr na základě toho, co si vybaví z paměti, nebo vybrat pro svou odpověď nevhodnou možnost/kategorii odpovědi. [Tourangeau, Rips, Rasinski 2000: 7 - 8] Následující kapitoly představují tyto čtyři složky podrobněji.

### **1.2.1 Porozumění**

Porozumění zahrnuje procesy jako je vyslechnutí nebo přečtení otázky a doprovodných instrukcí, přiřazení významu její formě a vyvození, o co v otázce jde, jakou informaci zjišťuje. V rámci této fáze může dojít k řadě problémů spojených s nesprávnou interpretací otázky například v důsledku snížené pozornosti respondenta, nezaregistrování nebo ignorování instrukcí, ale i díky samotnému položení otázky – ta



může být například dvojitá (neúmyslně zjišťující dvě různé informace zároveň), příliš komplikovaná, obsahovat pro respondenta neznámé či vágní pojmy apod. [Tourangeau, Rips, Rasinski 2000: 9] V důsledku tak „otázka, na kterou respondent odpovídá, nemusí být tou, kterou výzkumník zamýšlel položit.“ [Tourangeau, Rips, Rasinski 2000: 23; překlad autorka] Otázka také již může u respondenta předpokládat nějakou charakteristiku a ptát se až na její aspekt, což v případě, že tato charakteristika neplatí, vede k obtížím. [Tourangeau, Rips, Rasinski 2000: 25] Různé typy otázek kladou na respondenty různé nároky; v porovnání s otázkami typu ano/ne znamenají otázky zjišťující kdo?, kdy?, kde?, proč?, jak? pro respondenta větší zátěž. [Tourangeau, Rips, Rasinski 2000: 38] Přetížení respondentovy procesní kapacity, pracovní paměti (například v důsledku přílišné komplexity otázky, pokrytí příliš mnoha oblastí), může mít dva důsledky - vypadnutí položky z pracovní paměti, tj. vynechání části výzkumníkem zamýšleného významu otázky ze zpracování, nebo zpomalení kognitivního procesu. K oběma těmto efektům může dojít také zároveň, tedy že respondent stráví hodně času při zodpovídání nekompletní interpretace otázky. [Tourangeau, Rips, Rasinski 2000: 40]

Důležité je také rozlišit okamžité pochopení otázky od její interpretace. Samotný pojem *porozumění* je ambivalentní, může znamenat obojí. [Tourangeau, Rips, Rasinski 2000: 30] Kde nedojde k problému s porozuměním otázky, může dojít k obtížím způsobeným její interpretací. V případě této práce je toto rozlišení důležité – zatímco porozumění jednotlivým slovům ani celé otázky „Do jaké míry byste doporučil(a) svým přátelům a známým následující společnost?“ nemusí komplikace způsobovat, její interpretace se již mohou mezi jednotlivými respondenty lišit. Ptá se otázka na sílu vztahu k dané instituci? Nebo na to, zda by respondent v reálné životní situaci svému příteli doporučil společnost, kterou sám využívá? Hraje roli, zda by ji doporučil aktivně, sám od sebe nebo pouze na vyzvání tohoto přítele? Měl by toto promítnout do své odpovědi? Autoři kognitivního modelu rozlišují dvě složky porozumění – jednu nutnou, druhou volitelnou. Obě jsou mentálními reprezentacemi<sup>4</sup> právě vyslechnuté/přečtené věty (otázky), liší se ale ve svém obsahu. První z nich, *reprezentace věty*, se týká struktury věty a významu slov a je víceméně konstantní napříč jazykově kompetentními

---

<sup>4</sup> Porozumění otázky lze chápat prostřednictvím psychologické kategorie mentální „reprezentace“, což je představa zastupující v mysli jev, který v daném okamžiku není vnímán smysly. Otázka v respondentovi tedy vyvolá představu reprezentující podstatné aspekty. [Vinopal 2009: 406] Je však třeba počítat s tím, že reprezentace je zatížena chybami vznikajícími při jejím ukládání i následném vyvolávání,

jednotlivci. Druhá, *representace o větě*, se liší na základě úhlu pohledu, znalosti tématu a kontextu, znalosti vyprávějící/tázající se osoby a pravděpodobně dalších faktorů, mj. také podle toho, kolik času a úsilí je osoba ochotna věnovat její interpretaci; může tedy být různá mezi respondenty i u jednoho respondenta napříč různými situacemi. Respondenti nepřestávají otázku interpretovat ve chvíli, kdy jsou hotovi s určením *representace věty*, tedy po kognitivních operacích jako je reprezentace logické formy otázky, rozeznání, na co se otázka ptá, nebo přiřazení klíčových pojmů k relevantním konceptům, ale pokračují dále, například tím, že se podívají na kategorie odpovědí a redefinují podle nich svou interpretaci otázky. Takový a další typy závěrů potom spoluvytváří *representaci o větě*. [Tourangeau, Rips, Rasinski 2000: 31 - 33] Kromě okamžitého porozumění otázce, porozumění jednotlivým slovům tedy dochází ještě k dalším procesům utvářejícím respondentovo chápání, interpretaci otázky na základě kontextu, zkušeností a podobně. Významy otázek tedy nejsou jejich fixními vlastnostmi, utváří je každý respondent při dotazování mj. také dle svých dojmů o celkovém významu rozhovoru, záměru otázky v rámci tohoto kontextu a v případě dotazování tazatelem také dle jeho chování při pokládání otázky. Výzkum jako takový může fungovat, protože pro mnoho otázek všichni respondenti přisuzují otázkám významy podobným způsobem; pokud tomu tak není, pak odpovědi respondentů nejsou vhodným indikátorem jediného atributu, který má otázka za úkol změřit. [Groves 2004: 419 - 420]

Jelikož má znění otázky velký dopad na výsledky výzkumu, bývají výzkumné nástroje podrobeny pretestům, například kognitivním rozhovorům. Odborné texty také někdy nabízí vodítka pro designování otázek tak, aby bylo sníženo riziko vzniku výše uvedených problémů. Tourangeau s kolegy doporučuje:

- Zachovat otázky jednoduché, vyhnout se komplikované větné skladbě
- Rozložit otázky zahrnující více oblastí do jednoduchých otázek pokrývajících vždy jen jednu oblast
- Definovat víceznačné nebo neznámé pojmy
- Vyhnout se vágním konceptům nebo v případě, že jsou nevyhnutelné, doplnit je příklady
- Nahradit vágní kvantifikátory kategoriemi specifikujícími přesné pravděpodobnosti, frekvence apod.

---

individuálním vnímáním, a tak nikdy stoprocentně neodpovídá prožité situaci, na kterou si respondent vzpomíná. [Vinopal 2009: 408]

- Školit tazatele, aby rozpoznali a napravili neporozumění na straně respondenta [Tourangeau, Rips, Rasinski 2000: 61]

### 1.2.2 Vyvolání informací z paměti

Tento kognitivní krok spočívá ve vyhledání relevantní informace v dlouhodobé paměti a její převedení do paměti pracovní, kde s ní může být dále nakládáno. Zahrnuje procesy jako je zaujetí strategie k získání informací z paměti, vytvoření podnětů k vyvolání vzpomínek, posbírání jednotlivých vzpomínek a doplnění dílčích vzpomínek odvozením. Přesnost a úplnost této složky kognitivního modelu může ovlivnit několik charakteristik z paměti vyvolaného materiálu, stejně jako položených otázek. [Tourangeau, Rips, Rasinski 2000: 9] Otázky pokládané ve výzkumu vyžadují, aby respondent nahlédl do své paměti; jsou i takové otázky, které jsou v tomto ohledu poměrně obtížné – čím vyšší nároky na paměť otázka klade, tím víc lze očekávat, že budou respondentovy odpovědi (a tedy i výsledky výzkumu) méně přesné. [Tourangeau, Rips, Rasinski 2000: 62]

Respondent musí hledat informace v paměti jak v případě faktických či behaviorálních otázek, tak při odpovídání na otázky postojové. Zjišťované (a tedy i vyvolávané) informace se ale různí - postojová otázka může zjišťovat názor na něco, o čem respondent doposud nikdy nepřemýšlel. [Tourangeau, Rips, Rasinski 2000: 62] Na rozdíl od faktických nebo behaviorálních se málokdy odkazuje na nějaký dobře definovaný soubor faktů; přestože existují fakta, která jsou pro rozhodování o odpovědi na postojovou otázku relevantní (např. u doporučení může být relevantní, zda respondent podal nějakou stížnost, plánuje od využívané společnosti odejít apod.) a mohou tedy být nepřímými vodítky při posuzování „přesnosti“ respondentovy odpovědi, jsou stěží definitivní a otázka se na ně přímo nezaměřuje. [Tourangeau, Rips, Rasinski 2000: 165 - 166]

Co přesně si tedy respondenti musí vybavit, aby byli schopni odpovědět na otázku týkající se jejich postoje? Odpovědi na postojové otázky mohou respondenti opírat například o minulé události či svá dřívější rozhodnutí v dané problematice. [Vinopal 2009: 407] Jak již ale bylo v této práci uvedeno, mít postoj automaticky neznamená mít předem (před dotazováním) soud o určitém problému, tedy že postoj je vždy předem v hlavě respondenta připraven; spíše existuje určité kontinuum zformování postoje. [Tourangeau, Rips, Rasinski 2000: 12] Zatímco respondenti, kteří mají v dané otázce silný postoj, odpovídají tak, že si vybaví a sdělí výzkumníkovi existující

hodnocení, ostatní respondenti musí projít odlišnými procesy. Zdokumentované problémy spojené s postojovými otázkami (efekt způsobený pořadím otázek, kontextem apod.) bývají připisovány právě této druhé skupině respondentů. Jak ale respondenti se slabým/žádným postojem otázky zodpovídají? Jako alternativní zdroj odpovědi mohou posloužit 1) dojmy nebo stereotypy (stejně jako může mít respondent u behaviorální otázky pouze mlhavou představu, jak často provádí určitou činnost, tak může mít jen vágní dojem z osoby nebo problému, na postoj k němuž se ho výzkumník ptá, a na tomto dojmu založit svou odpověď), 2) obecné postoje nebo hodnoty, z nichž respondent odvodí svou pozici ve specifickější, konkrétnější otázce, nebo naopak 3) specifické názory nebo pocity, z nichž respondent odvodí svou pozici v obecnější otázce. [Tourangeau, Rips, Rasinski 2000: 172] Ať již je při zodpovídání otázky využit jakýkoliv typ informace (v závislosti na jejich dostupnosti, ale například i na motivaci respondenta dospět k obhajitelné pozici), stabilita odpovědí v čase bude záviset nejen na konzistenci respondentova pohledu na problém, ale i na tom, zda při každé příležitosti vyvolává stejné informace a stejně si je vykládá. [Tourangeau, Rips, Rasinski 2000: 195]

### 1.2.3 Úsudek

Vyvolání relevantních informací často nepřináší explicitní odpověď na položenou otázku – někdy je nutné zkombinovat jednotlivé vyvolané informace, někdy je doplnit. V rámci této složky modelu dochází k procesům, jako je ohodnocení kompletnosti, relevance a přesnosti vybavených vzpomínek, vyvození závěrů na základě procesu vzpomínání nebo odhadů vyplňujícího mezery v získaných informacích (například na základě typického chování v situaci, která je předmětem otázky) a integrace veškerého vyvolaného materiálu do celkového úsudku, rozhodnutí. Respondenti tak přeměňují vyvolané informace do příslušné odpovědi.<sup>5</sup> [Tourangeau, Rips, Rasinski 2000: 8 - 11] Příkladem úsudku v rámci postojových otázek může být také situace, kdy respondent zmírní nebo nesdělí svůj názor, pokud cítí, že v paměti dostupné informace nejsou dostatečné. Může rovněž odpovědět na základě toho, co mu

---

<sup>5</sup> Úsudek obvykle probíhá na základě z paměti získaných informací, někdy ale může i zcela nahradit fázi vyhledávání – respondenti se jí mohou vyhnout, pokud například nikdy neslyšeli o organizaci, která je předmětem otázky, a díky tomu rovnou pouze na základě tohoto faktu usoudí, že nejsou jejími členy, nebo zodpovědním retrospektivní otázky pouze s ohledem na obecnou věrohodnost odpovědi. [Tourangeau, Rips, Rasinski 2000: 11]

nejednodušší přijde na mysl, nebo doplnit informace, které si nemůže vybavit, stereotypy nebo schémata. [Tourangeau, Rips, Rasinski 2000: 13]

V rámci tohoto modelu jsou úsudky dočasnými konstrukcemi – ačkoliv mohou stát na dlouhotrvajících hodnoceních, úsudky samotné musí být obvykle vytvořeny v odpověď na otázku, tedy poměrně rychle. Úsudky při měření postojů jsou tak poměrně závislé na kontextu – stejná otázka často přináší různé odpovědi v různých kontextech. Tourangeau s kolegy věří, že stabilita odpovědí na postojové otázky napříč různými příležitostmi a kontexty reflektuje 1) homogenitu jednotlivých respondentových úvah vztahujících se k danému tématu, 2) překryv souborů úvah, které jsou využívány pro odpovědi na různé otázky nebo v různých kontextech, 3) reliabilitu procesu ohodnocení těchto úvah; přitom kontext otázky může ovlivnit všechny tyto tři aspekty. [Tourangeau, Rips, Rasinski 2000: 197] Mimoto kontext může působit také na kombinaci specifických vybavených úvah do celkového úsudku. Není vždy zřejmé, jak vybavené informace při konstrukci odpovědi na postojovou otázku využít; respondenti mohou v některých případech různým úvahám přiřadit různou váhu nebo například zvolit nějaké pravidlo nebo proces, pomocí něž vyvolané informace zkombinují. Všechny takové procesy ale mohou být ovlivněny kontextem. [Tourangeau, Rips, Rasinski 2000: 210]

#### 1.2.4 Editace odpovědi

Poslední složkou modelu je výběr a uvedení odpovědi, což znamená zařazení výsledného úsudku do příslušné odpovědní kategorie, přizpůsobení odpovědi nabízeným možnostem. [Tourangeau, Rips, Rasinski 2000: 13, Vinopal 2009: 413]

Ač může mít respondent jasnou odpověď, kterou chce uvést, nemusí být vždy jasné *jak* ji uvést. Možnosti odpovědí u postojových otázek bývají vágní (rozhodně souhlasím, spíše souhlasím...) a není jasné, co přesně znamenají, kde je hranice jedné a druhé odpovědi. Kromě toho se mohou lišit přístupy respondentů k výběru odpovědi, pokud shledají více než jednu možnost vhodnou – zatímco někteří se velmi snaží vybrat tu nejlepší možnou, jiní se spokojí v první vhodnou zvažovanou odpovědí. Stejně tak se může lišit ochota respondentů odpověď vůbec uvést nebo raději říci, že neví. Z těchto důvodů se výzkumníci zabývají formátem možností odpovědí, jako je například vhodný počet bodů na škále, zda by měla být každá z nich slovně označena, jaké by mělo být jejich pořadí a zda nabídnou možnost „nevím“. [Tourangeau, Rips, Rasinski 2000: 13 - 14] Respondenti mají sklon zjednodušovat si reportování své odpovědi, jak jen je to

možné, například uváděním rozmezí nebo kulatých hodnot při kvantifikaci, při hodnocení sledovat pouze několik jednoduchých principů, zcela obejít seriózní zvážení otázky apod. Může dojít také k tzv. *satisficingu*, kdy se respondent snaží svou odpovědí pouze zavděčit tazateli/výzkumníkovi, tedy uvést přijatelnou, nikoliv však optimální odpověď, kvůli které by musel zcela pochopit otázku, vybavit si všechny relevantní informace atd. Kromě únavy, nezájmu o výzkum či neschopnosti vyrovnat se s požadavky dané úlohy, jsou tedy respondenti navíc náchylní k tomu, aby redukovali zátěž, kterou na ně otázka klade. V zájmu výzkumníků je proto pokládat jednoduché otázky a tuto zátěž minimalizovat. Respondenti totiž nemají žádný důvod, aby tvrdě pracovali na odpovědích na složité otázky. [Tourangeau, Rips, Rasinski 2000: 254]

Otázky v dotaznících mají různé podoby (otevřené, uzavřené s číselnými kategoriemi, slovním označením nebo obojím, s různým počtem kategorií a jejich řazením) a každá z nich slouží jinému účelu. Dle formátu otázky se liší také procesy, ke kterým dochází v rámci editace odpovědi. Formátem relevantním pro tuto práci jsou uzavřené otázky, jejichž odpovědi jsou seřazeny, konkrétně do formy hodnotící škály. V tomto případě hrají roli záchytné body a další vlivy, které mohou zkomplikovat vztah mezi úsudkem a odpovědí. [Tourangeau, Rips, Rasinski 2000: 230 - 232] Je zdokumentováno několik efektů, ke kterým může dojít při využití hodnotících škál. Jednak dochází k tzv. *positivity bias*, tedy že odpovědi bývají vychýleny k pozitivnímu konci škály, zdroj tohoto vychýlení však není jasný. [Tourangeau, Rips, Rasinski 2000: 240 - 241] Dále je zřejmá také určitá neochota respondentů používat extrémy, krajní body škály – hodnocení tedy zmírňují, podceňují objekty spadající do horního konce škály a naopak přeceňují ty, které spadají na spodní konec. Tento trend je nazýván *response contraction bias*. Dochází k němu v případech, kdy respondenti snadno rozeznají střed škály, například v případě stupnice od 0 do 10 bývají extrémy využívány málo a odpovědi se shlukují okolo střední hodnoty 5. [Tourangeau, Rips, Rasinski 2000: 244] Na druhou stranu Groves upozorňuje na tendenci, kdy je neutrální střední kategorie se zvyšujícím se počtem kategorií vybírána méně často – větší zastoupení střední kategorie je tedy na kratších škálách. [Groves 2004: 465] Kromě nadužívání středních kategorií může dojít také k nadužívání tzv. *kognitivních referenčních bodů*, tedy specifických čísel jako je 0, 1, 10, 100 atd. Jedná se o speciální, snadno identifikovatelná čísla. Hodnoty na škále, které jsou referenčními body, mohou být nejen nadhodnoceny, ale jelikož jsou vnímány jako speciální, mohou se jim respondenti i vyhýbat, například v případě nuly, která má v numerickém systému speciální místo.

[Tourangeau, Rips, Rasinski 2000: 245] S jednotlivými body mohou mít respondenti také spojeny určité konotace, nadužívat kategorie, které obsahují standardní nebo typickou položku. Standard může poskytnout buď již přímo otázka, nebo jej respondent zná ze zkušenosti. [Tourangeau, Rips, Rasinski 2000: 244 - 245] V neposlední řadě je úloha ovlivněna i prostým počtem kategorií, délkou hodnotící škály. Pokud je příliš krátká, nemusí rozlišit různé úsudky respondentů, je-li naopak příliš dlouhá, mohou nastat komplikace s odlišením sousedících kategorií. [Tourangeau, Rips, Rasinski 2000: 249]

### 1.2.5 Shrnutí a aplikace modelu

Ačkoliv představený model obsahuje řadu kognitivních procesů, trvá zodpovězení postojové otázky obvykle méně než pět sekund. Na místě je otázka, zda lze uvažovat o sekvenci čtyř hlavních složek modelu, tedy zda jde o oddělené, na sebe navazující fáze, anebo jsou to třídy příbuzných procesů. Bohužel pro příklon k jedné z variant není dostatek důkazů. Na jednu stranu lze předpokládat, že u většiny otázek budou respondenti sledovat logickou návaznost, tedy že nejdříve pochopí otázku, poté si vybaví relevantní informace, provedou na jejich základě úsudek a ten poté editují do vhodné možnosti odpovědi. Stejně tak je ale zřejmé, že proces odpovídání může různě variovat. Autoři modelu předpokládají, že jednotlivé komponenty se budou často alespoň částečně překrývat. Také připouští možnost vrátit se v procesu z „pozdější“ fáze do „dřívější“ nebo vynechat či zkrátit některou ze složek. Přestože tedy předpokládají, že porozumění – vyvolání informací – úsudek – odpověď jsou nejčastějším sledem komponentů, nelze je popsat jako čtyři nepřekrývající se fáze. [Tourangeau, Rips, Rasinski 2000: 14 - 16] Příkladem může být pochopení otázky – jde sice o první fázi procesu, není však ohraničená, smysl otázky může být v průběhu odpovídání respondentem reinterpretován, ať již z důvodu nadbytečnosti vybavených informací pro původní pojetí otázky nebo protože odpověď nevystihuje žádná z nabízených možností. [Vinopal 2009: 406]

Model je užitečný jako rámec pro porozumění zdrojům response effects a jelikož nabízí metody jejich redukce. [Tourangeau, Rips, Rasinski 2000: 19] Poskytl však nový pohled také na studium postojů jako takových a na tradiční předpoklad, že respondent má předem existující odpověď na většinu postojových otázek a stačí mu tedy si otázky pouze přečíst/vyslechnout. [Tourangeau, Rips, Rasinski 2000: 22]

## 2. Metodika výzkumu

Tato kapitola pojednává o využitých metodách a postupech při hledání odpovědi na výzkumné otázky. Otázky lze rozdělit do dvou základních okruhů:

- **Ovlivňuje délka škály výsledná data?** Přináší pětibodová a jedenáctibodová škála rozdílné výsledky? Pokud ano, jak se liší?
- **Jak funguje proces odpovídání při různých délkách škál?** Jaké jsou strategie respondentů při odpovídání? Jak náročné je vybrat odpověď na škále? Jsou respondenti schopni rozlišit mezi jednotlivými body?

Aby bylo možné přinést odpovědi na oba okruhy výzkumných otázek, výzkum kombinuje dvě různá řešení: *split ballot experiment* a *kognitivní interview*. Obě tyto metody slouží k testování a evaluaci výzkumných nástrojů. [Presser et al. 2004] Zatímco účelem experimentálního designu je porovnat dvě verze testované položky (zde dvě různé délky škály) z hlediska distribuce odpovědí, tj. odhalit vliv na výsledná data, kognitivní interview přináší hlubší porozumění procesu odpovídání a jeho obtížnosti pro respondenta.

### 2.1 Kognitivní interview

Metoda kognitivních rozhovorů existuje zhruba od poloviny osmdesátých let a od té doby se prosadila jako nástroj k ohodnocování dotazníků. Je plodem vzájemné diskuze mezi výzkumníky realizujícími výběrová šetření a kognitivními psychology. [Beatty 2004: 45; Tourangeau, Rips, Rasinski 2000: 326] Metoda vychází z perspektivy CASM<sup>6</sup> (stručně představené v teoretické části práce), dle níž je odpověď na otázku řízena kognitivními procesy respondenta a porozumění těmto procesům je tedy klíčové pro sestavování výzkumného nástroje. [Willis 2004: 23] Kognitivní interview se tak od běžných rozhovorů liší ve svém záměru: „Běžné rozhovory se zaměřují na produkování kódovatelných odpovědí na otázky. Kognitivní rozhovory naopak poskytují pohled na procesy otázkami vyvolané.“ [Presser et al. 2004: 4; překlad autorka] Jejich cílem je zmapovat myšlenkový proces vedoucí od interpretace otázky až k jejímu zodpovězení, strategii respondenta při odpovídání a diagnostikovat tak případné problémy, které

---

<sup>6</sup> Prvotním impulzem k provádění kognitivních rozhovorů byl interdisciplinární seminář na téma kognitivních aspektů metodologie výběrových šetření (The Advanced Research Seminar on Cognitive Aspects of Survey Methodology) v roce 1983. Obsah tohoto setkání shrnul Jabine et al. [Jabine et al. 1984]



otázka/instrument vyvolává. [Presser et al. 2004: 4] Kromě toho je prostřednictvím kognitivních rozhovorů možné zjistit, zda otázky v dotazníku generují zamýšlené odpovědi [Beatty, Willis 2007: 288], tj. jestli je otázka správně pochopena. Z hlediska modelu Rogera Tourangeau a kolegů [Tourangeau, Rips, Rasinski 2000], popsáno v teoretické části práce, je tedy cílem kognitivního interview posoudit, zda jsou respondenti schopni 1) porozumět otázce tak, jak ji výzkumník zamýšlel, 2) vybavit si relevantní informace vztahující se k otázce, 3) udělat si úsudek ohledně vyvolaných informací, 4) odpovědět pomocí porovnání tohoto vnitřního úsudku s nabízenými kategoriemi odpovědí. [Willis 2004: 25]

Pojem kognitivní interview je však poměrně široký a může v sobě zahrnovat celou řadu kognitivně inspirovaných technik a procedur. Neexistují všeobecně sdílené standardy pro provádění kognitivních rozhovorů, „best practices“ o tom, které techniky využít. [Tourangeau, Rips, Rasinski 2000: 326 - 327; Conrad, Blair 2004: 67] I přes tyto rozdíly však lze v jednotlivých technikách najít společného jmenovatele. Každé kognitivní interview zahrnuje individuální interakci tazatele s respondentem, která může sledovat předem připravený protokol/scénář. Již se ale může lišit míra, do jaké se tazatelé mají scénáře držet. [Conrad, Blair 2004: 72] Kognitivní rozhovory tak mohou být strukturované, polo-strukturované či zcela nestrukturované, improvizované na základě diskuze s respondentem. [Beatty, Willis 2007: 289] Nejčastěji formou jsou polo-strukturované rozhovory. [Conrad, Blair 1996: 1]

Jednotlivé praktiky jsou pak založeny na dvou základních paradigmatech – *thinking-aloud* a *probing*. „Jedno zahrnuje kognitivního tazatele, jehož rolí je usnadnit účastníkům verbalizaci jejich myšlenkových procesů, ale zasahovat do generování této verbální informace co nejméně je to možné. Druhé zahrnuje tazatele, který vede interakci aktivněji, obvykle pokládající dodatečné, přímé otázky o základu odpovědi.“ [Beatty, Willis 2007: 289; překlad autorka] Aplikací těchto dvou technik – nabádání respondenta k přemýšlení nahlas (tzv. *think-alouds*) a pokládání dodatečných otázek, sond (tzv. *follow-up probes*) – se kognitivní interview liší od běžných rozhovorů [Willis 2004: 24]

### 2.1.1 Thinking-aloud

Původní paradigma kognitivního dotazování se velmi silně spoléhá na „přemýšlení nahlas“, které výzkumníkovi přináší vhled do myšlenkových procesů při plnění nějakého úkolu (zde odpovědi na otázku). Základními předpoklady této

procedury jsou samozřejmě dostupnost tohoto typu informací respondentovi, schopnost je přesně vystihnout a že vypovídání o těchto procesech zásadně nezmění samotné myšlenkové aktivity, o nichž je referováno. Pokud tyto předpoklady platí, technika think-alouds přináší informace o tom, jakým způsobem si respondenti vybavují vzpomínky. Role tazatele není v případě této techniky příliš aktivní – žádá respondenty, aby říkali nahlas, co si myslí při odpovídání na otázku, a pokud je třeba, připomíná jim, aby v tom pokračovali. [Beatty, Willis 2007: 290] Přemýšlení nahlas však může být pro respondenta náročné, někteří jednotlivci navíc nejsou v této aktivitě zdatní, nemusí neobvyklý úkol přemýšlet nahlas zvládnout. [Willis 1999: 4 - 5] S těmito nesnáze se setkala i autorka této práce; z toho důvodu nebyla technika think-alouds v kognitivních rozhovorech příliš využívána.

### 2.1.2 Verbal probing techniques

Alternativním paradigmatickým k thinking-aloud je tzv. *probing*, někdy nazývaný také jako *intenzivní dotazování*. Při kognitivním rozhovoru založeném na této technice je tazateli dovoleno pokládat respondentovi přímé otázky, sondy – například o chápání položené otázky či konkrétního termínu. [Beatty, Willis 2007: 290 - 291] „Potom, co tazatel položí otázku z dotazníku a subjekt odpoví, tazatel se zeptá na další, specifickou informaci relevantní pro otázku nebo pro získanou odpověď. Tazatel v podstatě ‚sonduje‘ dále po podstatě odpovědi.“ [Willis 1999: 5; překlad autorka, uvozovky v originále] Role tazatele je tedy v tomto případě aktivnější.

Přímé otázky či sondy (anglicky *probes*) mohou mít různý charakter. Kognitivní psycholog a metodolog Gordon B. Willis [Willis 1999: 6] uvádí šest základních kategorií kognitivních dotazů:

- Sondy týkající se porozumění/interpretace, tedy pochopení otázky respondentem
- *Parafrázování* – respondent je požádán, aby položenou otázku zopakoval vlastními slovy
- *Posouzení jistoty* – respondent je dotázán, do jaké míry si je jist svou odpovědí
- *Sondy týkající se paměti, vyvolání vzpomínky* – příkladem může být otázka: „Jak si pamatujete, že jste šel/šla za doktorem v posledních 12 měsících pětkrát?“ Respondent je tedy požádán, aby vysvětlil proces, jakým došel ke své odpovědi, tj. právě k číslu pět.

- *Specifikující sondy* – doptávání se na detaily, pozadí a důvody dané odpovědi
- *Obecné sondy* – další typy dotazů, jako např. jak respondent zvolil odpověď, zda to bylo náročné, proč při odpovídání zaváhal apod. Paul Beatty, výzkumník amerického National Center for Health Statistics (NCHS), tuto skupinu dotazů označuje jako *kognitivní sondy*<sup>7</sup>, jejichž účelem je porozumění výpočetním procesům, odhalení informací zvažovaných při odpovídání a zjištění míry obtížnosti, na jakou respondent při přemýšlení o odpovědi naráží. [Beatty 2004: 58]

Kromě výše uvedených kategorií, představených Willisem, lze najít ještě další. Podle Beattyho [Beatty 2004: 58 - 59] se jedná o:

- *Potvrzující sondy* – tazatel se ujistí o správnosti doposud získaných odpovědí; jednou z možností je tzv. *nastavení zrcadla*, kdy tazatel zopakuje doslova, co respondent řekl jako implicitní požadavek potvrzení
- *Rozšiřující sondy* zjišťující dodatečné informace, nad rámec doposud dodané odpovědi
- *Funkční poznámky* navracející respondenta k diskutované otázce v dotazníku
- *Zpětná vazba* – poznámky, které podporují respondenta v žádoucím stylu odpovídání, např. „Děkuji, toto je přesně ten typ informace, který hledám.“

Kromě různých technik sondování je důležité rozlišit, zda se jedná o *souběžný* či *retrospektivní probing*. V prvním případě probíhá kognitivní interview tak, že tazatel položí otázku z dotazníku, respondent na ni odpoví a následuje fáze sondování. To se opakuje pro každou z testovaných otázek. V případě retrospektivního probingu tazatel s respondentem nejdříve projde celý dotazník a až potom mu položí sondovací otázky. Jedná se tedy o takový „debriefing“. Výzkumníci častěji využívají souběžný probing, a to z důvodu čerstvosti zjišťovaných informací. [Willis 1999: 7]

### 2.1.3 Kombinace technik

„Cílem obou paradigmat probingu a think-alouds je generovat verbální informace, které obvykle zůstávají v rámci dotazování skryté, za účelem zhodnocení, jak dobře plní otázky svůj úkol.“ [Beatty, Willis 2007: 292; překlad autorka] Vzhledem ke společným cílům nemusí oba přístupy existovat pouze odděleně; think-alouds a verbal probes lze v rámci kognitivního rozhovoru kombinovat [Beatty 2004: 57],

<sup>7</sup> Do této skupiny Beatty řadí také otázky týkající se interpretace pojmů, ty jsou v této práci ze skupiny vyčleněné dle Willisova vzoru.

čehož také výzkumníci v praxi využívají. [Willis 2004: 36] Vzhledem k odlišnému charakteru obou technik se tato strategie ukazuje jako praktická; zatímco think-alouds umožní nahlédnout do myšlenkových procesů při vzpomínání, vyvolávání informací, pomocí cílených sond výzkumník získá dodatečné informace jako je respondentova znalost pozadí zkoumaného tématu či porozumění termínům, znění otázky. [Willis 2004: 36] Technika think-alouds může v odhalení toho, jak respondent interpretuje či dezinterpretuje použité pojmy, selhat. [Presser et al. 2004: 6] Stejně tak není jistá užitečnost přemýšlení nahlas při odhalování nevhodných možností odpovědí či dalších problémů dotazníku. [Beatty, Willis 2007: 294] Také z toho důvodu autorka této práce využila především sondující otázky, kombinované s občasnými instrukcemi k přemýšlení nahlas.

#### **2.1.4 Získaná data**

„Kognitivní rozhovory produkují verbálně poskytované informace a tyto považujeme za syrová data pro další analýzu.“ [Conrad, Blair 2004: 72; překlad autorka] Jedná se o data kvalitativního charakteru. [Willis 1999: 28] Je však otázkou, co všechno je součástí těchto dat – a čím aktivnější roli přejímá tazatel, tím je tato problematika nejasnější. Je zřejmé, že součástí dat jsou respondentovy think-alouds, stejně jako odpovědi na sondující otázky; pravděpodobně sem patří i další respondentovy poznámky. Měly by ale být zahrnuty i výroky tazatele, jako například rekapitulace toho, jak chápe respondentovo vyjádření? Tato otázka prozatím nebyla uspokojivě zodpovězena. Je však důležité odlišit sběr dat (získání verbálních zpráv) od analýzy (interpretace těchto zpráv). [Conrad, Blair 2004: 73] Shrnutí respondentových odpovědí, vysvětlení jejich významu je v této práci chápáno jako interpretace. Tento typ informací není zahrnut mezi primární data; může sloužit jako podnět pro další verbální informace podané respondentem, nicméně není samostatně analyzován.

Z praktického hlediska může být analýza dat z kognitivních rozhovorů založena na poznámkách pořízených během interview nebo systematických prepisech rozhovorů. [Beatty, Willis 2007: 289] Nahrávání rozhovorů a poslech těchto nahrávek samozřejmě vyžaduje dodatečné úsilí a čas; při objasnění procesu odpovídání a identifikaci problémů se však vyplatí. [DeMaio, Landreth 2004: 107]

### 2.1.5 Analýza dat

Data z kognitivních rozhovorů je potřeba interpretovat a odhalené problémy zařadit do kategorií. Kódovací schéma by mělo být vyčerpávající a dobře definované, aby bylo možné zařadit jakýkoliv problém. [Conrad, Blair 2004: 73 - 74] Analýza dat by měla být časově oddělena od sběru dat, tedy od realizace samotných rozhovorů - analytik se tak nemusí zabývat vedením rozhovoru a může svou pozornost plně zaměřit na obsah verbálních informací získaných od respondentů. [Conrad, Blair 1996: 1]

Analýza kognitivních interview se může stát zdrojem limitované reliability této metody: „Je snadné vidět, že různí tazatelé mohou dojít k různým závěrům o podstatě a lokaci problémů v dotazníku v závislosti na tom, jak prováděli interview a na jaký typ náznaků jsou citliví.“ [Conrad, Blair 1996: 1; překlad autorka] Je proto důležité věnovat jí náležitou pozornost. Někteří autoři se pokusili o vytvoření kódovacích schémat, která dovolují důslednější analýzu. Tato schémata obvykle reflektují čtyř-složkový kognitivní model představený v teoretické části této práce. [Tourangeau, Rips, Rasinski 2000: 328]

Dvojice amerických metodologů Frederick G. Conrad a Johnny Blair [Conrad, Blair 1996] vyvinula taxonomii problémů určenou ke klasifikaci verbálních zpráv respondentů o obtížích s otázkou. Kromě identifikace aspektu otázky, který problém způsobuje (slovní, logický apod.) jejich kódovací schéma umožňuje přiřadit jej ke kognitivnímu procesu, při kterém k němu dochází (pochopení otázky, vybavení si relevantní informace apod.). „Identifikováním fáze odpovědi, při níž se problém pravděpodobně vyskytl, se jistá řešení problému stanou slibnějšími, zatímco jiná se stanou méně přijatelnými.“ [Conrad, Blair 1996: 1; překlad autorka] Tento model není pro účel této práce bezvýhradně vhodný, autorka se jím však inspirovala. Ač Conrad a Blair akceptují čtyř-složkový model, který představil Tourangeau s kolegy, pro účely klasifikace verbálních zpráv respondentů spojují vyvolání informace a jejího následného posouzení do jedné – provedení úkolu. Analytik podle nich nemůže v think-alouds a odpovědích na sondující otázky rozlišit získávání informace z dlouhodobé paměti a zhodnocením, co bylo takto načteno. [Conrad, Blair 1996: 2 - 3] Zejména to může být komplikované při převládajícím využití techniky probingu jako v případě této práce – respondent předkládá jednotlivé vyvolané informace i jejich případné doplnění již integrované, společně s rozhodnutím o svém postoji. Navíc se mohou obě složky prolínat – jak již bylo uvedeno, nejde o oddělené, na sebe navazující fáze. Při analýze

dat z kognitivních rozhovorů jsou proto v této práci složky vyvolání informace a úsudku spojeny.

## **2.2 Split ballot experiment**

Pomocí kognitivních interview výzkumníci dokážou identifikovat vlastnosti otázek, které mohou způsobit problém či ovlivnit kvalitu dat. Metoda je však limitovaná – neposkytuje žádnou přímou informaci o tom, jak jsou data ovlivněna nebo u jaké části populace se problém projeví (kognitivní interview se provádí na malých výběrech, u nichž není usilováno o reprezentativitu). Navíc, je-li otázka na základě kognitivních rozhovorů upravena, není jasné, zda bude kvalita dat získaných alternativní verzí otázky vyšší. Přínos kognitivních interview se tak může plně projevit až při realizaci společně s experimentem testujícím různé varianty. [Fowler 2004: 174 - 175] Metody lze využít společně při pretestování výzkumného nástroje – v takovém případě je nejdříve dotazník podroben kognitivním rozhovorům, které adresují problémy. Po jejich úpravě je otestována původní a alternativní verze nástroje pomocí experimentu. [Presser et al. 2004: 11] V této práci jde o odlišné využití kombinace obou metod – nejdříve je proveden experiment, následně kognitivní rozhovory. Tento design vychází z výzkumného záměru autorky, kterým není pretest dotazníku pro zvýšení validity a reliability měření, ale metodologický experiment testující dvě různě dlouhé hodnotící škály a jejich vliv na výsledky výzkumu. Pětibodová škála, která je alternativní k (pro NPS originální) jedenáctibodové škále, tak není reakcí na kognitivní interview; její použití je motivováno záměrem porovnat dlouhou a krátkou hodnotící škálu.

Jednoduchým experimentálním designem, který umožňuje porovnání různých verzí určité položky, je tzv. *split ballot experiment*, nazývaný také *split sample experiment*. [Groves 2004: 305] Cíl takového experimentu může být ryze praktický (vybrat lepší verzi dotazníku pro nadcházející šetření nebo například posoudit dopad přechodu na novou verzi dotazníku v longitudiální studii), ale i obecnější - identifikovat charakteristiky nebo výzkumné principy vedoucí k otázkám, které přináší reliabilnější a přesnější odpovědi [Tourangeau 2004: 223 - 224], případně otestovat efekt určitého aspektu otázky na měření.

Koncept stojící za designem je jednoduchý – dvěma<sup>8</sup> nebo více podskupinám výběru jsou položeny odlišné verze otázky se stejným cílem; distribuce odpovědí na různé verze otázky jsou následně porovnány. Podstatou designu je – stejně jako v případě jiných experimentů – náhodné přiřazení jednotek do jedné z těchto experimentálních skupin. [Fowler 2004: 176] Právě princip randomizace odlišuje společně s manipulací nezávisle proměnnou (verze otázky, dotazníku) experiment od ostatních výzkumných strategií. [Kirk 2009: 23] Randomizace zaručuje náhodné rozložení specifických charakteristik respondentů napříč experimentálními skupinami, takže výsledek experimentu není těmito charakteristikami zkreslen – skupiny jsou v době rozdělení srovnatelné. [Kirk 2009: 24] Náhodným rozdělením respondentů do skupin je zaručena vysoká interní validita designu. [Kirk 2009: 27; Fienberg, Tanur 1989: 1018] Kromě randomizace je také nezbytné, aby měření v obou experimentálních skupinách probíhala ve stejný čas a za stejných podmínek [Kirk 2009: 28], tj. například aby byla aplikována stejná metoda sběru dat. V opačném případě by nebylo možné rozlišit efekt metody sběru dat od vlivu sledované variace otázky. [Tourangeau 2004: 213] Je proto důležité testovat vždy jen jeden prvek.

### 2.2.1 Výběrový soubor

Experimenty jsou často prováděny pouze v jedné nebo několika vybraných lokalitách a s relativně malými vzorky. Nebezpečí skrývající se za malým vzorkem je nedostatečná síla testu, možnost selhání při hledání rozdílů mezi skupinami. Nebezpečí nerepresentativního vzorku jsou samozřejmě nerepresentativní výsledky experimentu. [Tourangeau 2004: 212 - 213] Pro porovnání obou skupin je klíčová randomizace, tedy nějaký mechanismus, který náhodně přiřadí každého respondenta do jedné z experimentálních skupin. V případě, že je experiment použitý k pretestování výzkumného nástroje před začátkem dotazníkového šetření, je třeba zajistit také to, aby výběr reprezentoval cílovou populaci daného šetření.

---

<sup>8</sup> Split ballot experimenty byly prováděny již v době, kdy nebyly informační technologie běžnou součástí sběru dat v dotazníkových šetřeních (tj. před představením „computer-assisted“ metod). S ohledem na náklady za tisk tedy z praktických důvodů bývaly pouze dvě verze dotazníku. Samozřejmě existují výjimky, nicméně jedna variace jednoho faktoru – tj. dvě verze dotazníku - jsou pro split ballot design nejobvyklejší. [Sniderman, Grob 1996: 379]

## 2.2.2 Porovnání dat

Co se týče porovnání variací testované otázky, nejjednodušší řešení se zde zdá být zároveň klíčové: „Nejdůležitější porovnání je často jednoduše porovnat rozložení odpovědí. Pokud se zdají být distribuce odpovědí prakticky totožné, lze vyvodit, že tyto dvě otázky jsou prakticky stejné.“ [Fowler 2004: 176; překlad autorka] Jak ale porovnat distribuce odpovědí v případě, že testovanou variací je počet bodů na odpovědní škále?

Výzkumníci před porovnáním distribucí odpovědí proměnné rekódují/přeškálují, aby měly stejný počet bodů. [viz např. Dawes 2008; Preston, Colman 2000] Jedním z příkladů je formule využitá Carolyn Preston a Andrew Colmanem pro přeškálování na hodnoty do 100:  $(\text{hodnota} - 1) / (\text{počet bodů na škále} - 1) * 100$ . [Preston, Colman 2000: 7] Pokud například chceme převést hodnotu 4 na pětibodové škále, výpočtem  $(4 - 1) / (5 - 1) * 100$  dojdeme k nové hodnotě 75. Je-li nejnižší hodnota na škále 1, vždy je převedena na hodnotu 0 a nová škála tak nabývá hodnot od 0 do 100. John Dawes navrhuje malou úpravu tohoto výpočtu, a totiž násobení 10 (přeškálování na 0 až 10). [Dawes 2008: 7] Tento alternativní výpočet tak lze využít v této práci pro přeškálování pětibodové odpovědní škály na jedenáctibodovou od 0 do 10, aby bylo možné porovnání s druhou variantou. Druhým způsobem je rekódování delší škály, tj. sloučení některých jejích hodnot. Existují různé možnosti slučování, které mohou vycházet například z teorie nebo z vnímání bodů na škále respondenty (v této práci zjišťováno prostřednictvím kognitivních rozhovorů). Vzhledem k aplikačnímu příkladu využitému v této práci se nabízí výpočet Net Promoter Scoru (představeno dále).

Jakmile máme data takto upravena, kromě distribuce odpovědí můžeme porovnávat také deskriptivní statistiky jako je průměr, rozptyl, šikmost, špičatost. [Dawes 2008]

## 2.3 Diskuze použitých metod

Na výzkumné otázky bylo nutné odpovědět pomocí výzkumu kombinujícího vzhled do kognitivních procesů spojených s odpovídáním na hodnotící škále a přístupu umožňujícího porovnání odpovědí při využití různých délek škál. Byla tedy zvolena kognitivní interview a split ballot experiment.

Doplňkem k experimentu může být místo kognitivních interview také kódování chování respondentů a tazatelů při dotazování (tzv. *behavior coding*). Otázky, které jsou při využití této metody spojeny s velkým množstvím určitých kódů (např. tazatel



nepřečetl celé znění otázky, respondent požádal o ujasnění), jsou považovány za hodné úpravy. [Presser et al. 2004: 8] Tato metoda ovšem musí být spojená s face-to-face nebo alespoň telefonickým dotazováním, protože vyžaduje interakci tazatele s respondentem. To by však experiment učinilo podstatně nákladnějším; mimoto i kódování chování jako takové je dražší metodou než kognitivní rozhovory [Tourangeau, Rips, Rasinski 2000: 332]. Kromě toho je výhodou kognitivních interview možnost klást detailní sondy nebo zapojit think-alouds, pokud respondent sám neposkytne informace, které jsou pro výzkumníka důležité. Zatímco behavior coding tak může zachytit pouze zjevné komplikace při dotazování, především problémy týkající se tazatele [Tourangeau, Rips, Rasinski 2000: 332], kognitivní interview mohou jít dál, což je pro tuto práci podstatné.

Je však důležité mít na paměti, že i kognitivní rozhovory mají svá omezení – mohou sice pomoci odhalit problémy spojené s různými otázkami, to ale neznamená to samé jako najít jejich řešení: „Kognitivní interview mohou poskytnout přesnou diagnózu, diagnóza ale ne vždy ukazuje na lék.“ [Tourangeau, Rips, Rasinski 2000: 334; překlad autorka] I z tohoto důvodu je přínosné spojení dvou různých metod – otázky jsou prozkoumány z více hledisek a je vyšší šance nalezení lepšího řešení.

Kromě jednoduchého split ballot experimentu existují i další experimentální designy, které by bylo možné použít. Namísto úplné randomizace (každá jednotka je náhodně přiřazena do jedné ze skupin a testuje právě jednu variantu), je možné zvolit odlišný postup, a totiž aplikovat znáhodnění na „bloky“ respondentů – každý blok následně testuje všechny varianty, případně více variant otázky<sup>9</sup>. Split ballot je sice méně efektivním řešením, protože od každého respondenta získáme pouze jedno měření, má však výhodu v tom, že respondenta neovlivňuje opakované odpovídání na stejnou otázku. [Tourangeau, 2004: 220 - 221]

### 2.3.1 Validita a reliabilita

Kognitivní interview mají vysokou zjevnou validitu – je obecně přijímáno, že přispívají k lepším otázkám. Bývají však spojována s nízkou reliabilitou, protože tazatelé/výzkumníci k nim přistupují různě, metoda je velmi málo standardizovaná. [Tourangeau, Rips, Rasinski 2000: 333; Conrad, Blair 1996: 1] Částečným řešením

---

<sup>9</sup> Testování více variant otázky na stejných respondentech zahrnuje také robustní multitrait-multimethod (MTMM) design nebo jeho novější adaptace SB-MTMM (split ballot multitrait-multimethod design). [Saris, Satorra, Coenders 2004]

může být použití některého z dostupných kódovacích schémat. [Tourangeau, Rips, Rasinski 2000: 333]

Split ballot experiment má díky randomizaci zaručenou vysokou interní validitu. [Kirk 2009: 27; Fienberg, Tanur 1989: 1018] Externí validita může být díky širokému využití NPS ve výzkumu trhu posouzena porovnáním dat s jiným zdrojem (např. s číslem uvedeným v tiskové zprávě).<sup>10</sup> Odhad reliability metody však není díky designu umožněn (nedochází k opakovanému měření).

## 2.4 Příklad postoje: Net Promoter Score

Jak již bylo v průběhu práce několikrát zmíněno, experiment s délkou škály byl proveden na konkrétním příkladu, tzv. *Net Promoter Score* (dále jen NPS). Je tedy třeba tento zvolený příklad měřeného postoje stručně představit.

NPS vyjadřuje ochotu doporučit využívanou společnost, například banku, a je vypočítán na základě odpovědí na jedinou otázku: „Do jaké míry byste doporučil(a) společnost XY svým přátelům a známým?“ Ochota doporučit je zde měřena na jedenáctibodové škále od 0 „určitě nedoporučil/a“ do 10 „určitě doporučil/a“. Ostatní hodnoty respondentovi slouží k odstupňování postoje a jsou označeny pouze číselně, nikoliv verbálně. Podle odpovědí jsou respondenti rozřazeni do tří skupin: promoters (10 + 9), passives<sup>11</sup> (8 + 7) a detractors (6 – 0).<sup>12</sup> [Reichheld 2003: 1] NPS je rozdíl promoters a detractors a vyjadřuje tedy jakýsi „čistý“ podíl klientů, kteří by danou společnost doporučili. Jeho hodnoty se mohou pohybovat v rozmezí od -100 do +100.

NPS byl představen v roce 2003 Frederikem F. Reichheldem v článku „The One Number You Need to Grow“ v *Harvard Business Review* v odpověď na tradiční komplexní metriky k měření loajality a spokojenosti, které dle něho často přinášejí nejednoznačné výsledky, na základě kterých je složité jednat, a příliš nekorelují

<sup>10</sup> Výzkumná agentura IPSOS v tiskové zprávě z 16. května 2013 uvádí, že NPS (měřené v dubnu na tradiční jedenáctibodové škále) je v průměru pro banky na úrovni -2, pro pojišťovny -5 a pro obchody s potravinami +2. [IPSOS 2013] V rámci této práce byly na jedenáctibodové škále o půl roku později stejnou metodou a na vzorku vybraném ze stejné populace naměřeny tyto hodnoty: banka +11, pojišťovna -1, obchod +17, pomyslné pořadí typů institucí je tedy stejné jako v tiskové zprávě, rozdíl mezi nimi jsou ale vyšší. Naměřené hodnoty se od těch z jara 2013 u banky a obchodu významně liší, u pojišťovny nikoliv (dle jednovýběrového t-testu na 95% hladině významnosti – viz příloha č. 8 na CD). Není však jasné, zda část rozdílů nelze připsat posunu v čase, novější data pro porovnání bohužel nejsou k dispozici.

<sup>11</sup> V původním článku z roku 2003 nazývá autor tuto skupinu „passively satisfied“. Více se ale ujal později (v knize *The Ultimate Question* z roku 2006) uvedený název „passives“.

<sup>12</sup> V původním znění otázka zní: „How likely is it that you would recommend company XXX to a friend or colleague?“ Extrémní body na škále jsou popsány „Not at all likely“ (0) a „Extremely likely“ (10).

s reálným chováním zákazníků, zisky společnosti a s jejím růstem. V kontrastu s nimi je NPS velmi jednoduchý, založený na jedné otázce. [Reichheld 2003] Právě pro svou nenáročnost a snadnou interpretovatelnost se stal v prostředí výzkumu trhu velmi oblíbeným a často využívaným. Co přesně ale má NPS měřit? Jak Reichheld ochotu doporučit určitou společnost chápe?

NPS se zaměřuje primárně na zákazníky, kteří jsou dostatečně nadšení, potěšení, aby danou společnost doporučili (promoters). Toto ochotu doporučit chápe Reichheld jako jeden z nejlepších indikátorů jejich loajality k této společnosti, protože doporučením dávají lidé v sázku svou reputaci; takové riziko jsou ochotni podstoupit, pouze pokud jsou velmi loajální. Loajalita zákazníků a konkrétně jejich ochota společnost doporučit podle Reichhelda pohání růst společnosti. Promoters se tak stávají „marketingovým oddělením“ společnosti. Na rozdíl od měření retence zákazníků, kdy je v podstatě sledováno to, jak společnost o zákazníky přichází, se tedy NPS zaměřuje na tendenci stávajících zákazníků přivádět společnosti nové zákazníky. [Reichheld 2003: 3 - 4] Důležité je, že NPS dle Reichhelda koreluje s reálným jednáním zákazníků. [Reichheld 2003] NPS je tedy příkladem postoje, jehož definice obsahuje spojitost s jednáním, ať již jde o ochotu společnost doporučit nebo nákupní chování zákazníka. To se pak promítá také do charakteristiky tří skupin zákazníků definovaných odpověďmi na otázku: „*Promoters* (...) jsou loajální nadšenci, kteří u společnosti stále nakupují a vyzývají své přátele, aby dělali to samé. *Passives* jsou spokojení, ale nenadšení zákazníci, kteří mohou být snadno zlákáni konkurencí. A *detractors* jsou nespokojení zákazníci chycení ve špatném vztahu.“<sup>13</sup> [Reichheld 2006: 19; překlad autorka, kurzíva v originále]

---

<sup>13</sup> Validizace metriky probíhala prostřednictvím propojení výzkumných dat s tvrdými daty o jejich nákupním chování. Reichheld s kolegy testoval zhruba 20 otázek na zákaznících 14 společností napříč 6 sektory a zjišťoval, která z nich nejsilněji koreluje s opakovanými nákupy. Ve většině společností se ukázala jako první nebo druhá nejlepší právě otázka o doporučení. Pravidlo pro rozlišení promoters, passives a detractors bylo odvozeno rovněž během této studie – prozkoumáním nákupního chování dle odpovědí na hodnotící škále od 0 do 10 našel Reichheld s kolegy tři logické clusterly. Posledním krokem byl tracking NPS u více než 400 společností z více než tuctu sektorů, aby bylo možné říci, zda NPS vysvětluje růst konkurentů v rámci sektoru. [Reichheld 2003: 5 - 6] Reichheldův článek rozpoutal velké diskuze a skupina marketingových výzkumníků se pokusila tuto validizační studii replikovat. Jeho závěry o tom, že NPS je nejlepším prediktorem růstu společnosti, však nepodpořila. [Keiningham et al. 2007]

## 2.5 Research design

Po představení využitích metod lze přejít k research designu samotného výzkumu. Jak již bylo nastíněno, proběhl ve dvou fázích – nejdříve byl proveden split ballot experiment, následně série kognitivních interview s některými z jeho účastníků.

### 2.5.1 První fáze: split ballot experiment

Kvantitativní část testování byla provedena ve spolupráci s výzkumnou agenturou IPSOS v České republice – prostřednictvím jejího online access panelu<sup>14</sup>. Experiment byl zařazen do jednoho z omnibusových šetření agentury, tedy do pravidelného výzkumu sdružujícího více výzkumných témat. [Krejčí 2008: 128] Sada otázek (resp. obě její varianty), které byly využity pro tuto práci, byla umístěna na samý začátek dotazníku, aby bylo předejito efektům jako je únava respondentů, kontext předchozích otázek apod. Agentura měřila Net Promoter Score prostřednictvím této studie již v minulosti, otázky byly tedy pouze přesunuty na začátek dotazníku a byla zařazena také druhá varianta hodnotící škály (pětibodová). Sledované typy institucí tak vycházely z širšího výzkumného záměru agentury – respondenti byli dotázáni, do jaké míry by doporučili svou hlavní banku (tj. nejčastěji využívanou, kde respondent provádí nejvíce transakcí), hlavní pojišťovnu (kde mají nejvíce pojistných smluv) a obchod s potravinami, kde měsíčně utratí nejvíce peněz. Nejdříve bylo zjištěno, které společnosti jsou respondentovou hlavní bankou, pojišťovnou a v jakém řetězci nejčastěji nakupuje potraviny. Dále mu byla položena otázka „*Do jaké míry byste doporučil/a svým přátelům a známým následující společnost?*“ s postupným výčtem v předchozím kroku vybraných společností, přičemž pořadí banky, pojišťovny a obchodu napříč respondenty rotovalo, aby nebylo žádné z hodnocení zatíženo pořadím. Jedna experimentální skupina odpovídala na škále od 0 do 10, druhá na škále od 1 do 5. V obou případech byly olabelovány pouze krajní hodnoty (0/1 jako „určitě nedoporučil/a“, 10/5 jako „určitě doporučil/a“) a respondentovi nebyla nabídnuta žádná „úniková“ možnost (nevím, nechci uvést, nedovedu ohodnotit apod.). Dotazník je k nahlédnutí v příloze č. 1, v naprogramované podobě jsou potom ukázky obou variant škály ukázány v příloze č. 2.

<sup>14</sup> Řada výzkumných organizací z komerční i nekomerční sféry pre-rekrutuje respondenty pro účast v online studiích prostřednictvím tzv. online access panelů. Jejich registrovaní účastníci se pak příležitostně účastní výzkumů dané organizace. [Görizt 2004: 411]

### 2.5.1.1 Cílová populace a výběr

Pravidelné omnibusové šetření, jehož součástí byl provedený experiment, je zacíleno na obecnou populaci České republiky využívající internet. Výběrový soubor o celkové velikosti  $N = 1\ 070$  byl vybrán pomocí kvótního výběru z respondentů registrovaných v online access panelu výzkumné agentury<sup>15</sup>. Kvótní zadání bylo sestaveno na základě informací o složení české populace zveřejněných Českým statistickým úřadem tak, aby výběr reprezentoval obecnou českou populaci; kvótními znaky byly pohlaví, věk, kraj a velikost místa bydliště (počet obyvatel). Přesto je třeba si uvědomit, že vzhledem ke zvolené metodě sběru dat vzorek zahrnuje pouze obyvatele využívající internet. Pro internetovou populaci ČR však není k dispozici uspokojivá opora výběru.

S kvótním výběrem bývá rovněž spojováno riziko nižší přesnosti měření, proto je tzv. „dobrým standardem“ (uznávaným doporučením pro dobrý výzkum) aplikace některé z pravděpodobnostních výběrových metod, které jsou však nákladnější a časově náročnější. Jak ale ukazuje Jindřich Krejčí, není zcela zřejmé, že pravděpodobnostní techniky výběru zaručují vyšší přesnost měření.<sup>16</sup> Mimoto se čeští výzkumníci potýkají s problémem nedostatečných opor pravděpodobnostních výběrů. [Krejčí 2008] Volba kvótního výběru má tedy své opodstatnění.

Navíc, jelikož účelem provedeného experimentu je porovnání distribucí odpovědí u jednotlivých variant odpovědní škály, tedy porovnání experimentálních skupin spíše než získání odpovědí reprezentativních pro nějakou cílovou skupinu (nebo pretest nástroje pro tuto skupinu určenou), podstatné je především zajištění vzájemné srovnatelnosti experimentálních skupin. Toho bylo dosaženo randomizací, tj. náhodným zařazením jednotek do skupin.

<sup>15</sup> V IPSOS online access panelu je cca 25 000 aktivních respondentů (tj. těch, kteří jsou nejen registrovaní, ale také o sobě pravidelně poskytují aktuální informace v osobním dotazníku a účastní se výzkumů. Více viz <https://www.instantresearch.cz/info/o-panelu/>

<sup>16</sup> Jindřich Krejčí na porovnání výsledků z povolebních výzkumů (respektive otázek věnujících se této problematice) založených na kvótním (CVVM červenec a listopad 2002) i pravděpodobnostním výběru (ISSP září/říjen 2002, ESS listopad/prosinec 2002) s reálnou podporou stran ve volbách do Poslanecké sněmovny v roce 2002 ukazuje, že „samotná výběrová metoda nezaručuje přesnější výsledky.“ Šetření ESS a ISSP, využívající pravděpodobnostní metody výběru, byla realizována výrazně draž a vykazovala nižší přesnost měření než výzkum CVVM. [Krejčí 2008: 156 - 158]

### **2.5.1.2 Randomizace**

Vybraným respondentům byla přiřazena náhodně generovaná čísla pomocí funkce v MS Excel; následně byl výběr rozdělen na dvě skupiny podle toho, zda bylo číslo liché nebo sudé. Byla tak splněna základní podmínka experimentálního designu.

Výsledkem procesu randomizace jsou dvě skupiny o přibližně stejné velikosti – variantu dotazníku s pětibodovou škálou vyplnilo 534 respondentů, s jedenáctibodovou 536 respondentů. Skupiny se mezi sebou na 95% hladině významnosti významně neliší dle pohlaví, věku, vzdělání, osobního příjmu, kraje ani velikosti místa bydliště (dle chí kvadrát testu/t-testu/Mann-Whitney testu dle typu proměnné – viz příloha č. 8 na CD).

### **2.5.1.3 Sběr dat**

Sběr dat proběhl v termínu od 18. do 24. září 2013 prostřednictvím IPSOS online access panelu, metodou sběru dat tedy bylo CASI (Computer-Assisted Self Interviewing - samovyplňovací elektronický dotazník). Elektronický dotazník umožnil jednoduchou aplikaci experimentu (náhodné rozdělení jednotek do experimentálních skupin) a díky jeho samovyplňovací povaze nejsou odpovědi ovlivněny tazatelem a snaha respondentů odpovídat sociálně žádoucím způsobem (tj. overreporting, underreporting) je tedy oproti dotazování tazatelem nižší. [Tourangeau, Rips, Rasinski 2000: 295 - 298]

Respondenti byli - stejně jako v případě jiných dotazníkových šetření realizovaných agenturou skrze online access panel – za vyplnění dotazníku odměněni finančními incentivy.

## **2.5.2 Druhá fáze: kognitivní rozhovory**

Druhá fáze výzkumu proběhla již bez spolupráce s výzkumnou agenturou, všechny rozhovory realizovala sama autorka, pouze s vědomím agentury, že oslovila respondenty registrované v jejím online access panelu.

### **2.5.2.1 Cílová populace a výběr**

Kognitivní rozhovory proběhly s účastníky první fáze výzkumu, kteří byli schopni ohodnotit všechny tři sledované typy institucí (tj. jmenovali svou hlavní banku,

pojišťovnu a obchod s potravinami) a souhlasili s účastí<sup>17</sup> na následném rozhovoru. Vzhledem k časovým a finančním omezením autorky byla druhá fáze výzkumu omezena pouze na obyvatele Prahy.

Jelikož kvantitativní data ukazují v případě obou experimentálních skupin na homogenitu odpovědí (nebyly nalezeny systematické rozdíly v mediánech dle socio-demografických charakteristik respondentů, tj. že by některá z podskupin volila pozitivnější/negativnější hodnocení<sup>18</sup>), nebyly socio-demografické charakteristiky zohledněny při vzorkování. Respondenti byli vybráni 1) na základě sledované charakteristiky (vzorců odpovídání na dané škále), 2) z obou experimentálních skupin. Autorka oslovila respondenty s co nejrůznějšími vzorci odpovídání, ať již z hlediska příklonu k vyššímu/nížšímu ohodnocení, tak co se týče variability odpovědí mezi třemi typy institucí.<sup>19</sup>

Ve chvíli, kdy bylo provedeno 8 rozhovorů (4 s respondenty, kteří odpovídali na pětibodové škále, 4 s respondenty, kteří odpovídali na jedenáctibodové škále), autorka sběr dat ukončila, protože nebyly generovány další informace.

### 2.5.2.2 Sběr dat

Rozhovory proběhly v období od 7. října do 6. listopadu 2013, sběr dat tedy začal s dvoutýdenním zpožděním oproti kvantitativní fázi. Toto časové období bylo nutné věnovat průzkumu kvantitativních dat a oslovení vybraných respondentů pro kvalitativní fázi.

Respondenti byli osloveni telefonicky, vždy v preferované době, kterou uvedli v dotazníku. Autorka telefonicky objasnila účel výzkumu, tedy že se jedná o výzkum k diplomové práci, nikoliv o aktivitu agentury IPSOS. Respondentům byla slíbena nefinanční incentiva<sup>20</sup> a byl domluven termín a místo setkání.

<sup>17</sup> Po vyplnění sady otázek na doporučení byli respondenti požádáni o účast na další vlně výzkumu a telefonní číslo společně s časovým rozmezím, kdy si přejí být kontaktováni – viz otázky X1\_1 a X1\_2 v příloze č. 1.

<sup>18</sup> Až na jednu výjimku se na 95% hladině významnosti mediány významně nelišily dle pohlaví, věku, vzdělání, osobního příjmu, kraje a velikosti místa bydliště (využití Kruskal Wallis testu/Mann-Whitney testu v závislosti na počtu podskupin). Výjimkou byla otázka na ochotu doporučit obchod, která se na jedenáctibodové škále lišila dle pohlaví (viz příloha č. 8 na CD).

<sup>19</sup> Kognitivních rozhovorů se tak zúčastnili tři muži a pět žen, jednalo se o osoby ve věku 23, 27, 32, 34, 39, 47, 50 a 51 let. Dvě z těchto osob byly vyučené, jedna absolvovala střední školu s maturitou, jedna nástavbové studium a čtyři magisterské studium. Z hlediska čistého měsíčního osobního příjmu šlo o tři osoby s příjmem nižším než 8 tis. Kč, jednu osobu s příjmem 10 – 15 tis. Kč, dvě osoby s příjmem 20 – 25 tis. Kč, jednu osobu s příjmem 30 – 40 tis. Kč a jednu s příjmem 40 – 50 tis. Kč.

<sup>20</sup> Na každé setkání autorka přinesla tři různé dárky, respondent si vždy vybral ten, který se mu nejvíce zamlouval. Jednalo se vždy o dárky v hodnotě 80 – 100 Kč (čokoláda, sada čajů, láhev vína apod.).

Rozhovory probíhaly osobně, a to buď v domácnosti respondenta nebo v kavárně či restauraci. Všechny byly za souhlasu respondenta nahrávány a následně přepsány pro účely analýzy (přepisy rozhovorů tvoří přílohu č. 9 na CD); sběr dat a jejich analýza tedy byly časově oddělené. Délka rozhovorů se pohybovala okolo 25 minut, celkově setkání zabralo kolem 40 minut – autorka nejdříve představila sebe i výzkum a informovala respondenta o průběhu rozhovoru i o dalším nakládání s daty, až následně bylo přikročeno k nahrávanému rozhovoru.

### **2.5.2.3 Scénář rozhovoru a další využívané materiály**

Scénář byl na základě prvních dvou rozhovorů mírně doplněn (o otevřenou otázku na ochotu doporučit), jeho aktualizovaná verze tvoří přílohu č. 3.

Na úvod rozhovoru byl respondent požádán, aby odpověděl znovu na tři otázky týkající se doporučení (nejprve vlastními slovy, následně s využitím hodnotící škály stejné délky jako v první fázi výzkumu) a aby při odpovídání na otázky přemýšlel nahlas (technika think-alouds). Zbytek interview byl již založen především na verbal probes. Vzhledem k tomu, že dotazník byl velmi krátký, autorka přistoupila k retrospektivnímu probingu, tzn. nejdříve prošla s respondentem otázky (což bylo důležité i kvůli osvěžení tématu vzhledem k časovému rozmezí mezi oběma fázemi sběru dat) a následně pokládala sondující otázky. Šlo zejména o kognitivní, specifikující, rozšiřující sondy, ale i sondy týkající se interpretace otázky a posouzení jistoty odpovědi.

Při rozhovoru autorka ukazovala respondentovi dvě vytištěné verze dotazníku – nejprve verzi s délkou škály, na kterou odpovídal, ve druhé části rozhovoru také druhou variantu. Oba dokumenty tvoří přílohu č. 4.

## **2.6 Etické souvislosti výzkumu**

První fáze výzkumu proběhla ve spolupráci s výzkumnou agenturou IPSOS, členem sdružení agentur pro výzkum trhu a veřejného mínění SIMAR.<sup>21</sup> SIMAR přejímá mezinárodní etické kodexy sdružení ESOMAR. [SIMAR 2001] Realizátor sběru dat první fáze tohoto výzkumu je tedy vázán sdílenými standardy na české i mezinárodní úrovni. Provedení experimentu v rámci pravidelného omnibusového

---

<sup>21</sup> Pro více informací o SIMAR viz [simar.cz](http://simar.cz).



šetření bylo autorce laskavě povoleno vedením agentury, písemné svolení k využití dat pro odbornou publikaci je k nahlédnutí v příloze č. 5.

Ve druhé fázi výzkumu byli respondenti autorkou informováni o účelu výzkumu, o nahrávání rozhovoru a dalším nakládání s daty včetně možné anonymizované citace jejich slov v diplomové práci či odborném článku. Autorka s každým z respondentů podepsala informovaný souhlas (viz příloha č. 6) ve dvou vyhotoveních, z nichž jedno si nechala a druhé zůstalo respondentovi. Autorka je zaměstnankyní výzkumné agentury IPSOS, bylo proto důležité respondenty zároveň ujistit, že výzkum je realizován pro účely diplomové práce autorky a bude tedy publikován, že se však nejedná o aktivitu výzkumné agentury. Žádný z respondentů nebyl zaměstnáním autorky znepokojen a všichni souhlasili s rozhovorem za těchto podmínek.

### 3. Analýza dat

V následujících kapitolách jsou představeny výsledky výzkumu, pozornost je přitom věnována oběma jeho částem – split ballot experimentu i kognitivním rozhovorům. Praktická část této práce je organizována dle představeného kognitivního modelu; autorka však společně s Conradem a Blairem [Conrad, Blair 1996] věří, že procesy vyvolání informace a následného úsudku nelze v primárních datech spolehlivě rozlišit. Obě tyto složky jsou proto zpracovány společně jako provedení úlohy. Analýza procesů porozumění a provedení úlohy je postavena na kvalitativních datech z kognitivních rozhovorů; z hlediska délky škály a tedy i této práce je však nejpodstatnější proces editace odpovědi, kde již autorka kombinuje zjištění z kvalitativní i kvantitativní fáze výzkumu.

Na základě studia literatury byly formulovány a testovány tyto hypotézy:

- **H<sub>1</sub>:** Jedenáctibodová a pětibodová škála přináší různé výsledky. [v kontrastu s Dawes 2002: 2; Dawes 2008: 9 - 10]
- **H<sub>2</sub>:** Na pětibodové škále je střední hodnota zastoupena více než na jedenáctibodové škále. [viz Groves 2004: 465]
- **H<sub>3</sub>:** Na jedenáctibodové škále jsou extrémní hodnoty zastoupeny méně než na pětibodové škále. [podobně jako Tourangeau, Rips, Rasinski 2000: 244]

- **H<sub>4</sub>:** Hodnotě 0 na jedenáctibodové škále se respondenti vyhýbají, jelikož má v numerickém systému své speciální místo. [viz Tourangeau, Rips, Rasinski 2000: 245]
- **H<sub>5</sub>:** Jedenáctibodová škála znamená pro respondenty vyšší kognitivní zátěž - dochází zde ke komplikacím s rozlišením sousedících hodnot. [podobně jako Tourangeau, Rips, Rasinski 2000: 249]

### 3.1 Porozumění

V rámci procesu porozumění je pro tuto práci klíčová *interpretace* otázky „Do jaké míry byste doporučil/a svým přátelům a známým následující společnost?“ spíše než porozumění jednotlivým slovům i celé otázce. Na základě kognitivních rozhovorů je patrné, že základní pochopení otázky respondentům potíže nečiní, v čem se však rozchází, jsou její interpretace. Někteří si představí reálnou situaci doporučení (na dotázání někoho blízkého), jiní považují dotaz na doporučení spíše za slovní obrat, ale hodnotí svou spokojenost nebo vztah k dané instituci.

„No jako pokud se mě někdo z mejh přátel jakoby konkrétně... jako já bych asi nechodila a neříkala ‚jdi do BANKY<sup>22</sup>‘ nebo do nějaký jiný... POJIŠŤOVNY... ale jako když by se mě zeptali, asi bych odpověděla stejně, jak jsem odpověděla tady. Když se mě někdo zeptá ‚chci jít k BANCE, myslíš, že tam mám jít? Máš s tím nějaký zkušenosti?‘, tak to ano, já s nima nemám žádnéj problém, je to v pohodě.“

„Vy vlastně se mě ptáte, jak jsem s tou bankou spokojený. To, jestli bych jí doporučil, to už je jenom takovej opis. Vy se mě ptáte ‚tady máte banku, jste v ní, jak jste s ní spokojený, co se Vám líbí, nelíbí, porovnejte, dejte číslo.‘“

Pokud respondent chápe otázku jako dotaz na spokojenost či vztah k dané instituci, je pro něj úloha o něco více přímočará - vyvolá relevantní informace k vyjádření tohoto postoje, tedy dosavadní klientské zkušenosti, výhody a nevýhody spolupráce apod.

„Je to v zásadě pozitivní odpověď, já jsem tam spokojený, mám slušný internetový bankovníctví, jsou pobočky, který maj otevířno, je spousta bankomatů, to jsem spokojený. Ale třeba se mi nelíbí to, že... i když to dělá spousta bankovních domů, spíš těch starších... to, že když u nich mám hypotéku, tak mám úplně jiný úroky jako stávající klient, než co nabízej tým novým. Taky jsem tam posoudil ty poplatky, které tam mám a nejsou úplně malé, ale bohužel ten účet tam musím mít, protože mám u nich hypotéku. Takže nemůžu odejít třeba k jiné bance. A tohle, tohleto dohromady, to mi dalo to číslo.“

---

<sup>22</sup> Jelikož cílem této práce není ohodnotit jednotlivé společnosti existující na českém trhu, byly jejich názvy v prepisech nahrazeny univerzálním označením BANKA, POJŠŤOVNA a OBCHOD.

Odlišná situace však nastává v případě interpretace otázky doslovně, tedy při hodnocení vlastní ochoty doporučit instituci známému. Úvahy respondenta jsou komplexnější; jeho vlastní spokojenost či pozitivní vztah k instituci jsou nutným předpokladem, nikoliv však zárukou jejího doporučení. Navíc jsou do výsledného hodnocení promítány dodatečné informace související s aktuálními podmínkami na trhu nebo například s tím, komu konkrétně by měla být instituce doporučena, co tato osoba hledá. Výsledné hodnocení tak může být kombinací spokojenosti, ale zdráhání se z určitých důvodů doporučit instituci někomu jinému.

„Před šesti lety, když jsme si zakládali to pojištění, tak to bylo v tu dobu jakoby to nej, pro nás nejvýhodnější, ale ona takovou nabídku v tuto chvíli už ani nedává, nemá, takže bych jí ani nedoporučovala. Zmínila bych jí ‚mám jí, mám jí s touhleto POJISTKOU máme, jsem spokojená, máme tam jakoby možnost nahoru – dolů, ale je to stará pojistka, takže v tuto chvíli bych jí nedoporučila.‘“

„Tam je to spíš než, že by tam bylo špatně jako u té banky a pojišťovny, ale spíš je tam ten parametr dosažitelnost a přátelé nejsou jenom po Praze, ale třeba po Jižních Čechách a podobně a já nevím, jak to tam maj daleko, tam jde spíš vo dojezd.“

Data tedy naznačují, že NPS nefunguje spolehlivě jako nástroj k měření jediného postojů – respondenti otázku chápou a zodpovídají různě.<sup>23</sup> Interpretaci dalších výsledků tak může tento fakt komplikovat.

### **3.2 Provedení úlohy: vyvolání informací a úsudek**

Na pomezí porozumění a procesů souvisejících se samotným provedením úlohy stojí další zjištění. Chápání otázky respondentem může ovlivňovat také set nabízených odpovědí – v případě experimentování s délkou hodnotící škály může počet nabízených odpovědí respondentovi napovědět, jak moc přesnou odpověď od něj výzkumník očekává. S různou délkou škály pak respondenti vyvolávají různé množství informací potřebných k úsudku.

<sup>23</sup> Z Reicheldovy konceptualizace není jednoznačné, která z těchto dvou interpretací je správná (pokud některá), tedy zda má být NPS indikátorem loajality (případně spokojenosti, jak se zdá na základě zde provedeného výzkumu) nebo spíše zjišťovat, zda zákazníci společnost skutečně doporučují; vzhledem k důrazu na korelaci s reálným chováním (doporučování, opakované nákupy) však jde zřejmě spíše o tu druhou, doslovnou, týkající se reálné situace doporučení. U respondentů, kteří chápou otázku takto, může být pravděpodobnost reálného doporučení (tj. spojitost postojů s jednáním) vyšší, ovšem i zde jde pouze o spekulaci, zde prezentovaný výzkum se na tento aspekt nezaměřuje.

### 3.2.1 Množství vyvolaných informací

Při kognitivních rozhovorech byla respondentům ukázána vždy nejdříve škála, na které odpovídali v první, kvantitativní fázi výzkumu; byli požádáni o opětovné ohodnocení institucí a zejména pomocí sondovacích otázek<sup>24</sup> byly zjišťovány důvody tohoto hodnocení, vybavené informace.<sup>25</sup> V průběhu kvalitativní fáze výzkumu autorka do scénáře přidala také úvodní otevřenou otázku na doporučení, v některých případech je tedy k dispozici porovnání tří „stupňů“ odpovídání: 1. na otevřenou otázku, 2. pomocí pětibodové škály a 3. pomocí jedenáctibodové škály.<sup>26</sup> S každým stupněm uváděli respondenti více důvodů, vybavovali si více informací.

**1. otevřená otázka:** „Tak určitě bych doporučil v rámci... Když já jsem jako student, když by můj kamarád třeba byl také student, tak určitě bych doporučil, protože vedení vlastně běžného účtu je zadarmo, výběr z bankomatů také.“

**2. pětibodová škála (hodnocení 4):** „Prostě jako asi většina bank prostě mají nějaké poplatky a většinou, když člověk se tam chce třeba na něčem domluvit, že potřebuje přímo se zeptat na pobočce na něco, tak většinou čeká dlouhý čas, tam stráví.“

**3. jedenáctibodová škála (hodnocení 7):** „Tak právě že asi jsem málo informován od nich přímo, že většinou se musím sám někde dotazovat nebo podívat se na stránky a zjišťovat si tam informace, abych se je dozvěděl.“

**1. pětibodová škála (hodnocení 5):** „BANKA, POJIŠŤOVNA, obchod tuším OBCHOD... A všechny bych určitě doporučovala. (...) Protože jsem skutečně spokojená. A kdybych nebyla spokojená, tak budu přemýšlet o změně.“

**2. jedenáctibodová škála (hodnocení 7):** „No to bych spíš dala tu devítku. (...) Nebo možná ten OBCHOD už bych nedala ani tu devítku, když už je to takhle konkrétní, protože oproti OBCHODU2 třeba nemaj tu zavázkovou službu, kterou já už bych chtěla vod toho OBCHODU2.“

Podrobnější výčet odpovědí (delší hodnotící škála) nutí respondenty více přemýšlet, někteří sami uvádí, že musí zvážit více informací, aby byli schopni vybrat

<sup>24</sup> Ač by k tomuto účelu byla vhodnější technika think-alouds, ukázala se pro respondenty jako příliš náročná.

<sup>25</sup> Co se týče typu vyvolávaných informací, obvykle je u ochoty doporučit zvažována vlastní spokojenost s institucí, výhody a nevýhody spolupráce, konkrétní zkušenosti a dále porovnání s konkurencí, ať již na základě vlastní zkušenosti (například s nákupy v jiných potravinových řetězcích) nebo dalších informací, získaných například od rodiny či přátel. Porovnání s konkurencí hraje větší roli v případě, že respondent interpretuje otázku jako dotaz na reálné doporučení, není ale vyloučeno ani v případě hodnocení spokojenosti, vztahu k instituci.

<sup>26</sup> Respektive 1. na otevřenou otázku, 2. pomocí jedenáctibodové škály a 3. pomocí pětibodové škály. V tomto pořadí však nelze vzorec odhalit, protože již při odpovídání na jedenáctibodové škále si respondent vybaví velké množství informací a u následně pětibodové škály pak nelze spolehlivě ohodnotit, zda využil stejné množství informací nebo pouze část z nich. Je tak možné porovnat pouze rozdíl mezi procesy vyvolanými otevřenou otázkou a jedenáctibodovou škálou. I při tomto uspořádání kognitivního rozhovoru však respondenti sami uváděli, že při hodnocení na delší škále musí víc přemýšlet.

příslušnou odpověď. Oproti tomu u krátké škály se jedná spíše o rychlé zařazení do kategorie.<sup>27</sup>

„No, nemusím nad tím tolik jako přemejšlet (*pětibodová škála*), protože přeci jenom tady (*jedenáctibodová škála*), těch víc bodů je možnost, že tam je... mezi pětkou a desítkou, že tam je víc... takový jako jestli jsem spokojená hodně, málo...“

„Je to takový rychlý (*pětibodová škála*), ale je to takový zařazení jenom do takového chlívěčku...“

„Tak stejná otázka, ale desetistupňová škála. No tak pro mě zase by ten průměr byl ta pětka, takže já bych to vyhodnocovala stejným způsobem s tím, že tady už bych víc jako ty nuance zvažovala. Tady celkem jako mám jenom ty dva stupně (*pětibodová škála – dva stupně intenzity*), takže bych to vzala v takovech těch balíčcích jako. Plusy a mínusy že jo, dva balíčky. A ten balíček, kterej by byl větší/menší, tak podle toho bych dala třeba tu dvojku nebo jedničku, rozumíte. Tady už samozřejmě mně to dá víc přemýšlení, je to konkrétnější, je to přesnější a i já o tom musím víc přemejšlet, jestli třeba i v tom balíku těch negativ, že jich tam sice budu mít hodně, ale jestli už ta negativa všechna maj stejnou váhu, třeba jestli jsem přišla o peníze anebo jenom že mně to neposlali ten den, ale poslali to druhý den. Tak samozřejmě i ty mínusy každéj má jinou jako váhu, tak bych se snažila to náh tak vyhodnotit. A zase stejným způsobem, tak jak by to pro mě bylo závažný třeba v tom mínusu, tak bych šla tou škálou dolu.“

„Protože je tady víc bodů a musel bych se opravdu ještě více, hloub zaměřit, co pro mě třeba udělali nebo když já něco potřebuju od nich, jestli opravdu se to dozvim nebo opravdu mě k tomu dopomůžou... Třeba taky už vybírám více bodů, protože se opravdu zamyslím nad tím, co oni mi poskytují nebo když něco ode mě chtějí, tak že mi třeba napíšu sami.“

Délka škály tedy ovlivňuje nejen editaci odpovědi, ale v podstatě celý kognitivní proces odpovídání.

### 3.2.2 Strategie respondentů při provádění úlohy

Respondenti při zpracovávání informací zaujímají velmi rozmanité strategie, lze však nalézt některé společné jmenovatele. Jelikož proces odpovídání trvá jen pár vteřin, záleží na tom, které informace si respondent aktuálně vybaví. Často může být výsledné hodnocení založené na pocitu, dojmu z instituce. Z konkrétních zkušeností se vybavují ty neobvyklé nebo čerstvé. Zejména v případě dlouhé, jedenáctibodové škály mívá proces dvě fáze. Respondent si nejdříve vytvoří na základě vyvolaných informací úsudek, takové hodnocení odpovídá spíše intervalu bodů nebo určité části škály. Následně si hodnocením pozitiv a negativ, aktuálně vybavených zkušeností, pocitu apod. vybere příslušný bod. Procesy vyvolání informací, úsudku a editace odpovědi tedy skutečně nelze označit za tři oddělené fáze, jelikož se vzájemně prolínají.

„To tam vnímám, tyhle ankety já vnímám pocitově jo. Zamyslím se OBCHOD, jaký z toho mám pocitu, vyjde mi z toho třeba ten interval no a pak se zaměřím a jedno

<sup>27</sup> V rámci kvantitativní fáze bohužel nebyl měřen čas strávený odpovídáním na NPS otázky, aby bylo možné testovat předpoklad, že na desetibodové škále odpovídají respondenti déle. Byl měřen pouze celkový čas omnibusového dotazníku, který se mezi experimentálními skupinami nelišil.

vyberu. Může to být třeba toto rozhodnutí už ovlivněný tím, že ráno jsem tam byl a měli tvrdý rohlíky. Druhej den, kdyby je měli měkký, tak bych dal to vyšší číslo. Ale to rozhodnutí v tom intervalu, to už je takové hodně pocitové v ten okamžik.“

„Tak i tak vlastně když si přečtu otázku ‚do jaké míry byste doporučila tadada...‘, tak si řeknu BANKU třeba konkrétně... spíš bych doporučila. A teprve potom se kouknu na tu škálu a rozhoduju se prostě na tý horní škále, že bych doporučila.“

„...a jako já teda, když tam vidim otázku, tak dám na ten jako první dojem, jako nebo prostě na tu mojí první zkušenost. Tou odpovím a teď to teda musím vybírat na tý škále. A jako říkám, buď to je špatný nebo dobrý a třeba v tom OBCHODĚ jsem si říkala tak pade na pade, ale zase jako maj dobrý jako ty produkty, tak jim přihodim bod jo.“

Nelze však tvrdit, že tento postup je vlastní všem respondentům, v rámci kognitivních rozhovorů byly zaznamenány i další strategie individuálního charakteru, bez společných aspektů. Zaměření se na různé strategie by proto bylo zajímavým podnětem pro další výzkum.

### **3.3 Editace odpovědi**

Pro tuto práci je stěžejním kognitivním procesem editace odpovědi, její zakódování do jednoho z bodů na hodnotící škále. V této kapitole je již pozornost věnována také kvantitativním datům pořízeným pomocí split ballot experimentu; všechny statistické testy jsou prováděny na 95% hladině významnosti, jejich výsledky jsou uvedeny v příloze č. 8 na CD a ve většině případů také stručně v textu (případně je uvedena alespoň informace, zda je rozdíl statisticky významný, tj. zda jej lze zobecnit nad rámec výběrového souboru).

#### **3.3.1 Vlastnosti proměnných**

Rozložení proměnných není ani v jednom z případů normální (dle Kolmogorov-Smirnov testu – sig. vždy 0,000), data ale nejsou příliš šikmá ani špičatá. Tabulka č. 1 ukazuje rozdíly mezi proměnnými s pěti a jedenácti hodnotami z hlediska koeficientů šikmosti a špičatosti.

**Tabulka č. 1: Šikmost a špičatost**

Otázka	11bodová škála	5bodová škála	Rozdíl (11bodová – 5bodová škála)
<b>Koeficient šikmosti</b>			
A6_1 Doporučení banky	-0,90	-0,62	-0,28
A6_2 Doporučení pojišťovny	-0,35	-0,46	+0,11
A6_3 Doporučení obchodu	-0,79	-0,59	-0,20
<b>Koeficient špičatosti</b>			
A6_1 Doporučení banky	0,31	-0,12	+0,43
A6_2 Doporučení pojišťovny	-0,71	-0,04	-0,67
A6_3 Doporučení obchodu	0,44	-0,22	+0,66

Jak je vidět, z hlediska tvaru dat se chová odlišně otázka týkající se ochoty doporučit svou pojišťovnu. Zatímco u banky a obchodu generuje jedenáctibodová škála data o něco více negativně zešikmená (tj. seskupená u pozitivního konce škály) a špičatá, u pojišťovny je tomu naopak – data získaná pomocí jedenáctibodové škály jsou méně šikmá a více plochá (jejich koeficient špičatosti je záporný) než u pětibodové škály. Ač jsou koeficienty šikmosti a špičatosti, stejně jako rozdíly mezi proměnnými, poměrně malé, ukazují na mírnou odlišnost otázky na doporučení pojišťovny. Takových specifik má otázka více, jak bude ukázáno dále v této práci.

Co se týče variability, data potvrzují, že více variability lze najít v datech získaných na jedenáctibodové škále. Tabulka č. 2 ukazuje variační koeficienty<sup>28</sup> jednotlivých proměnných.

**Tabulka č. 2: Variabilita**

Otázka	11bodová škála	5bodová škála	Rozdíl (11bodová – 5bodová škála)
<b>Variační koeficient</b>			
A6_1 Doporučení banky	30,87	24,73	+6,14
A6_2 Doporučení pojišťovny	28,48	23,87	+4,61
A6_3 Doporučení obchodu	24,10	21,35	+2,75

Tento výsledek není překvapením, jelikož delší škála nabízí větší množství možných odpovědí; data potvrzují, že mají-li respondenti na výběr větší počet bodů, také jich více využijí. Zatímco na pětibodové škále pro ohodnocení všech tří institucí<sup>29</sup> využili průměrně 1,89 odpovědi, na jedenáctibodové škále to bylo 2,16 odpovědi. Dle

<sup>28</sup> Po vzoru Dawese jsou koeficienty pro větší přehlednost vynásobené stem. [Dawes 2002: 7]

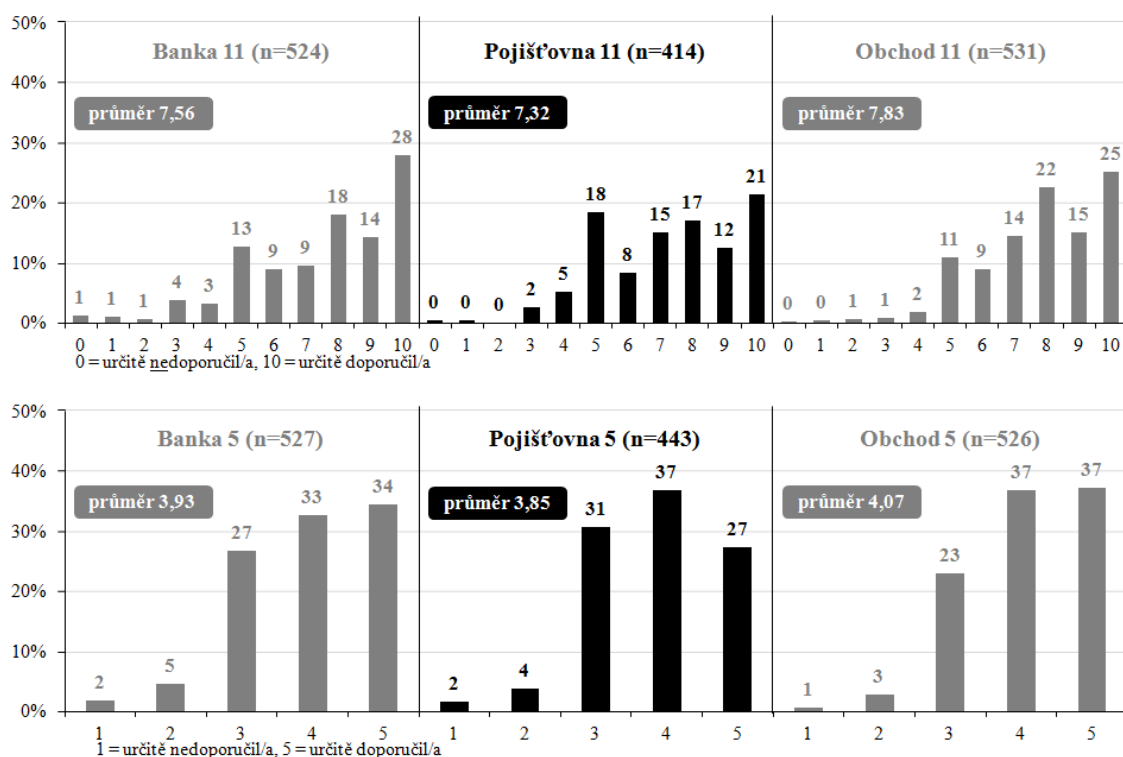
<sup>29</sup> Do výpočtu vstupovaly pouze odpovědi respondentů, kteří hodnotili všechny tři typy institucí.

Mann-Whitneyho testu je rozdíl v počtu použitých odpovědí statisticky významný (sig. 0,000, Z -5,893).

### 3.3.2 Rozložení odpovědí

Jak již bylo nastíněno dříve v této práci, pro porovnání variací testované otázky je klíčové právě to nejjednodušší řešení – porovnání distribucí odpovědí. Nejdříve je zde ukázáno rozložení původních proměnných, následně je přikročeno k různým rekódům a přeškálování proměnných umožňujícím přímé porovnání. Graf č. 1 ukazuje zastoupení jednotlivých hodnot u původních proměnných.

**Graf č. 1: Rozložení odpovědí – původní proměnné (relativní četnosti)**



Jak je vidět z grafu č. 1, zastoupení hodnot je u otázky na ochotu doporučit pojišťovnu odlišné od zbylých dvou typů institucí; v případě jedenáctibodové i pětibodové škály je zde méně zastoupená maximální pozitivní hodnota (10, resp. 5) a více naopak neutrální střední hodnota (5, resp. 3), obě škály se tedy „chovají“ jinak než u banky či obchodu. Důvodem se zdá být méně těsný vztah respondentů právě k pojišťovně – zatímco s bankou a obchodem jsou v kontaktu na denní bázi (vybírají peníze z bankomatu, využívají internetové bankovníctví, nakupují potraviny), u pojišťovny tomu tak není. Situaci navíc komplikuje fakt, že někteří z nich nemají s pojišťovnou plnou zákaznickou zkušenost, protože zatím neměli žádnou pojistnou



událost. Volí tak spíše neutrální ohodnocení (více k vnímání jednotlivých bodů viz dále).

„No protože zatím jsme jenom uzavřeli a nic jako jsme neřešili, takže ehm jenom jsem slyšela nebo nám jí všichni jako doporučili, ale víc jako nemůžu jakoby úplně s ní spokojená, když ještě se mi neprojevila ňák nebo neukázala. (...) No protože... ehm... s těma bankama nebo s tím obchodem jsem stále každé den de facto ve styku nebo tam něco řeším, ale díky tý pojišťovně že prostě tam mně nic nenastalo, jenom se uzavřely věci, tak z tohodle důvodu jako jsem to takhle dala lepší a nižší.“

„Takovej, že prostě nevím, co oni pořádně nabízí, ani že mě ničím nezklamali, že zatím u nich funguji, ale nevím o nich nic pořádně, nějaké výhody, nevýhody. Takový průměr.“ „Jo. A kdyby tam třeba byla možnost ještě že nevíte, tak vybral byste si radši to nevím?“ „To asi ano.“

Přesto jsou některé rozdíly mezi dlouhou a krátkou škálou společné všem typům institucí a potvrzují efekty zdokumentované v literatuře. Jak je vidět z grafu č. 1, u hodnotících škál dochází k *positivity bias* – odpovědi jsou vychýleny směrem k pozitivnímu konci škály, zatímco negativní hodnoty jsou voleny málo [Tourangeau, Rips, Rasinski 2000: 240 - 241]; to platí pro pětibodovou i jedenáctibodovou škálu. Liší se však zastoupení jednotlivých typů hodnot. U krátké škály respondenti výrazně častěji než u dlouhé volí neutrální střední hodnotu, což potvrzuje hypotézu H<sub>2</sub>. Tabulka č. 3 pro přehlednost ukazuje podíly negativních (0 – 4, resp. 1 – 2), neutrálních (5, resp. 3) a pozitivních hodnot (6 – 10, resp. 4 – 5), které se mezi škálami významně liší (dle chí-kvadrát testu – sig. vždy 0,000 - a adjustovaných reziduí) - v případě všech tří typů institucí je na pětibodové škále více zastoupena střední hodnota, zatímco na jedenáctibodové jsou častější pozitivní hodnoty. Zastoupení negativních hodnot je na obou škálách srovnatelné.

**Tabulka č. 3: Zastoupení negativních, neutrálních a pozitivních odpovědí (relativní četnosti)**

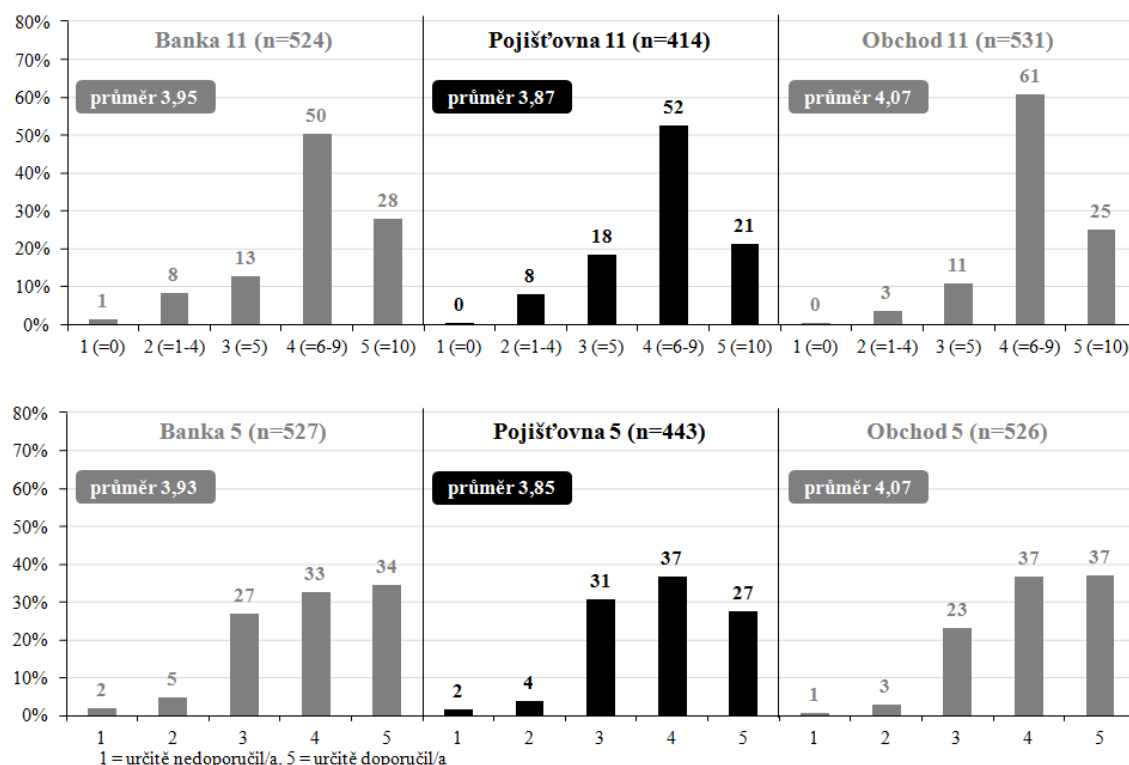
Otázka	Negativní hodnoty		Neutrální hodnoty		Pozitivní hodnoty	
	11bodová škála	5bodová škála	11bodová škála	5bodová škála	11bodová škála	5bodová škála
A6_1 Doporučení banky	9 %	6 %	13 % ↓	27 % ↑	78 % ↑	67 % ↓
A6_2 Doporučení pojišťovny	8 %	5 %	18 % ↓	31 % ↑	74 % ↑	64 % ↓
A6_3 Doporučení obchodu	4 %	3 %	11 % ↓	23 % ↑	86 % ↑	74 % ↓

↑ = signifikantně vyšší četnost (adjustovaná rezíua > + 2)

↓ = signifikantně nižší četnost (adjustovaná rezíua < - 2)

Z grafu č. 1 je dále patrné, že ačkoliv je na jedenáctibodové škále maximální hodnota 10 módem, tedy nejvíce zastoupenou hodnotou, v porovnání s pětibodovou škálou zvolila maximum menší část respondentů; hypotéza  $H_3$  tedy pro pozitivní konec škály platí (ne však pro negativní - zastoupení negativní krajní hodnoty se výrazně neliší). Jinými slovy, jedenáctibodová škála podléhá efektu *response contraction bias*, kdy respondenti nevyužívají tak často pozitivní extrém, ale své hodnocení zmírňují [Tourangeau, Rips, Rasinski 2000: 244], uchylují se spíše k méně silným pozitivním hodnotám 6 – 9. Tento trend ukazuje graf č. 2 zobrazující původní pětibodové proměnné společně s rekódem jedenáctibodové škály, kde jsou ponechány krajní a střední hodnoty samostatně a méně silné hodnoty jsou spojeny (1 – 4 negativní, 6 – 9 pozitivní).

**Graf č. 2: Rekód – spojení méně silných pozitivních a negativních hodnot (relativní četnosti)**



Průměry proměnných v grafu č. 2 jsou téměř shodné, rozložení na rekódované dlouhé škále je však vizuálně špičatější (nejvíce u obchodu – koef. špičatosti 1,40 vs. - 0,22). Zatímco na pětibodové škále je více zastoupena střední (3) a maximální pozitivní (5) hodnota, na rekódované dlouhé škále jsou to naopak méně silné pozitivní hodnoty, tedy část škály zde odpovídající bodu 4. V případě všech tří typů institucí je tento rozdíl statisticky významný (dle chí kvadrát testu – sig. vždy 0,000 - a adjustovaných

reziduí).<sup>30</sup> Sklon volit na dlouhé škále méně silné pozitivní hodnoty a na krátké škále naopak pozitivní extrém odpovídá tomu, jak respondenti vnímají body na škále. V kognitivních rozhovorech respondenti uváděli, že na krátké škále je mezi hodnotami 4 a 5 velký rozdíl, pro dobré hodnocení tedy nelze zvolit čtyřku, protože tou již respondent instituci příliš „ublíží“.

„To je jakože... že takhle, když už je tam takhle málo zase těch, tak se mi zdá, že když jí dám tu čtyřku, tak jí hodně ublížím jakoby té bance, i kdybych jí doporučila. Že něco mezi bych tam dala.“

„*Takže ten rozdíl mezi pětkou a čtyřkou je poměrně velký?*“ „Pro mě jo. A určitě to není vo tom, že bych jako... možná doporučil.“ „*Takže by to znamenalo co, ta čtyřka?*“ „Čtyřka by znamenala, že už bych měl nějakou negativní zkušenost s bankou, ať je to cokoliv.“

Oproti tomu jedenáctibodová škála nabízí komfort nechat si maximální hodnotu 10 jako rezervu a zvolit méně silnou odpověď.

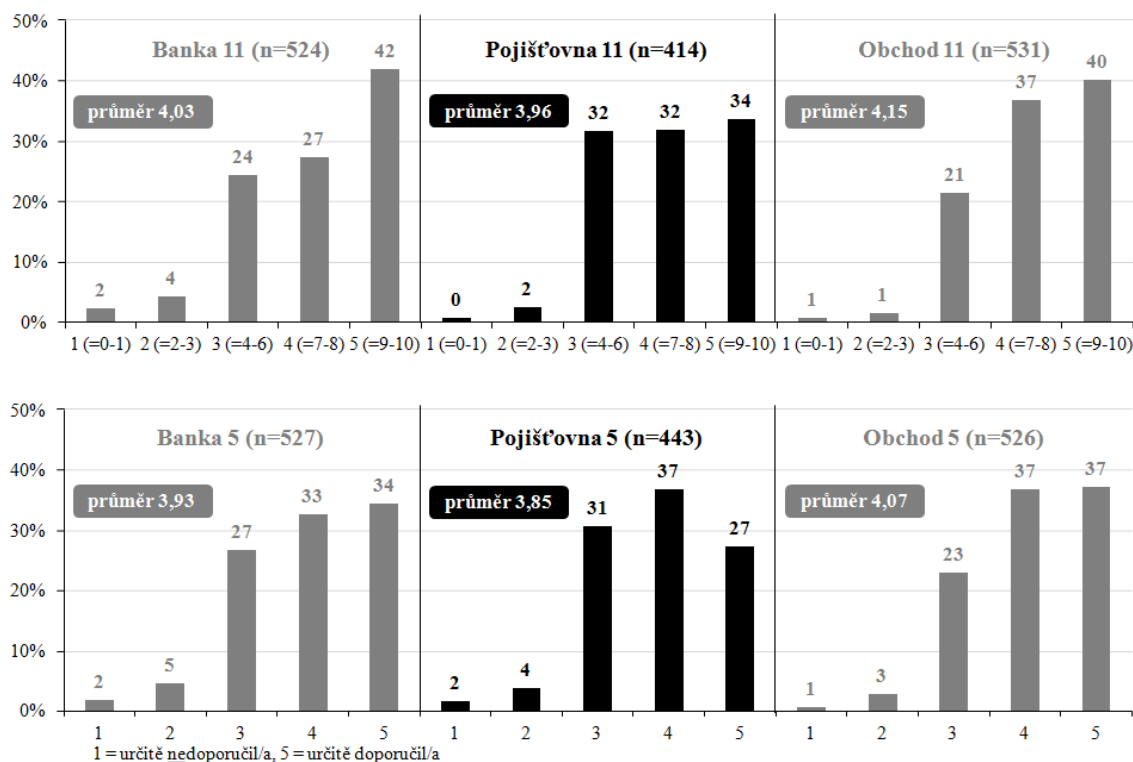
„Tak třeba, tady třeba tu desítku, už bych možná v této škále váhala dát tu desítku, už bych třeba dala devítku, protože to už je konkrétnější a ten jeden bodík bych si nechala vlastně takovou rezervu, že může se stát nějaký renonc, že to nemůžeme vyloučit, anebo teďka s Váma skončím průzkum a uvidím, že někde mají supr nabídku. A už budu na vážkách.“

„Jak už jsem řekla, desítka jako... Vždycky je tam... V jakýmkoliv obchodě nikdy bych asi nedala tu desítku. Je tam co zlepšit, nikdy to není stoprocentní.“ „*Takže desítku byste vůbec nedala?*“ „Neuvědomuju si, že bych to mohla dát jakýmkoliv obchodu, bance ani pojišťovně.“

Pro porovnání jedenáctibodové a pětibodové škály se nabízí ještě další logiky rekódování. Jelikož někteří respondenti vnímají jako neutrální hodnoty jakýsi rozšířený střed jedenáctibodové škály nebo že krajní hodnoty pětibodové škály pojmu více bodů na dlouhé škále (viz dále), ukazuje graf č. 3 rekód jedenáctibodové škály, kdy jsou sousední body spojeny.

---

<sup>30</sup> U banky je statisticky významný navíc také mírný rozdíl v zastoupení slabě negativních hodnot, u pojišťovny se významně liší zastoupení všech hodnot.

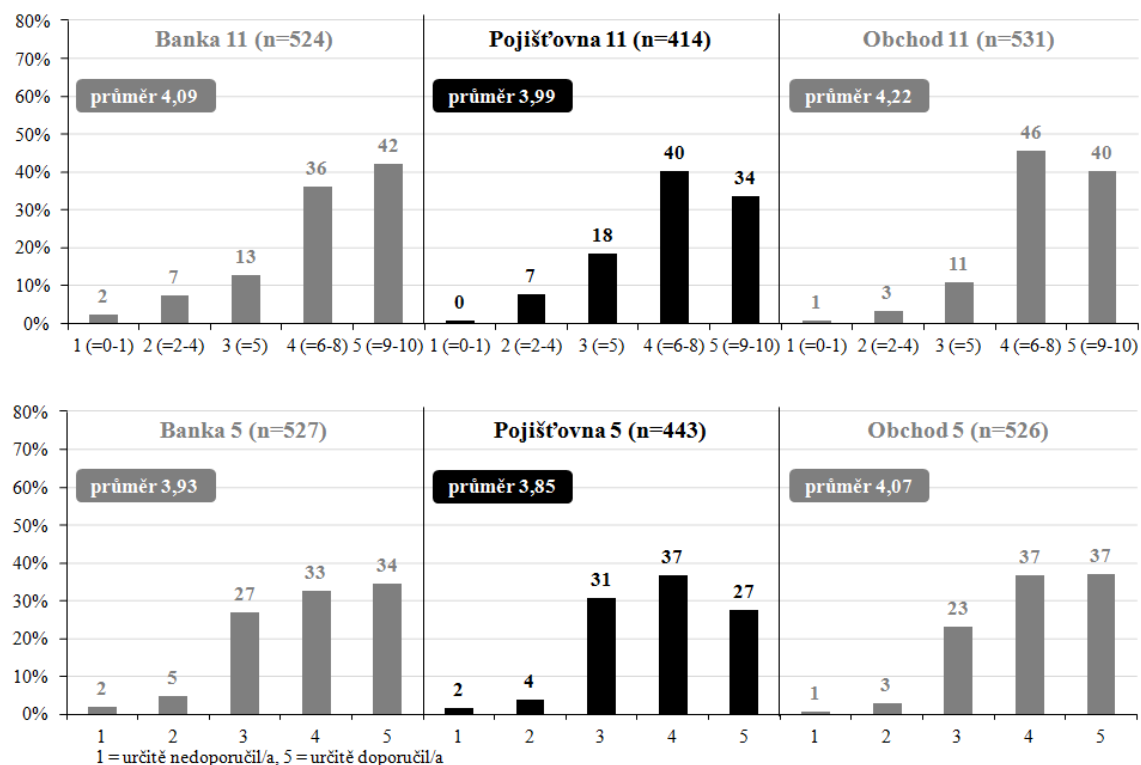
**Graf č. 3: Rekód – spojení sousedních hodnot (relativní četnosti)**

Ve výběrovém souboru jsou průměry u rekódované škály mírně (zhruba o 0,1 bodu) vyšší než u původní pětibodové, rozdíl mezi proměnnými ale není statisticky významný (dle Mann Whitneyho testu<sup>31</sup> – banka sig. 0,057, Z -1,905; pojišťovna sig. 0,124, Z -1,539; obchod sig. 0,185, Z -1,325). Z hlediska distribuce odpovědí se však škály vizuálně liší – zatímco u banky a obchodu je na pětibodové škále zastoupení pozitivních hodnot 4 a 5 stejné, na rekódované jedenáctibodové škále je u nich krajní hodnota 5 zastoupena častěji než bod 4 (zejména u banky, u obchodu jen mírně); u pojišťovny, kde hodnocení funguje mírně odlišně, je na pětibodové škále mírně pozitivní hodnota 4 zastoupena častěji než krajní 5, u rekódu dlouhé škály je pak jejich podíl vyrovnaný. Na dlouhé škále tedy respondenti u všech tří institucí častěji volili hodnoty 9 a 10 (tj. častěji je lze označit za promoters) než na pětibodové škále hodnotu 5. Rozdíl je nejvíce znatelný u banky, ani zde však není statisticky významný (dle chí kvadrát testu – sig. banka 0,118, pojišťovna 0,100, obchod 0,495).

Při rekódování je však potřeba mít na vědomí, že střední a extrémní hodnoty (tj. body 0 – 5 – 10, resp. 1 – 3 – 5) jsou v jistém smyslu speciální, mohou sloužit respondentům jako záchytné / referenční body. Proto rekód, který tyto hodnoty spojuje

s jinými, nemusí být ideální. Autorka proto provedla ještě další dva rekódy – jeden ponechávající střední hodnotu samostatně (graf č. 4), druhý ponechávající samostatně extrémy (graf č. 5).<sup>32</sup>

**Graf č. 4: Rekód – střed samostatně, krajní hodnoty spojeny se sousedními (relativní četnosti)**

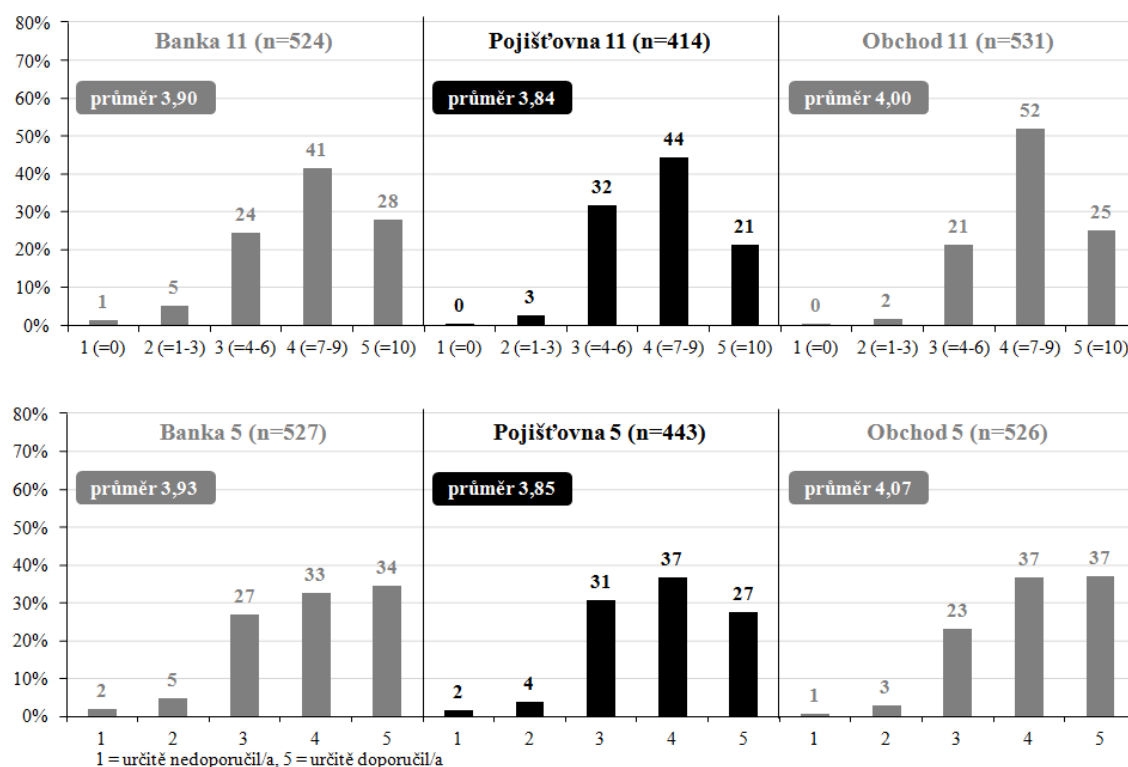


V tomto případě (graf č. 4) se již proměnné statisticky významně liší (dle Mann Whitneyho testu - banka sig. 0,002, Z -3,172; pojišťovna sig. 0,011, Z -2,555; obchod sig. 0,005, Z -2,801), průměr je na rekódované dlouhé škále o zhruba 0,15 bodu vyšší, výsledky jsou tedy mírně pozitivnější. Je to způsobeno především méně často zastoupenou střední hodnotou. Rozložení odpovědí se u všech tří institucí statisticky významně liší (dle chí kvadrát testu – sig. vždy 0,000 - a adjustovaných reziduí), ne však na úrovni stejných bodů – u banky je na rekódované jedenáctibodové škále méně zastoupený střed a více hodnota 5, to samé platí pro pojišťovnu, kde je ale navíc významný také malý rozdíl u mírně negativních hodnot (2). V případě obchodu je na rekódované jedenáctibodové škále významně více zastoupena hodnota 4 a méně střed.

<sup>31</sup> Vzhledem k ordinálnímu charakteru proměnných jsou v této práci uváděny výsledky neparametrických testů, tj. jsou testovány nulové hypotézy o shodnosti mediánů, nikoliv průměrů. Autorka přesto pokládá za důležité zobrazit v grafu podrobnější průměr, který lépe reflektuje případné rozdíly v odpovědích.

<sup>32</sup> Samostatný střed a zároveň extrémy jsou již zobrazeny v rekódu v grafu č. 2.

**Graf č. 5: Rekód – rozšířený střed, krajní hodnoty samostatně (relativní četnosti)**



Z grafu č. 5 je vidět, že spojení bodů 4 – 6 na jedenáctibodové škále do hodnoty 3 na pětibodové škále vyrovnává zastoupení neutrálního středu, na rekódu dlouhé škály je však stále vidět převaha méně silných pozitivních hodnot (zde 7 – 9) a menší zastoupení maximální pozitivní hodnoty. Tento rozdíl je statisticky významný (dle chí-kvadrát testu – sig. banka 0,035, pojišťovna 0,028, obchod 0,000 - a adjustovaných reziduí).<sup>33</sup> Průměry se zde neliší.

Opačným přístupem k rekódování je přeškálování krátké stupnice do rozmezí 0 až 10; k tomu autorka využila formuli Preston a Colmana upravenou dle Dawese:  $(\text{hodnota} - 1) / (\text{počet bodů na škále} - 1) * 10$ .<sup>34</sup> [Preston, Colman 2000: 7; Dawes 2008: 7] Porovnání průměrů takto vzniklé proměnné a původní jedenáctibodové škály ukazuje tabulka č. 4.

<sup>33</sup> U pojišťovny je významný navíc také malý rozdíl v zastoupení negativního extrému.

<sup>34</sup> Hodnoty pětibodové škály tedy byly přeškálovány na 0 – 2,5 – 5 – 7,5 – 10.

**Tabulka č. 4: Převedení pětibodové škály na rozmezí 0 – 10 (průměry)**

Otázka	11bodová škála	5bodová škála	Rozdíl (11bodová – 5bodová škála)
	<b>Průměr</b>		
A6_1 Doporučení banky	7,56	7,33	+0,23
A6_2 Doporučení pojišťovny	7,32	7,12	+0,20
A6_3 Doporučení obchodu	7,83	7,68	+0,15

Tabulka č. 4 ukazuje, že výsledky získané na kratší škále jsou (ve výběrovém souboru) po přeškálování mírně pozitivnější. Rozdíl však není statisticky významný (dle Mann-Whitneyho testu - banka sig. 0,081, Z -1,745; pojišťovna sig. 0,282, Z -1,077; obchod sig. 0,319, Z -0,996).

Na konec tabulka č. 5 ukazuje vypočtené hodnoty NPS pro jednotlivé typy institucí na dlouhé i krátké škále.<sup>35</sup>

**Tabulka č. 5: Porovnání škál z hlediska NPS**

Otázka	11bodová škála	5bodová škála	Rozdíl (11bodová – 5bodová škála)
	<b>NPS</b>		
A6_1 Doporučení banky	11	1	+10
A6_2 Doporučení pojišťovny	-1	-9	+8
A6_3 Doporučení obchodu	17	11	+6

Hodnota NPS je mírně vyšší na jedenáctibodové škále, rozdíl však není statisticky významný (dle t-testu porovnávajícího průměry proměnných vzniklých přiřazením hodnoty -100 skupině detractors, 0 skupině passives a +100 skupině promoters – banka sig. 0,054, t 1,931; pojišťovna sig. 0,169, t 1,377; obchod sig. 0,205, t 1,268). Navíc při pohledu na jednotlivé společnosti není tento trend jednoznačný – jsou i takové, které mají NPS získané na dlouhé škále nižší než NPS z krátké škály (viz tabulka č. 6 dále).

Celkově tedy pohled na distribuce odpovědí na jedenáctibodové a pětibodové škále ukazuje, že:

- Na krátké škále respondenti častěji volí střední, neutrální hodnotu, na dlouhé škále jsou naopak více zastoupené méně silné pozitivní hodnoty. Celkový podíl pozitivních hodnot je tak vyšší v případě využití dlouhé škály.

<sup>35</sup> NPS je rozdíl podílu promoters a detractors, na originální jedenáctibodové škále tedy (9 až 10) – (0 až 6). Na pětibodové škále byla hodnota NPS vypočtena takto: (5) – (1 až 3).

- Zastoupení maximální pozitivní hodnoty je ale častější na pětibodové škále.
- Průměry (resp. mediány) proměnných se liší pouze mírně (jsou o něco vyšší na jedenáctibodové škále) a rozdíly nejsou při většině zvolených rekodů statisticky významné. Hypotéza  $H_1$  o tom, že škály přináší různé výsledky, tak byla potvrzena na úrovni distribuce odpovědí – zastoupení jednotlivých hodnot, nelze ji však potvrdit na úrovni celkového výsledku (průměr, hodnota NPS) a pohledu přes typy institucí. (K výsledkům jednotlivých společností bude přihlédnuto dále.)
- Rozložení odpovědí se liší dle předmětu otázky – u pojišťovny je odlišné od banky a obchodu.

Lze však stejný trend vysledovat také v různých podskupinách výběrového souboru? Nebo je rozdíl „tažen“ pouze určitou skupinou respondentů?

### 3.3.3 Rozložení odpovědí v podskupinách

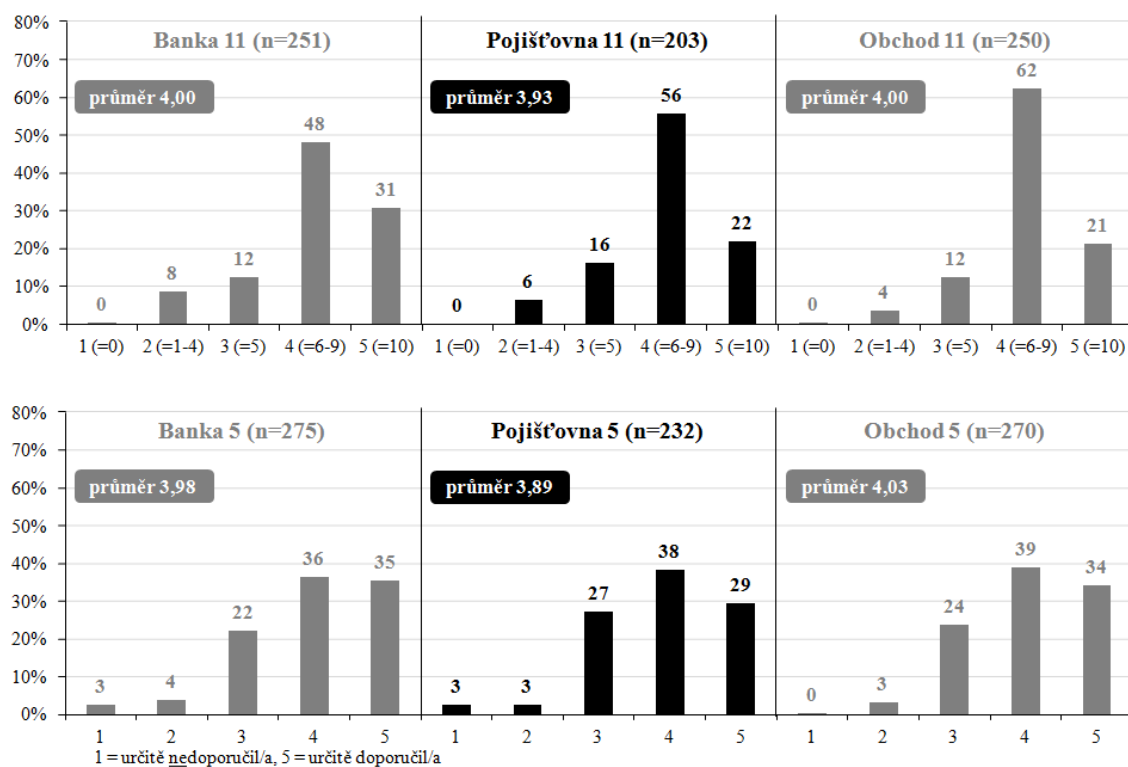
Z hlediska pohlaví lze sice v rámci výběrového souboru vyzorovat mírné odlišnosti v rozložení odpovědí na pětibodové škále (ženy volí u banky dle chí kvadrát testu - sig. 0,033 – a adjustovaných reziduí významně častěji než muži střední hodnotu  $3^{36}$ ), samotné fungování škál je však stejné jako při pohledu na celý výběrový soubor. Jako nejvhodnější pohled k porovnání obou typů škál na podskupinách autorka považuje rekod ponechávající střed i krajní hodnoty samostatně – zde je nejlépe vidět, jaký podíl respondentů zvolil střed, jaký méně silnou pozitivní hodnotu a jaký pozitivní extrém, tedy zda se výše uvedené principy týkají také podskupin. Grafy č. 6 (muži) a č. 7 (ženy) tedy ukazují distribuce odpovědí na původní pětibodové škále a na jedenáctibodové škále v tomto rekódu.

---

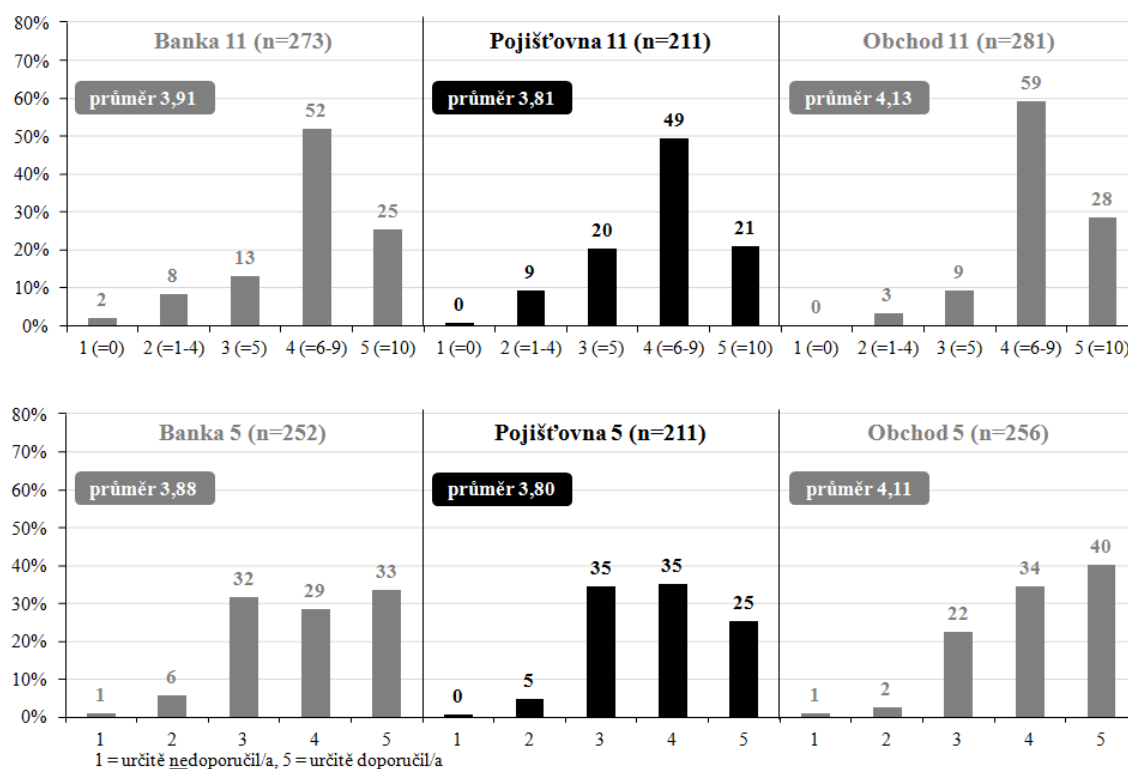
<sup>36</sup> Ostatní rozdíly nejsou statisticky významné, viz příloha č. 8 na CD.



Graf č. 6: Rozložení odpovědí – muži (relativní četnosti)



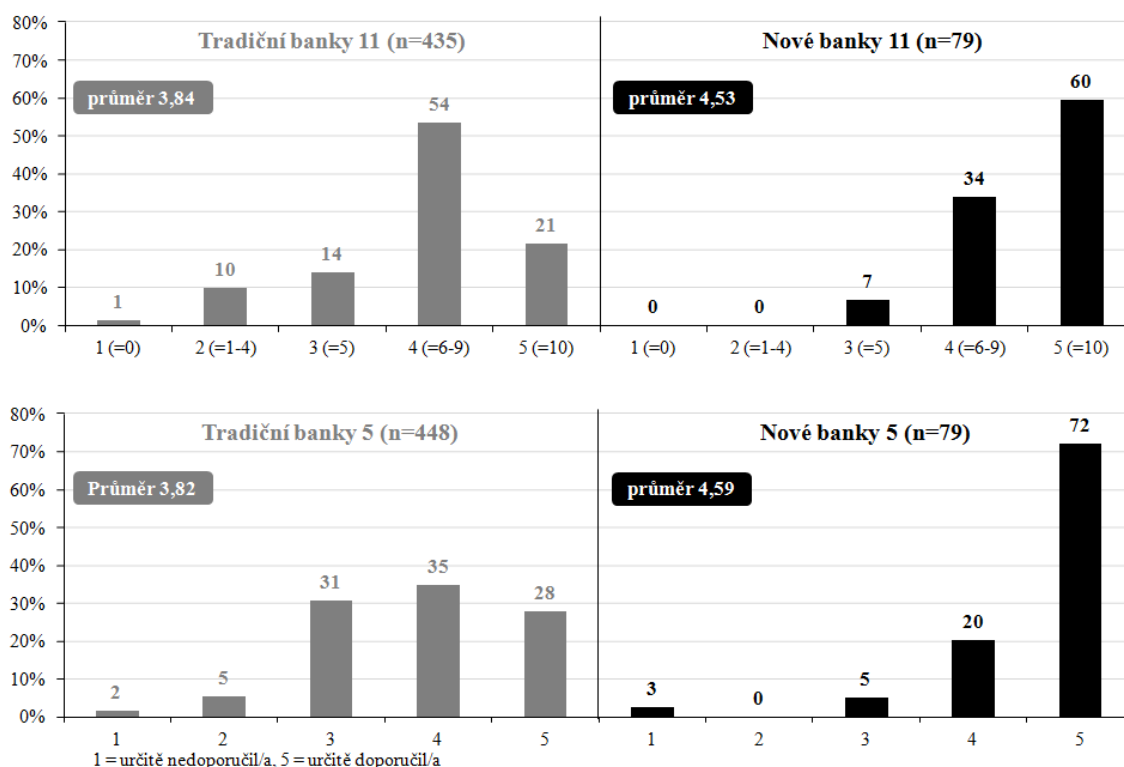
Graf č. 7: Rozložení odpovědí - ženy (relativní četnosti)



I v rámci podskupin výběrového souboru dle pohlaví tedy platí, že při hodnocení na pětibodové škále je častěji zastoupena střední hodnota (3) a krajní pozitivní hodnota (5), zatímco hodnotí-li respondenti na jedenáctibodové škále, vybírají častěji méně silné pozitivní hodnoty (rekódováno na 4).<sup>37</sup> Stejná situace nastává v rámci věkových (do 30 let, 31 – 45 let, 46 a více let) a vzdělanostních skupin (ZŠ, SŠ a vyučen/a, VŠ a VOŠ) – rozložení odpovědí je uvedeno v příloze č. 7.

Z hlediska situace na českém bankovním trhu je zajímavé se podívat, zda odpovídání funguje stejně také v rámci zákaznických skupin různých typů bank – tradičních „kamenných“ a nových „internetových“. Rozložení odpovědí těchto dvou klientských skupin na otázku o doporučení banky ukazuje graf č. 8.

**Graf č. 8: Rozložení odpovědí – tradiční vs. nové banky (relativní četnosti)**



Zatímco v případě tradičních bank fungují škály stejně jako v rámci celého vzorku (resp. vzhledem k převaze klientů tradičních bank fungují na úrovni celého vzorku škály stejně jako na úrovni doporučování tradičních bank), to znamená rozložení odpovědí se významně liší (dle chí kvadrát testu – sig. 0,000 - a adjustovaných reziduí)

<sup>37</sup> Výsledek chí kvadrát testu (v příloze č. 8 na CD) ukazuje, že ve všech případech (tj. u všech institucí a obou skupin rozlišených pohlavím) jsou rozdíly v rozložení odpovědí na krátké a dlouhé škále statisticky významné (vždy se liší hodnoty 3 a 4, v některých případech také krajní hodnota 5 nebo negativní hodnoty). V případě obchodu jsou však minimální očekávané četnosti v kontingenční tabulce nižší než 1. Konzervativní pravidlo pro využití chí-kvadrát testu vyžaduje, aby očekávané četnosti

– na dlouhé škále jsou častější méně silné pozitivní (6 – 9, rekódováno na 4) a mírně i negativní (1 – 4, rekódováno na 2) hodnoty, na krátké je oproti tomu častěji zastoupený neutrální střed (3) a maximální pozitivní hodnota (5). U nových bank nebylo vzhledem k nízké bázi a rozložení odpovědí (minimální očekávané četnosti byly nižší než 1, ve 38 % nižší než 5) možné přihlídnout k výsledku chí kvadrát testu. Z grafu č. 8 je však patrné, že ochota doporučit banku je u klientů natolik vysoká, že rozložení odpovědí je na obou typech škál podobné – tedy že i na dlouhé škále většina respondentů volí maximální pozitivní hodnotu (10, rekódováno na 5), ač je ve výběrovém souboru stále o něco méně zastoupená než na krátké škále. Na obou škálách pak následují méně silné pozitivní hodnoty 6 – 9, resp. 4 (více zastoupeno na dlouhé škále) a až potom neutrální střed. V případě tradičních bank jsou přitom hodnoty 3 – 5 na pětibodové škále rovnoměrně zastoupené. Zákazníci nových bank tedy využívají téměř výhradně pozitivní hodnoty odpovědí a zejména ty extrémní, čímž se vymaňují z celkového hodnocení ochoty doporučit banku.

### 3.3.4 Schopnost škál diskriminovat

Jednou z důležitých vlastností jakéhokoliv postojového znaku je jeho schopnost diskriminovat, tj. rozlišit respondenty s pozitivními a negativními postoji. V rámci úvah o NPS se nabízí rozlišení na promoters (9+10, resp. 5), passives (7+8, resp. 4) a detractors (0-6, resp. 1-3). Z hlediska sociodemografických charakteristik nejsou v podílech těchto skupin ani na jedné ze škál systematické rozdíly, tj. nelze říci, že by některá skupina (např. ženy, vzdělaní lidé, obyvatelé velkých měst atd.) obecně častěji patřila k promoters, jsou zde pouze ojedinělé případy, kdy se rozložení liší.<sup>38</sup> V zastoupení pouze top boxů (tj. hodnot 10, resp. 5) nejsou z hlediska sociodemografických charakteristik ani na jedné ze škál žádné významné rozdíly.<sup>39</sup>

---

v každé z buněk byly vyšší než 1 a ve většině buněk vyšší než 5. Pokud je ve více než 20 % buněk nižší očekávaná četnost než 5, je vhodné sloučit kategorie. [Norušis 2005: 167]

<sup>38</sup> Při měření pomocí pětibodové škály jsou u banky častěji detractors ženy. U pojišťovny jsou při využití jedenáctibodové škály mladí lidé do 30 let častěji passives a méně často promoters, na pětibodové škále jsou lidé z malých obcí do 1 000 obyvatel a lidé s čistým osobním měsíčním příjmem nad 15 000 Kč častěji passives a méně často detractors, obyvatelé z měst s 5 – 20 tis. obyv. méně často passives a lidé s příjmem do 8 000 Kč častěji detractors a méně často passives. U obchodu lze najít rozdíly na jedenáctibodové škále: muži jsou častěji detractors, ženy naopak promoters, lidé z malých obcí do 1 000 obyvatel jsou méně často passives, obyvatelé obcí s 5 – 20 tis. obyv. jsou passives naopak častěji a méně často promoters. Tyto rozdíly jsou statisticky významné (dle chí kvadrát testu a adjustovaných reziduí), viz příloha č. 8 na CD.

<sup>39</sup> Autorka pro otestování použila rekód z grafu č. 5, kdy jsou krajní hodnoty ponechány samostatně. Ve většině kontingenčních tabulek jsou však příliš nízké očekávané četnosti, výsledek chí kvadrát testu lze tedy brát pouze jako indikativní – viz příloha č. 8 na CD. Testovány byly rozdíly dle pohlaví, věku,

Poměry jednotlivých klientských skupin – promoters, passives a detractors - a tedy i výsledné NPS se ale samozřejmě liší mezi jednotlivými společnostmi (viz příloha č. 8 na CD) – zde je pro zjednodušení uvedena pouze vypočítaná hodnota NPS. V případě měření NPS je právě diskriminace na úrovni hodnocené společnosti (tj. rozlišení společností, jejichž zákazníci jsou ochotni je doporučit, a těch, jejichž zákazníci by je nedoporučili) důležitá. Zadavatel výzkumu potřebuje vědět, zda je jeho hodnota NPS vyšší/ srovnatelná/nížší než NPS konkurentů. Tabulka č. 6 ukazuje hodnoty NPS získané na obou typech škál dle jednotlivých značek.<sup>40</sup>

**Tabulka č. 6: Porovnání škál z hlediska NPS u jednotlivých společností**

Hodnocená společnost	11bodová škála	5bodová škála	Rozdíl (11bodová – 5bodová škála)
<b>Banka</b>	<b>NPS</b>		
BANKA1	56	62	-6
BANKA2	27	7	+20
BANKA3	9	-5	+14
BANKA4	2	-13	+15
BANKA5	-5	-1	-4
BANKA6	-12	-20	+8
<b>Pojišťovna</b>	<b>NPS</b>		
POJIŠŤOVNA1	37	38	-1
POJIŠŤOVNA2	-3	0	-3
POJIŠŤOVNA3	-6	0	-6
POJIŠŤOVNA4	-6	-8	+2
POJIŠŤOVNA5	-21	-28	+7
<b>Obchod</b>	<b>NPS</b>		
OBCHOD1	58	49	+9
OBCHOD2	41	34	+7
OBCHOD3	33	12	+21
OBCHOD4	10	10	0
OBCHOD5	9	16	-7
OBCHOD6	-10	-24	+14
OBCHOD7	-16	-6	-10

Podíváme-li se na hodnoty NPS ne pro banku, pojišťovnu a obchod obecně jako v tabulce č. 5, ale podle jednotlivých společností (tabulka č. 6), je zřejmé, že ne vždy je výsledná hodnota na jedenáctibodové škále vyšší než na pětibodové, u některých

---

vzdělání, regionu (Praha, Čechy, Morava – spojeno kvůli nízkým očekávaným četnostem u kraje), velikosti místa bydliště a čistého osobního příjmu.

<sup>40</sup> Jelikož cílem práce není porovnání jednotlivých společností na českém trhu, jsou jména společností nahrazena názvy „BANKA1“, „POJIŠŤOVNA2“ atd. Číslování společností vychází z pořadí hodnot NPS na jedenáctibodové škále a neodpovídá tedy hodnotám společností v dotazníku. Do tabulky vstupují pouze společnosti, které na krátké i dlouhé škále hodnotilo alespoň 30 respondentů.

společností naopak pětibodová škála generuje lepší výsledek, u jiných se obě hodnoty liší pouze minimálně nebo vůbec. Co se týče pořadí společností dle výše jejich NPS, je na jedenáctibodové a pětibodové škále podobné, u banky a obchodu jsou však i výjimky (BANKA5 má na pětibodové škále lepší pořadí než na jedenáctibodové, stejně jako OBCHOD5 a OBCHOD7).

Z tabulky č. 6 je však patrné, že rozpětí jednotlivých hodnot NPS v rámci typů institucí je u jedenáctibodové i pětibodové škály vysoké (u banky 68 bodů na jedenáctibodové vs. 82 bodů na pětibodové, u pojišťovny 58 vs. 66 bodů, u obchodu 74 vs. 73 bodů). Dle analýzy rozptylu (ANOVA) nejsou hodnoty NPS<sup>41</sup> jednotlivých bank, pojišťoven ani obchodů shodné (sig. 0,006 u pojišťovny na jedenáctibodové škále, jinak vždy sig. 0,000) – na jedenáctibodové i pětibodové škále se mezi sebou vždy alespoň dvě hodnoty statisticky významně liší.

Lze tedy říci, že jak jedenáctibodová, tak pětibodová škála dokážou rozlišit společnosti se zákazníky ochotnými je doporučit a společnosti se zákazníky neochotnými je doporučit, výsledné hodnoty NPS (a někdy i pořadí společností) se však mezi sebou liší, přičemž tento rozdíl není systematický (nelze například říci, že jedenáctibodová škála obecně generuje vyšší NPS). Hypotéza H<sub>1</sub> je tak na úrovni jednotlivých společností potvrzena – jedenáctibodová a pětibodová škála přináší rozdílné výsledky; nelze je však souhrnně uchopit. NPS získaná na různých typech škál tak není možné vzájemně porovnávat.

### 3.3.5 Vnímání bodů na škále

Vnímání bodů na jedenáctibodové ani pětibodové škále není mezi respondenty jednotné, pracují s nimi různě. Tabulka č. 7 ukazuje vnímání jednotlivých hodnot respondenty<sup>42</sup>, kteří se zúčastnili kognitivních rozhovorů – autorka jej zrekonstruovala na základě deklarácí respondentů, které hodnoty vnímají jako pozitivní/neutrální/negativní, v kombinaci s tím, jak odpovídali a hodnotili své odpovědi. Zatímco deklarace byly obvykle symetrické, s jedním středním bodem vnímaným jako neutrální, v hodnocení se od nich respondenti v některých případech odchylovali.

<sup>41</sup> Hodnoty NPS jsou průměry proměnných, jejichž hodnoty jsou -100 pro detractors, 0 pro passives a 100 pro promoters. Průměr takové proměnné odpovídá rozdílu podílu promoters a detractors.

<sup>42</sup> Autorka nakládá s informacemi získanými od respondentů důvěrně, v textu práce i příloze proto uvádí veškeré informace anonymizovaně. Pořadí respondentů v tabulce odpovídá pořadí přepisů jejich rozhovorů v příloze č. 9 na CD.

Tabulka č. 7: vnímání bodů na škále

Respondent	11bodová škála			5bodová škála		
	Negativní	Neutrální	Pozitivní	Negativní	Neutrální	Pozitivní
Respondent1	0 - 4	5 - 6	7 - 10	1	2 - 4	5
Respondent2	0 - 4	5 - 6	7 - 10	1 - 3	4	5
Respondent3	0 - 3	4 - 6	7 - 10	1 - 2	3	4 - 5
Respondent4	0 - 4	5	6 - 10	1 - 2	3	4 - 5
Respondent5	0 - 4	5	6 - 10	1 - 2	3 - 4	5
Respondent6	0 - 4	5	6 - 10	1 - 2	3	4 - 5
Respondent7	0 - 4	-	5 - 10	1 - 2	3 - 4	5
Respondent8	0 - 3	4 - 6	7 - 10	1 - 2	3	4 - 5

Jak je vidět z tabulky č. 7, respondenti vidí škály rozmanitě – je tak narušen předpoklad, že respondenti rozumí významu jednotlivých hodnot všichni stejně, čímž může trpět kvalita získaných dat. [Krosnick, Judd, Wittenbrink 2005: 35] V případě jedenáctibodové škály spočívá hlavní rozdíl v tom, zda je za neutrální považována pouze střední hodnota 5 nebo je střed rozšířený a hodnota 6 (případně i 4) je vnímána rovněž neutrálně.

„Šestka to už bych zas viděl víc, že to je ten průměr, takže to už...“

„A potom vod tý pětky vejš, to je takový vo tom, že samozřejmě si kdokoliv může najít, že teda ta banka není až tak dobrá, i když já jí doporučuju. A to samý s těma ostatníma no.“ „To znamená od šestky nahoru už je to pozitivní nebo?“ „Od šestky asi... spíš od sedmičky. Od sedmičky nahoru no.“ „Takže i ta šestka je ještě neutrální.“ „Tak ještě ta pětka, šestka... No, ta šestka možná ještě je, s tou pětkou jsou takový neutrální.“

„Co to znamená teda ta čtyřka pro Vás?“ „Ehm... No že v tuto chvíli já už, kdyby někdo se mě ptal, řekla bych pouze ‚mám jí tam, ale vím, že nic takového už Vám nenabídnou, zkuste to v jiné bance.‘“ „Dobrá. A proč jste třeba nedala nějaké nižší číslo ještě?“ „Ehm... protože s tím produktem, který u nich mám, já jsem jakoby spokojená. Ale na jedničku, na dvojku to určitě není, protože to by muselo bejt něco, co opravdu bylo špatně, nebylo tam plnění, nebylo já nevím... cokoliv by bylo špatně. A ta trojka je ještě furt blíž k tomu nedoporučila, takže ta čtyřka je taková ta jako... (...) ...tadyta čtyřka, pětka nebo ta pětka pro mě je ten prostředek jako nevím opravdu, já jsem s ní tak, jak jsem s ní, ale záleží tom daném příteli nebo tohle... Nevím, jestli bych jí doporučila no. Je to tak jako na vážkách.“

Pozitivní hodnoty nejsou vnímány všechny stejným způsobem, některé jsou mírně pozitivní, jiné znamenají v očích respondentů vysokou ochotu doporučit/spokojenost (v závislosti na tom, jak chápou otázku). Hodnoty 7 a 8 (případně ještě 6, není-li neutrální) bývají vnímány jako spíše nadprůměrné, lehce pozitivní. Právě u těchto dvou hodnot se v některých případech stává, že respondentům vzájemně splývají.

„Tak a kdybyste mi mohla k té bance říct, Vy jste dala sedmičku, tak co ta odpověď tedy znamená?“ „Je to jako... je to doporučení, je to banka, kterou jsem s ní spokojená, ale

některé výhrady mám, takže není to úplně ta desítka, ta stoprocentní ‚určitě doporučila‘.“

„Takže nakonec mi z toho vyjde nějaké číslo. Takže u té banky by mně vyšlo taková šestka, sedmička.“ „*Hmhm.*“ „Jsem tam spokojený, služby jsou docela dobrý, ale jsou tam věci, který se mi nelíbí a které bych třeba chtěl i nějak řešit, takže proto.“

„Ehm... je to, je to bliž k tomu... je to jako hodně mi to splývá tahleta sedmička, osmička dohromady. Není tam až takovej jakoby rozdíl, kterej bych řekla ‚tohle je osmička a když to nesplňuje, já nevím, čerstvost zeleniny, už je to sedmička.‘ Jakoby nemám tam úplně ten nějaký detail, kterej bych tam viděla v tom mínusu.“

„Tak tady, vzhledem k tomu, že to nemám s čím porovnat, tak pro Vás jako jestli je to sedmička nebo osmička, nevím, jestli to má nějaký význam.“ „*A pro Vás to má nějaký význam? Liší se ta sedmička s osmičkou?*“ „To je zase v daným okamžiku, co by mě napadlo všechno za věci... poskládaný z informací, co jakoby držím v krátkodobý paměti jo. Takže... význam to pro mě úplně nemá.“

Oproti tomu hodnoty 9 a 10 bývají vnímány jako silně pozitivní. Desítka je maximální hodnota a je i slovně označená jako „určitě doporučil/a“; zde tedy vnímání vychází z jejího postavení vzhledem k ostatním odpovědím, devítka je však také silná.

„...*OBCHOD, tak tam jste dala taky desítku.*“ „Tam jsme spokojený se všim no. S potravinama, s přístupem, s pokladníma, se všim no.“

„Tady devět, deset – určitě bych doporučila...“

„*A potom co ta osmička, liší se nějak od té devítky? Případně jak?*“ „To ne, já si myslím, že to je, že ta banka je jakoby dobrá ta BANKA, takže to už by...“ „*Ta osmička by už znamenala co?*“ „To by bylo takový, že bych už byla víc na vážkách, jestli bych jí opravdu měla doporučit nebo ne. Že jakoby ta devítka je přece jenom že je to takový ne stoprocentní, ale prostě devadesát procent.“

Vezmeme-li v úvahu také fakt, že rozložení odpovědí je vychýleno k pozitivnímu konci škály (tj. že dochází k *positivity bias* [Tourangeau, Rips, Rasinski 2000: 240 - 241]), zdá se rozlišení klientských skupin dle Reichelda opodstatněné – ti, kteří volí výrazně pozitivní odpovědi 9 a 10, jsou promoters, mírně pozitivní 7 a 8 passives a neutrální či negativní odpovědi detractors. Negativní hodnoty 0 – 4 jsou málo zastoupené (9 % u banky, 8 % u pojišťovny, 4 % u obchodu – viz tabulka č. 3). Jelikož je respondenti obvykle pro hodnocení nepoužívají, mohou na škále splývat. To může dle autorky souviset s již nastíněnými dvoustupňovými strategiemi odpovídání, kdy je nejdříve na základě z paměti vyvolaného materiálu vybrán interval bodů či určitá část škály (obvykle pozitivní či neutrální) a až následně respondent přikročí k upřesnění konkrétního bodu; negativní hodnoty tak zůstávají již v prvním kroku vyřazené a respondent je vnímá pouze celkově jako „balík“ negativních.

„Takhle v tenhle moment jsou asi hodně podobný, ale opravdu kdybych měl nějakou zkušenost, co by přímo mně jakoby snižovala u nich ten... tu bodovou škálu, tak bych asi i nad tím přemýšlel, že dám prostě menší hodnotu.“

„No, tady už jo, tady jo. Pětka je opravdu ten neutrální pro mě nebo ten střed a čtyřka... čtyřka je tak jako takový menší mínus, jako není to úplně tak špatný, někomu by možná vyhovoval ten produkt, ale potom tři, dva, jedna, to už jako je... bych ani svým přátelům neřekla.“

Nerozlišování jednotlivých negativních bodů ale nemusí být záležitostí pouze dlouhé škály, může se objevit i v případě hodnocení od 1 do 5.

„Já bych řekla, že neutrální je pak ta čtyřka no. Že to není takový to číslo uprostřed, ale že od té čtyřky už je to dolu... Už je to na tom, že bych nedoporučila nikomu z toho nic no.“ (...) „*Hmhm. A jsou tam nějaký rozdíly mezi těma třema bodama - ta trojka, dvojka, jednička - nebo to je pro Vás prostě...*“ „No, to je... už je to asi negativní no.“

Tabulka č. 7 rovněž ukazuje, že vnímání hodnot je mírně posunuté od klasického „matematického“ uvažování také na krátké pětibodové škále. Někteří respondenti sice vnímají jako neutrální pouze střední hodnotu 3, jiní však do neutrálního pole zahrnují také bod 4. Pozitivní tak zůstává pouze krajní hodnota 5, která se – jak již bylo popsáno dříve v této práci – od čtyřky velmi liší.

### 3.3.5.1 Krajní a střední hodnoty

Specifickými hodnotami na škále jsou extrémy (0 a 10, resp. 1 a 5) a střed (5, resp. 3). V závislosti na celkovém počtu bodů na škále se může měnit jejich postavení – vnímání hodnoty 10 na jedenáctibodové škále nemusí být stejné jako u hodnoty 5 na pětibodové škále. V porovnání s dlouhou škálou někteří respondenti uvádí, že pětka je méně silná než desítka, která je skutečně bezvýhradně pozitivní. (Toto vnímání však není univerzální, někteří považují krajní hodnoty 10 a 5 za srovnatelné.)

„Myslím, že jsou... stejný no. Tam se liší jenom to číslo, ale myslím si, že jakoby jako na hodnocení je to stejný.“ „*Je to stejný, je to ten jakoby maximální... Akorát že tady jste třeba tam zařadila i tu BANKU, kde tady jste jí desítku nedala.*“ „No, to jo no, tady jsem jí nedala. Tak v tomhle, když je to do té pětky, tak tu pětku už bych jakoby té BANCE dala no.“

„Tady (*pětibodová škála*), když jsem jako velmi spokojená, tak bych určitě dala tu pětku ‚určitě doporučila‘, protože nemám dát možnost tu devítku, kde je něco určitě... nějaká možnost na zlepšení, to nejde. Jo takže tady bych devět, deset (*jedenáctibodová škála*), když bych spojila dohromady, tak bych určitě tu pětku dávala.“ „*Takže tahle je, dejme tomu silnější ta desítka.*“ „No ta desítka je na téhle dlouhé škále, je opravdu stoprocentní, stála bych si za ní.“

„No není to maximálně, kdyby to bylo maximální, tak bych dal tady desítku. Jo ale když máte tu škálu takhle krátkou, tak jsem víc spokojenej než čtyřka. Jo takže když jsem víc než čtyřka, tak už je jenom pětka. A není tam zas nic zásadního, co by mě nutilo jít do té čtyřky. Takže tady bych šel do toho ‚určitě doporučil‘.“



Na druhém konci škály proti sobě stojí hodnoty 0 a 1.<sup>43</sup> Pro ty platí v zásadě to samé jako pro pozitivní extrémy 10 a 5 – někteří respondenti je vnímají stejně, jiní se domnívají, že nula je silnější, zatímco krajní hodnota 1 na pětibodové škále tak odpovídá hodnotám 0 + 1 na jedenáctibodové škále.

„Jako ‚určitě nedoporučila‘, tady zase nula, jedna bych spojila, možná nula, jedna, dva dohromady tady (*jedenáctibodová škála*). Že bych to nechtěla dát tak strašně negativní, jedničku, dvojku bych tady volila, protože určitě někdo je spokojenější nebo nějakou část má dobře, takže bych to... A tady bych dala tu jedničku (*pětibodová škála*).“

„No, jednička beru jako nula jedna, dvojku beru dva a tři, trojku čtyři a pět a tak dále.“

Nad rámec toho má však nula navíc v číselném systému své zvláštní postavení, což má na její vnímání jako hodnoty na škále vliv. Pro některé respondenty je zarážející, neobvyklá pro hodnocení, nadbytečná.

„Nula je... nevim no. Nevim, jak bych vodpověděla, ale ta nula pro mě není číslo jako. Jo to je... ve škole se taky nehodnotí nulou, takže nevim.“ „*Takže ta je tam taková problematická.*“ „Taková... problém no. Když bych jako chtěla jet od pětiky dolů, tak bych třeba dala, nevim, jedničku, dvojku. A tu nulu bych určitě nepoužívala no.“

„No úplně teda, abych... No je zajímavý, že jako se nula jako že definovala taky, většinou je to že jo vod tý jedničky do desítky.“ „*A kdyby to teda bylo od tý jedničky do desítky, tak byl by to nějaký rozdíl?*“ „Vono by se to asi nezměnilo.“ „*Nějak by se to moc nezměnilo.*“ „Vo to jedno číslo...“

Kromě toho může být nula matoucí – přestože právě ona činí jedenáctibodovou škálu lichou, tedy mající střední hodnotu 5, některé respondenty zmátla a domnívali se, že díky „přidané“ nule nad rámec rozmezí 1 – 10 není hodnota 5 uprostřed.

„Je to pravda, že člověk bere jakoby třeba tu pětku takovej ten pomyslej... tedy vlastně pětka pomyslej nějaký střed že jo a není to, je to mezi čtyřkou, pětou něco. Protože když máte jedenáctibodovou škálu, tak pětka není prostředek že jo, je to něco mezi. Ale i tak beru jakoby... i když je to od nuly, tak beru tak tu pětku tak jako střed.“

„No tak jako vzhledem k tomu, že jako to není sudý číslo, ale... tak jsem si říkala, že mezi nima je prostě ten střed ta trojka, tak jsme to tak nějak vnímala takhle.“ „*Tohle je taky lichý právě kvůli tý nule.*“ „Jo to je nula, aha, jo to je pravda.“

V souvislosti s hypotézou  $H_4$  se tak ukazuje, že nula je skutečně vnímána odlišně než ostatní čísla; je také v kvantitativních datech málo zastoupená (1 % u banky, 0,2 % u pojišťovny a obchodu). Stejně tak je ale málo zastoupená hodnota 1 na pětibodové škále; rozdíl mezi zastoupením negativních extrémů na dlouhé a krátké škále je malý - cca 1 % - a pouze u pojišťovny statisticky významný (dle chí kvadrát testu – sig. 0,028 - a adjustovaných reziduí). Hypotézu  $H_4$  tak nelze plně potvrdit ani vyvrátit.

<sup>43</sup> V tomto jediném aspektu se varianty dotazníku mezi sebou lišily nad rámec testované délky škály – zatímco originální jedenáctibodová škála má hodnoty od 0 do 10, pro krátkou škálu zvolila autorka tradiční 1 – 5. Pro zajištění úplné shodnosti by bylo třeba mít hodnoty od 0 do 4, ovšem tuto alternativu

### 3.3.6 Náročnost a preference respondentů

Preference respondentů pro délku hodnotící škály nejsou jednotné. Obecně se shodují, že odpovídání na pětibodové škále je jednodušší a rychlejší; někteří ji proto preferují, jelikož delší škála pro ně znamená o něco větší zátěž.

„Z tohoto už se vybírá líp (*pětibodová škála*) než tady z těch více bodů, ale...“ „*A vybírá se, protože... Z jakýho důvodu se Vám vybírá líp?*“ „No, nemusím nad tím tolik jako přemejšlet, protože přeci jenom tady, těch víc bodů je možnost, že tam je... mezi pětkou a desítkou, že tam je víc... takový jako jestli jsem spokojená hodně, málo... a vod tý pětky že jo to je, že nejsem spokojená.“

„No že to je hrozně dlouhý hodnocení teda. (smích)“ „*Hmm. A přijde Vám to lepší, horší?*“ „No, upřímně teda těch pět bodů bohatě stačí, protože teď jako si vůbec nedokážu vyhodnotit, jak bych tohle měla jako hodnotit v podstatě no.“

Jiní ale upřednostňují dlouhou škálu, protože ač u ní musí dle svých slov více/déle přemejšlet, považují svou odpověď za přesnější; zároveň jsou ale stále schopni svůj postoj zakódovat do jedné z nabízených hodnot, není to pro ně příliš složité.

„Tak asi ta desetistupňová škála je určitě asi lepší, že člověk se opravdu musí více zamyslet a vybrat... opravdu přesnější možnost.“

„No určitě ta jemnější, protože tam můžete... můžete to umístit přesněji.“ „*Hmm. A není to pro Vás příliš složité?*“ „No není právě. Navíc si myslím, že pokud se dělá takovýhle výzkum, tak statisticky je určitě mnohem lepší tohle (*jedenáctibodová*) než tohle (*pětibodová*). Jo protože tady máte těch možností málo, a tak málo škatulek. Já bych určitě teda raději vyhodnocoval tenhle výzkum (*jedenáctibodová*) než tenhle (*pětibodová*).“

V obou případech – při obou variantách délky škály – si však respondenti pomáhají přirovnáním škály k lépe uchopitelnému číselnému systému. Zatímco škála od 0 do 10 evokuje procenta (0 % – 100 %), u pětibodové škály přináší zjednodušení přirovnání ke školní klasifikační stupnici (a to přestože je školní hodnocení opačné – 1 nejlepší, 5 nejhorší). Právě díky otočení hodnot oproti školním známám však může tato asociace respondenty zmást.

„Ne, pro mě to znamená nula procent. V mém vnímání. A tohle je pro mě sto procent, že jsem maximálně spokojená nebo má to pro mě jako supr reference. (*10 na jedenáctibodové škále*)“

„Tak asi podle... jak je to jako ve škole skoro známkování, akorát že tady to je naopak, že tady jednička je to nejhorší, takže asi jako bodově, kdybych to známkoval nebo hodnotil...“

„No, každopádně mě to ze začátku mate že jo, protože jsem byla zvyklá jako... ehm nad tím musím chvíli teda přemejšlet, protože jsem byla zvyklá že od jedničky do pětky ve škole jako... že to je, takže jsme vždycky musela tak jako...“

Celkově však platí, že jedenáctibodová škála je pro respondenty o něco náročnější a je tedy potvrzena první část hypotézy H<sub>5</sub> – delší škála znamená vyšší kognitivní zátěž. Druhá část hypotézy uvádí jako důvod komplikace s rozlišením sousedních hodnot, což se zde ukázalo mezi hodnotami 7 a 8, u ostatních bodů však ke komplikacím obvykle nedocházelo. Na druhou stranu však z kognitivních rozhovorů nevyplývá, že by respondenti nebyli schopni hodnocení na jedenáctibodové škále provést – hodnotu vyberou a uvedou důvody své odpovědi. Při úvahách, zda využít či nevyužít dlouhou, jedenáctibodovou škálu, by však bylo vhodné zvážit její vhodnost pro konkrétní předmět výzkumu. Jak ukazuje rozložení odpovědí z kvantitativní fáze výzkumu, škála funguje odlišně u ochoty doporučit pojišťovnu, se kterou nejsou lidé v běžném kontaktu. Úloha je tak složitější již jen z toho důvodu, že respondent hodnotí něco méně známého, jeho postoj může být předem méně zformovaný či být formován až v návaznosti na otázku. Stejně tak někteří respondenti uvádí, že dlouhou škálu by více uvítali v případě hodnocení něčeho konkrétnějšího – například produktu, nikoliv celé společnosti.

„Možná že kdyby... možná že... já nevím no. Možná kdyby se tam objevil nějaký jako úplně konkrétní produkt nebo něco, tak tam už bych detailnějš to hodnotila, ale takhle v globále mi to přijde hodně podobný no.“

„Co se týká pojišťovny, tam bych dala... dvojku. A nevím teď, jestli bych neváhala... možná kdyby to bylo konkrétně k životnímu... protože takhle jako obecně pojišťovna to nevím, protože má různý. Třeba důchodový bych určitě třeba nedoporučila tam, ale třeba životní ano, takže...“

„Hm, to je zase přesně o tom, pokavaď tam následuje jakoby možnost výběru z těch deseti (*deseti různých společností, které by měl pán všechny hodnotit, tzn. porovnávat přiřazováním různých hodnot*), tak samozřejmě preferuju tohle (*jedenáctibodová škála*), protože tady (*pětibodová škála*), tam toho moc nevymyslím; a pokavaď bude vyložene jenom nějaká jednoduchá otázka na to, jestli bych doporučil konkrétní, jednu, tak je tohle, těchle pět.“

## Závěr

Autorka v této práci testovala vliv zvolené délky škály na měření, a to z hlediska kognitivních procesů probíhajících v myslech respondentů i dopadu na získaná data. Na základě studia literatury formulovala a na datech získaných pomocí split ballot experimentu a kognitivních rozhovorů testovala hypotézy týkající se vzájemné srovnatelnosti výsledků, rozložení odpovědí a náročnosti úlohy. Pro porovnání dat na obou typech škál bylo využito rekódování či přeškálování proměnných tak, aby měly stejný počet bodů. Ukázalo se, že odpovídání na jedenáctibodové a pětibodové škále

znamená pro respondenty dvě odlišné kognitivní úlohy a že získaná data se od sebe liší. Žádná z hypotéz nebyla vyvrácena, některé se však nepodařilo plně potvrdit.

Jedenáctibodová škála od 0 do 10, která je originální pro zde zvolený aplikační příklad (Net Promoter Score vyjadřující míru ochoty doporučit využívanou banku, pojišťovnu a obchod s potravinami), je pro respondenty o něco náročnější; důvodem se zdá být spíše než v literatuře uváděné nesnáze s rozlišením sousedních bodů (i když z rozhovorů vyplývá, že u hodnot 7 a 8 k nim může docházet) nutnost více se nad hodnocením zamyslet, vyvolat větší množství relevantních informací. Strategie respondentů přitom bývá dvoufázová – nejdříve si výběr odpovědí na základě vyvolaných informací zúží pouze na část škály či interval bodů, pro volbu konkrétní odpovědi následně zvažují další informace. Výsledkem je však dle vnímání respondentů přesnější odpověď, což koresponduje s *informační teorií*, která říká, že delší škála poskytuje respondentovi efektivnější rámec pro přenos informací, ten tak může přesněji vyjádřit svůj postoj. [Alwin 1997: 322]

Ačkoliv kvantitativní data ukazují na určitou tendenci, že jedenáctibodová škála generuje mírně pozitivnější výsledky než pětibodová (vyšší průměr, vyšší hodnotu NPS při pohledu přes typy institucí), tento trend se zde při zvolených rekódech až na jednu výjimku neukázal jako statisticky významný. Na dlouhé škále je sice signifikantně vyšší zastoupení pozitivních hodnot (6 – 10 vs. 4 – 5), na krátké však významně větší podíl respondentů volí pozitivní extrém, což celkový výsledek téměř vyrovnává. Není však vyloučeno, že by se podobný trend ukázal i při dalších opakováních výzkumu - ty by jej mohly potvrdit nebo naopak ukázat, že byl zachycen skutečně pouze v tomto výběrovém souboru.

Jedenáctibodová škála svým širokým bodovým rozpětím respondentům poskytuje prostor pro zmírňování svých odpovědí, dochází tedy k *response contraction bias*, kdy se respondenti zdráhají použít extrémní pozitivní hodnotu [Tourangeau, Rips, Rasinski 2000: 244], naopak významně více než na krátké škále se uchylují k méně silným pozitivním hodnotám (6, případně 7 – 9 vs. 4). Konkrétně body 10 i 9 jsou vnímány jako velmi pozitivní, což nabízí respondentům komfort ponechat si rezervu oproti maximálnímu hodnocení a stále využívanou společnost z hlediska doporučení velmi dobře ohodnotit devítkou. Na krátké, pětibodové škále je naopak druhá nejvyšší hodnota 4 od pozitivního extrému ve vnímání respondentů značně vzdálena (pro některé dokonce již ani není pozitivní, ale neutrální), není tedy možné zvolit stejný postup. To vysvětluje vyšší zastoupení extrémní hodnoty na pětibodové škále. S tím souvisí

i porovnání vnímání extrémních hodnot – na jedenáctibodové škále jsou silnější; zatímco desítka je tedy bezvýhradně pozitivní, pětku volí i ti, kteří nechtějí využívané společnosti „ublížit“ čtyřkou.

Pětibodová škála je na druhou stranu na odpovídání snadnější a především rychlejší – respondenti hodnocení popisují jako rychlé zařazení do kategorie. O volbě odpovědi na krátké škále tolik nepřemýšlí, nepotřebují vyvolat takové množství informací jako v případě hodnocení na dlouhé škále. Z hlediska distribuce odpovědí je na ní (kromě již zmíněného pozitivního extrému) významně více zastoupena střední hodnota (3) než na jedenáctibodové škále (5).

Tyto odlišnosti (větší zastoupení střední a maximální hodnoty na krátké škále, volba méně silných pozitivních hodnot na dlouhé škále) fungují stejně také v rámci podskupin dle socio-demografických charakteristik (rozložení odpovědí se mezi podskupinami systematicky neliší, jsou zde tedy zopakovány stejné odlišnosti jako při pohledu na celek). Obě škály mají schopnost diskriminovat, tedy rozlišit respondenty s různými postoji; za jediný nedostatek autorka považuje volbu maximální hodnoty na krátké škále, která je společná pro ty, kteří chtějí dát maximální hodnocení, i pro ty, kterým nic jiného nezbývá (jejich postoj není bez výhrad, ovšem hodnotu 4 již vnímají jako příliš nízkou pro jeho vyjádření). Z hlediska rozlišení hodnocených společností však obě škály fungují dobře – ukážou, které mají zákazníky ochotné je doporučit a které nikoliv. Porovnání hodnot NPS individuálních společností se však chová poněkud nepředvídatelně – zatímco na úrovni typů institucí (banka, pojišťovna, obchod) je mírně vyšší hodnota získaná na jedenáctibodové škále, u konkrétních společností to není pravidlem – jsou i takové, které mají obě hodnoty srovnatelné nebo dokonce vyšší z pětibodové škály. Mírně se liší také pořadí společností dle NPS hodnot spočítaných z odpovědí na jedenáctibodové a pětibodové škále. Zde mohou roli hrát různé tržní vlivy a charakteristiky, které autorce v rámci tohoto výzkumu zůstávají skryté. V této oblasti je tedy prostor pro další zkoumání. Zde získaná data však ukazují, že NPS z různých délek škál nelze vzájemně porovnávat. Výzkum tak nepotvrzuje přesvědčení Johna Dawese, že výsledky získané pomocí různých délek škál jsou srovnatelné. [Dawes 2002: 2; Dawes 2008: 9 - 10]

Analýza vnímání jednotlivých hodnot respondenty ukazuje, že rozlišení klientských skupin dle jejich odpovědí na promoters (9 + 10), passives (7 + 8) a detractors (0 – 6) je opodstatněné, zvláště je-li přihlédnuto také k tomu, že rozložení odpovědí je (na obou škálách) vychýleno směrem k pozitivnímu konci škály (*positivity*

*bias* [Tourangeau, Rips, Rasinski 2000: 240 - 241]). Hodnoty 9 a 10 jsou vnímány jako velmi pozitivní, 7 + 8 jako nadprůměrné a ostatní jako neutrální či negativní. Jelikož za NPS stojí záměr orientovat se pouze na velmi entusiastické zákazníky, je dle autorky skutečně vhodné definovat tuto skupinu pouze dvěma nejvyššími hodnotami. Negativní část škály (a to jak jedenáctibodové, tak pětibodové) je respondenty využívána málo, hodnoty proto mohou vzájemně splývat – bývají souhrnně označovány zkrátka jako negativní, bez upřesnění vzájemných rozdílů. Konkrétně krajní hodnota 0 na jedenáctibodové škále je ale specifická postavením nuly v číselném systému, některým respondentům tedy přijde zarážející, neobvyklá, nadbytečná. Navíc mohou být respondenti přidáním nuly k rozmezí 1 – 10 zmateni, neschopni rozlišit 5 jako střední hodnotu.

Je zdokumentováno, že respondenti mají tendenci si úlohu zjednodušovat [Tourangeau, Rips, Rasinski 2000: 254], což se potvrdilo také v této studii. V případě jedenáctibodové škály redukuje kognitivní zátěž přirovnáním hodnot k procentům (0 %, 10 % ... 100 %), u pětibodové si vypomáhají školní klasifikací. Ale zatímco procenta jsou ekvivalentním numerickým systémem se vzájemně stejně vzdálenými hodnotami, přirovnání ke školní klasifikaci shledává autorka problematické. Jednak je zde riziko zmatení, a tedy nesprávného odpovídání díky opačnému významu hodnot (na škále využité v dotazníku je 1 nejhorší, zatímco ve škole jde o nejlepší známku), jednak školní známky klasifikace nemusí být vnímána všemi stejně - například v závislosti na tom, jaké známky byli zvyklí ve škole dostávat (3 může být pro někoho průměrná, pro jiného špatná známka).

V neposlední řadě je zajímavý pohled na předměty jednotlivých otázek. V rámci výzkumu byla zjišťována ochota respondentů doporučit využívané společnosti, a to konkrétně jejich hlavní banku (tj. banku, jejíž služby využívají nejčastěji a provádí zde nejvíce transakcí), hlavní pojišťovnu (tj. pojišťovnu, kde mají nejvíce pojistných smluv) a obchod, kde měsíčně utratí nejvíce peněz za potraviny. Odpovědi na otázku týkající se pojišťovny se odlišují od rozložení u zbývajících dvou institucí – na pětibodové i jedenáctibodové škále je u pojišťovny méně zastoupená maximální pozitivní hodnota (10, resp. 5) a více naopak neutrální střed (5, resp. 3). Důvodem se na základě kognitivních rozhovorů zdá být jiný typ vztahu; zatímco s bankou a obchodem jsou zákazníci v běžném denním kontaktu (výběr z bankomatu, internetové/mobilní bankovníctví, nákup potravin), u pojišťovny je kontakt minimální. Někteří zákazníci navíc dosud neměli pojistnou událost, která je pro posouzení pojišťovny zásadní. Volí

tak spíše neutrální hodnocení. S tím souvisí také zmínky respondentů, že dlouhou škálu by lépe využili při hodnocení něčeho konkrétnějšího, například produktu. Je tak důležité si uvědomit, pro které situace je jedenáctibodová škála vhodná a pro které může znamenat riziko. Přináší sice větší variabilitu v datech, což je výhodné pro analýzy závislosti mezi proměnnými, konkrétně u NPS je zde také výhoda srovnatelnosti s jinými zdroji, jelikož jde o originální, běžně využívanou délku škály pro měření ochoty doporučit, může však v případě, že již samotná úloha je pro respondenty náročná (např. hodnotí něco méně známého, dotazník je dlouhý atd.), odpovídání ještě více zkomplikovat a snížit jejich motivaci najít optimální odpověď. Je tedy důležité vždy předem zvážit efekty délky škály a vybrat vhodnější pro konkrétní šetření. Zatímco dlouhá škála může přinést více informací za předpokladu, že jsou ji respondenti schopni využít, tj. při hodnocení něčeho konkrétního, kdy nemusí najednou zvažovat velké množství úhlů pohledu či atributů, či dobře známého (tj. například u zaměstnaneckých výzkumů či zjišťování spokojenosti se službou, produktem či běžně využívanou společností), krátká škála naopak může lépe posloužit, je-li předmětem otázky něco, u čeho očekáváme nízkou míru znalosti respondentů či méně těsný vztah.

Celkově práce přináší přehled o efektech škál se dvěma konkrétními počty bodů (jedenácti a pěti) na proces získávání dat, tj. na kognitivní procesy v myslích respondentů, i na výsledná data. Jde však o bohatou oblast, a jelikož se zde jedná o jednu studii a jeden aplikační příklad, je samozřejmé, že pro další a hlubší porozumění dané problematice je třeba realizovat více opakování, například s využitím různých aplikačních příkladů pro ověření, zda fungují efekty stejně i v případě jiných otázek; mimo jiné také proto, že NPS je zatíženo nejednotným porozuměním otázky respondenty. Další zkoumání se může zaměřit také na ověření hypotézy formulované nad zde prezentovanými kvalitativními daty, a totiž že jedenáctibodová škála znamená pro respondenty vyšší zátěž, což se projevuje také v délce kognitivního procesu, tj. v čase potřebném pro zodpovězení otázky. Při změření délky odpovídání na jednotlivé otázky lze tuto hypotézu testovat na kvantitativních datech. Na míru náročnosti by mohla ukázat také další úroveň analýzy na již dostupných datech, a totiž skrze pořadí hodnocení – pořadí hodnocených institucí při sběru dat rotovalo, lze tedy sledovat, jestli se vzorce odpovídání mění v závislosti na tom, je-li instituce hodnocena v pořadí první, druhá nebo třetí. Předběžná analýza však ukázala, že rozložení odpovědí ani mediány se u žádného z typů institucí dle pořadí významně neliší (dle chí kvadrát testu a Kruskal-

Wallis testu – viz příloha č. 8 na CD), autorka v tomto ohledu neodhalila ani žádné statisticky nevýznamné, ale systematické trendy v rámci výběrového souboru.

Další zajímavou úlohou by bylo zapojení další délky škály, například sedmibodové, která je kompromisem mezi zde testovanými škálami. Autorka pro první krok zvolila dvě kontrastní délky škály, aby bylo možné lépe odhalit případné odlišnosti kognitivních úloh. V dalším kroku by však již bylo možné situaci zkomplikovat a zde nabyté poznatky rozšířit o další délku škály. Poslední podoblastí, kde autorka vidí prostor pro další zkoumání, jsou strategie respondentů při odpovídání. Provedená série kognitivních rozhovorů poskytla dostatek dat pro rozlišení pouze jedné ze strategií (dvoufázové provedení úlohy), data však naznačují, že strategií bude více. Práce je tak příspěvkem ke zkoumání dané problematiky, zejména potom pro její zkoumání v českém prostředí. Neklade si však – a ani nemůže klást – za cíl být vyčerpávající. V průběhu výzkumu vyvstaly další otázky, na které lze dále hledat odpovědi.

## Summary

In this thesis, author tested the influence of rating scale length (number of scale points) on measurement, both in terms of cognitive processes in the minds of respondents and impact on the data. Based on the literature, she formulated and using data obtained from a split ballot experiment and cognitive interviews tested hypotheses related to the comparability of results, distribution of responses and task demands. In order to compare scales with different number of points, various recoding and rescaling techniques were used. Analysis shows that rating on eleven-point and five-point scales means two different cognitive tasks for respondents and that the data differ from each other. No hypothesis has been disproved, but some hasn't been fully confirmed.

Eleven-point scale, which is original for an application example used (Net Promoters Score expressing respondent's willingness to recommend bank, insurance company and grocery shop he/she uses) is slightly more demanding for respondents as they need to think about it more, retrieve more information and as some of them use two-step-strategy of rating – first they select part of scale or range of points based on the initial information retrieved, then consider more information in order to make judgment and map it onto response category. However, they perceive the answer as more accurate



than the one on five-point scale (which is quick classification of company used in category); this corresponds with *information theory*. [Alwin 1997: 322]

Analysis confirms two documented response effects on rating scales – *positivity bias* (on both scales) and *response contraction bias* (on eleven-point scale). [Tourangeau, Rips, Rasinski 2000: 240 – 241, 244] Although quantitative data indicate a tendency that eleven-point scale generates slightly more positive results than the five-point, this trend has not been proved to be statistically significant here. However, distribution of responses differ significantly – on short range scale, midpoint and extreme positive category are selected more often, while on eleven-point scale, milder positive points are used more often (i. e. respondents tend to moderate their answers and avoid extreme rating – *response contraction bias*). These differences work the same within subgroups according to socio-demographic characteristics. Both scales have power to discriminate companies with customers willing and unwilling to recommend them, but comparison of NPS scores for individual companies shows, that each scale produces different results while these differences are not systematic (one time the one from eleven-point scale is higher, second time it is comparable or lower).

As points 9 and 10 on eleven-point scale are perceived as highly positive, 7 and 8 as moderately positive and the rest of the scale as neutral or negative, author consider the distinction of client groups to promoters, passives and detractors to be justified. Here, different market influences, which remain hidden in this research, can play their role. However, these results shows that NPS scores obtained from scales with different length cannot be compared (in contrast with John Dawes belief [Dawes 2002: 2; Dawes 2008: 9 - 10]).

Last but not least, an interesting look is the one at the subject of individual questions. Distribution of responses to the question regarding insurance company differs from the distribution of the remaining two institution (bank, grocery shop) – on five-point as well as eleven-point scale maximum positive value (10, resp. 5) is less represented and midpoint (5, resp. 3) is more represented at insurance company. Different type of relationship seems to be the reason – people are in touch with their bank and grocery shop on daily basis (withdraw money, use internet/mobile banking, buy food), but contact with insurer is minimal. Moreover, some customers haven't had a claim yet which is important for insurance company evaluation. Thus, they select rather neutral rating. This is related also to mentions of some respondents that they would use eleven-point scale better for rating of more specific thing, e. g. product used.

It means that it is important to realize when long scale is useful and when may pose a risk.

## Použitá literatura

Ajzen, Icek. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* [online] 50: 179 – 211 [cit. 6. 4. 2014]. Dostupné z: [file:///C:/Users/Vendula/Downloads/Oct+19+Cited+%231+Manage+THE+THEORY+OF+PLANNED+BEHAVIOR%20\(1\).pdf](file:///C:/Users/Vendula/Downloads/Oct+19+Cited+%231+Manage+THE+THEORY+OF+PLANNED+BEHAVIOR%20(1).pdf).

Albarracín, Dolores, Blair T. Johnson, Mark P. Zanna (eds.). 2005. *The Handbook of Attitudes* [online] [cit. 6. 4. 2014]. Dostupné z: [http://www.google.cz/books?hl=cs&lr=&id=qoJ5AgAAQBAJ&oi=fnd&pg=PR5&dq=handbook+of+attitudes&ots=I3vkn0YE7k&sig=CBVTZWe87OE5VEKq\\_VvtISOflMs&redir\\_esc=y#v=onepage&q=handbook%20of%20attitudes&f=false](http://www.google.cz/books?hl=cs&lr=&id=qoJ5AgAAQBAJ&oi=fnd&pg=PR5&dq=handbook+of+attitudes&ots=I3vkn0YE7k&sig=CBVTZWe87OE5VEKq_VvtISOflMs&redir_esc=y#v=onepage&q=handbook%20of%20attitudes&f=false).

Alwin, Duane F., Jon A. Krosnick. 1991. "The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes." *Sociological Methods & Research* [online] 20 (1): 139 – 181 [cit. 10. 4. 2014]. Dostupné z: <http://smr.sagepub.com/content/20/1/139>.

Alwin, Duane F. 1997. „Feeling Thermometers Versus 7-Point Scales: Which Are Better?“ *Sociological Methods and Research* [online] 25 (3): 318 – 340 [cit. 4. 1. 2014]. Dostupné z: [http://141.213.232.243/bitstream/handle/2027.42/68989/10.1177\\_0049124197025003003.pdf?sequence=2](http://141.213.232.243/bitstream/handle/2027.42/68989/10.1177_0049124197025003003.pdf?sequence=2).

Beatty, Paul C. 2004. „The Dynamics of Cognitive Interviewing.“ Pp. 45 – 66 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.

Beatty, Paul C., Gordon B. Willis. 2007. „Research Synthesis: The Practice of Cognitive Interviewing.“ *Public Opinion Quarterly* [online] 71 (2): 287 – 311 [cit. 4. 1. 2014]. Dostupné z: <http://xa.yimg.com/kq/groups/22882378/1893572222/name/La%20pr%C3%A1ctica%20de%20la%20entrevista%20cognitiva..pdf>.

Bohner, Gerd, Nina Dickel. 2011. „Attitudes and Attitude Change.“ *Annual Review of Psychology* [online] 62: 391 – 417 [cit. 6. 4. 2014]. Dostupné z:

<http://neuron4.psych.ubc.ca/~schaller/Psyc590Readings/BohnerDickel2011.pdf>.

Conrad, Frederick G., Johnny Blair. 1996. „From Impressions to Data: Increasing the Objectivity of Cognitive Interviews.“ Příspěvek přednesený na *1996 Joint Statistical Meetings* [online] [cit. 10. 1. 2014]. Chicago, 4. 8. 1996. Dostupné z:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.4003&rep=rep1&type=pdf>.

Conrad, Frederick G., Johnny Blair. 2004. „Data Quality in Cognitive Interviews: The Case of Verbal Reports.“ Pp. 67 – 87 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.

Dawes, John. 2002. „Five Point vs. Eleven Point Scales: Does It Make A Difference To Data Characteristics?“ *Australasian Journal of Market Research* [online] 10 (1): 1 – 17 [cit. 20. 4. 2014]. Dostupné z: [http://www.johndawes.com.au/page6/files/page6\\_2.pdf](http://www.johndawes.com.au/page6/files/page6_2.pdf).

Dawes, John. 2008. „Do Data Characteristics Change According to the Number of Scale Points Used? An Experiment Using 5 Point, 7 Point and 10 Point Scales.“ *International Journal of Market Research* [online] 50 (1) [cit. 12. 1. 2014]. Dostupné z:

<http://ssrn.com/abstract=2013613>.

DeMaio Theresa J., Ashley Landreth. 2004. „Do Different Cognitive Interview Techniques Produce Different Results?“ Pp. 67 – 87 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.

Eagly, Alice H., Shelly Chaiken. 2007. „The advantages of an inclusive definition of attitude.“ *Social Cognition* [online] 25 (5): 582 - 602 [cit. 6. 4. 2014]. Dostupné z:

<http://eds.a.ebscohost.com.ezproxy.is.cuni.cz/ehost/pdfviewer/pdfviewer?sid=c69c2422-f7da-4ebc-bc09-f7bb0d223f87%40sessionmgr4002&vid=5&hid=4202>.

*Etické zásady činnosti v oboru marketingového výzkumu*. 2001. [online] [cit. 15. 1. 2014] Praha: SIMAR. Dostupné z: <http://simar.cz/assets/media/Clenstvi/guide.pdf>.

Fazio, Russell H. 2007. „Attitudes as object-evaluation associations of varying strength.“ *Social Cognition* [online] 25 (5): 603 – 637 [cit. 6. 4. 2014]. Dostupné z: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2677817/>.

Fienberg, Stephen E., Judith M. Tanur. 1989. „Combining Cognitive and Statistical Approaches to Survey Design.“ *Science, New Series* [online] 243 (4894): 1017 – 1022 [cit. 12. 1. 2014]. Dostupné z: <http://www.jstor.org/stable/1702967?origin=JSTOR-pdf>.

Fowler, Floyd Jackson Jr. 2004. „The Case for More Split-Sample Experiments in Developing Survey Instruments.“ Pp. 173 – 188 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.

Görizt, Anja S. 2004. „Recruitment for Online Access Panels.“ *International Journal of Market Research* [online] 46 (4) [cit. 12. 1. 2014]. Dostupné z: <http://www.goeritz.net/IJMR2.pdf>.

Groves, Robert M. 2004. *Survey Errors and Survey Costs*. Hoboken: John Wiley & Sons, Inc.

IPSOS. 2013. „Češi mnohem raději doporučí svou značku automobilu než banku či pojišťovnu.“ Tisková zpráva z 16. 5. 2013 [online] [cit. 8. 5. 2014]. Dostupné z: [http://www.ipsos.cz/sites/default/files/Tiskov%C3%A1%20zpr%C3%A1va%20Ipsos%2005\\_16\\_2013.pdf](http://www.ipsos.cz/sites/default/files/Tiskov%C3%A1%20zpr%C3%A1va%20Ipsos%2005_16_2013.pdf).

Jabine, Thomas B., Miron L. Straf, Judith M. Tanur, Roger Tourangeau. (eds.) 1984. *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington, D.C.: National Academy Press.

Kirk, Roger E. 2009. „Experimental Design.“ Pp. 23 – 45 in Millsap, Roger E., Alberto Maydeu-Olivares (eds.). *The SAGE Handbook of Quantitative Methods in Psychology* [online] [cit. 12. 1. 2014]. Dostupné z: [http://www.corwin.com/upm-data/29173\\_Millsap\\_Chapter\\_2.pdf](http://www.corwin.com/upm-data/29173_Millsap_Chapter_2.pdf).

Krejčí, Jindřich. 2008. *Kvalita sociálněvědních výběrových šetření v České republice*. Praha: Sociologické nakladatelství.

Keiningham, Timothy L., Bruce Cooil, Tor Wallin Andreassen, Lerzan Aksoy. 2007. „A Longitudinal Examination of Net Promoter and Firm Revenue Growth.” *Journal of Marketing* [online] 71 (3): 39 – 51 [cit. 15. 4. 2014]. Dostupné z: <http://journals.ama.org/doi/abs/10.1509/jmkg.71.3.39>.

Krosnick, Jon A., Charles M. Judd, Bernd Wittenbrink. 2005. „The measurement of attitudes.“ Pp. 21 – 76 in Albarracín, Dolores, Blair T. Johnson, Mark P. Zanna (eds.). *The Handbook of Attitudes* [online] [cit. 6. 4. 2014]. Dostupné z: <http://www.stanford.edu/dept/communication/faculty/krosnick/docs/The%20Measurement%20of%20Attitudes.pdf>.

Norušis, Maria J. 2005. *SPSS 13.0 Guide to Data Analysis*. Englewood Cliffs: Prentice Hall.

Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer. 2004. „Methods for Testing and Evaluating Survey Questions.“ Pp. 1 – 22 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.

Preston, Carolyn C., Andrew M. Colman. 2000. „Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power, and Respondent Preferences.“ *Acta Psychologica* [online] 104: 1 - 15 [cit. 12. 1. 2014]. Dostupné z: <http://rangevoting.org/optinumb.pdf>.

Reichheld, Frederik F. 2003. „The One Number You Need To Grow.“ *Harvard Business Review* [online] [cit. 4. 1. 2014]. Dostupné z: [http://www.thedatashop.co.uk/docs/NetPromoterScore\\_HBR.pdf](http://www.thedatashop.co.uk/docs/NetPromoterScore_HBR.pdf).

Reichheld, Frederik F. 2006. *The Ultimate Question: For Unlocking the Door to Good Profits and True Growth*. Boston: Harvard Business School Press.

Revilla, Melanie, Willem E. Saris, Jon A Krosnick. 2014. „Choosing the Number of Categories in Agree–Disagree Scales.“ *Sociological Methods & Research* [online] 34 (1): 73 – 97 [cit. 7. 5. 2014]. Dostupné z: <https://pprg.stanford.edu/wp-content/uploads/Choosing-the-Number-of-Categories-in-Agree-Disagree-Scales-Revilla-M.-Saris-W.-Krosnick-J..pdf>.

Saris, Willem E., Albert Satorra, Germà Coenders. 2004. „A New Approach to Evaluating the Quality of Measurement Instruments: The Split-Ballot MTMM Design.“ *Sociological Methodology* [online] 34 (1): 311–347 [cit. 15. 1. 2014]. Dostupné z: <http://www3.udg.edu/fcee/professors/gcoenders/pap22.pdf>.

Schwarz, Norbert. 2007. „Attitude Construction: Evaluation in Context.“ *Social Cognition* [online] 25 (5): 638 - 656 [cit. 6. 4. 2014]. Dostupné z: <http://guilfordjournals.com/doi/pdf/10.1521/soco.2007.25.5.638>.

Sniderman, Paul M., Douglas B. Grob. 1996. “Innovations in Experimental Design in Attitude Surveys.” *Annual Review of Sociology* [online] 22: 377 – 399 [cit. 12. 1. 2014]. Dostupné z: <http://www.jstor.org/stable/pdfplus/2083436.pdf?acceptTC=true&acceptTC=true&jpdConfirm=true>.

Tourangeau, Roger. 2004. “Experimental Design Considerations for Testing and Evaluating Questionnaires.” Pp. 209 – 224 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.

Tourangeau, Roger, Lance J. Rips, Kenneth Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.

Vávra, Martin. 2006. „Nesnáze s měřením postojů.“ *SDA Info* [online] 8 (1): 9 - 12 [cit. 6. 4. 2014]. Dostupné z: <http://archiv.soc.cas.cz/download/122/info0601.pdf>

Vinopal, Jiří. 2009. „Situace standardizovaného dotazování z hlediska kognitivních přístupů.“ *Sociologický časopis* [online] 45 (2): 397 – 420 [cit. 5. 4. 2014]. Dostupné z:

[http://sreview.soc.cas.cz/uploads/3e19bb597d035580a764ff158a494fc97b37d295\\_Vinopal2009-2.pdf](http://sreview.soc.cas.cz/uploads/3e19bb597d035580a764ff158a494fc97b37d295_Vinopal2009-2.pdf).

Willis, Gordon B. 1999. „Cognitive Interviewing: A ‚How To‘ Guide.“ Příspěvek přednesený na *1999 Meeting of the American Statistical Association* [online] [cit. 5. 1. 2014]. Dostupné z:

<http://www.dfpp.univr.it/documenti/OccorrenzaIns/matdid/matdid823948.pdf>.

Willis, Gordon B. 2004. „Cognitive Interviewing Revisited: A Useful Technique, in Theory?“ Pp. 23 – 43 in Presser, Stanley, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer (eds.). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken: John Wiley & Sons, Inc.



## **Seznam příloh**

**Příloha č. 1: Dotazník (text)**

**Příloha č. 2: Dotazník – naprogramované verze obou typů škál (obrázek)**

**Příloha č. 3: Scénář rozhovoru (text)**

**Příloha č. 4: Další materiály využívané při rozhovoru (text)**

**Příloha č. 5: Souhlas k využití dat agentury IPSOS (obrázek)**

**Příloha č. 6: Vzor informovaného souhlasu (text)**

**Příloha č. 7: Rozložení odpovědí dle věku a vzdělání (grafy)**

Další přílohy jsou k dispozici na přiloženém CD:

**Příloha č. 8: Výsledky statistických testů (tabulky)**

**Příloha č. 9: Přepisy rozhovorů (text)**

## Přílohy

### Příloha č. 1: Dotazník (text)

**PROG.: ZOBRAZIT TEXT: Nejdříve se budeme věnovat tématu doporučení.**

**A1.** Kterou z následujících bank považujete za svou hlavní banku? Hlavní bankou máme na mysli tu banku, jejíž služby využíváte nejčastěji, tu banku, ve které provádíte nejvíce transakcí.

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

1	Air Bank	
3	Citibank	
4	Česká spořitelna	
5	ČSOB	
6	Equa bank	
7	Fio Banka	
8	GE Money Bank	
9	Hypoteční banka	
10	ING Bank	
11	Komerční banka	
12	LBBW	
13	mBank	
14	Poštovní spořitelna/ERA	
15	Raiffeisenbank	
16	UniCredit Bank	
17	Volksbank/Sberbank	
18	Zuno	
98	Jiná banka	<b>PROG: NEPOKLÁDAT OTÁZKU A6_1</b>
99	<i>Nevím, bez odpovědi</i>	
96	<i>Nevyužívám žádnou banku</i>	

**A2.** Kterou z následujících pojišťoven považujete za svou hlavní pojišťovnu? Hlavní pojišťovnou máme na mysli tu pojišťovnu, u které máte sjednáno nejvíce pojistných smluv.

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

1	Allianz	
2	AXA	
3	Česká podnikatelská pojišťovna	
4	Česká pojišťovna	
5	ČSOB Pojišťovna	
6	Generali	
7	ING Pojišťovna	
8	Kooperativa	
9	MetLife Amcico	
10	Pojišťovna České spořitelny	
11	UNIQA pojišťovna	
98	Jiná pojišťovna	<b>PROG: NEPOKLÁDAT OTÁZKU A6_2</b>
99	<i>Nevím, bez odpovědi</i>	
96	<i>Nevyužívám žádnou pojišťovnu</i>	

**A3.** Ve kterém z následujících obchodů utratíte měsíčně nejvíce za potraviny?

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

1	Albert hypermarket	
2	Albert supermarket	
3	Billa	
4	COOP	
5	Globus	
6	Interspar	
7	Kaufland	
8	Lidl	
9	Makro Cash&Carry	
10	Penny Market	
11	Tesco Expres	
12	Tesco hypermarket	
13	Tesco supermarket	
98	Jiný místní obchod	<b>PROG: NEPOKLÁDAT OTÁZKU A6_3</b>
99	<i>Nevím, bez odpovědi</i>	
96	<i>Nenakupují potraviny</i>	

**PROG.: VÝBĚR JE ROZDĚLEN NÁHODNĚ DO DVOU SKUPIN. RESPONDENTŮM S PŘÍZNAKEM 1 ZOBRAZTE 11-BODOVOU ŠKÁLU, RESPONDENTŮM S PŘÍZNAKEM 2 5-BODOVOU ŠKÁLU. ZNĚNÍ OTÁZKY ZOBRAZTE VŠEM STEJNĚ.**

**A6.** Do jaké míry byste doporučil(a) svým přátelům a známým následující společnost?

**PROG: ROTUJTE MOŽNOSTI, V DATECH VYTVOŘTE PROMĚNNOU OZNAČUJÍCÍ POŘADÍ, V JAKÉM BYLY SPOLEČNOSTI RESPONDENTOVI ZOBRAZENY**

1	<BANKA>	PROG: NAČÍST ODPOVĚĎ Z A1
2	<POJIŠŤOVNA>	PROG: NAČÍST ODPOVĚĎ Z A2
3	<OBCHOD>	PROG: NAČÍST ODPOVĚĎ Z A3

### SKUPINA 1

Svůj názor vyjádřete prosím na jedenáctibodové škále, kde 0 znamená určitě nedoporučil(a) a 10 znamená naopak určitě doporučil(a).

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

0	0 Určitě nedoporučil(a)
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10 Určitě doporučil(a)

### SKUPINA 2

Svůj názor vyjádřete prosím na pětibodové škále, kde 1 znamená určitě nedoporučil(a) a 5 znamená naopak určitě doporučil(a).

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

1	1 Určitě nedoporučil(a)
2	2
3	3
4	4
5	5 Určitě doporučil(a)

**PROG.: X1\_1 A X1\_2 ZOBRAZTE NA JEDNU OBRAZOVKU. OTÁZKY X1 ZOBRAZTE V PŘÍPADĚ, ŽE RESPONDENT HODNOTIL V A6 ALESPŮŇ 2 DRUHY INSTITUCÍ (TJ. POUZE V JEDNOM PŘÍPADĚ SI NEVYBRAL NABÍZENOU BANKU/POJIŠŤOVNU/OBCHOD).**

**X1.** Děkujeme za Vaše odpovědi týkající se doporučení. Byl(a) byste ochoten(a) zúčastnit se zhruba půlhodinového osobního rozhovoru navazujícího na předchozí otázku? Pokud ano, uveďte prosím telefonní číslo a denní dobu, kdy Vás můžeme v pracovní dny s nabídkou rozhovoru kontaktovat.

**X1\_1.** Telefonní číslo, na kterém Vás můžeme kontaktovat

**PROG.: POVOLTE ZÁPIS PRÁVĚ 9 ČÍSLIC**

99	Nepřeji si být kontaktován(a) kvůli navazujícímu rozhovoru

**PROG.: POKUD X1\_1=99, ZAZNAMENEJTE 99 I DO X1\_2.**

**X1\_2.** Denní doba, kdy Vás můžeme v pracovní dny kontaktovat

99	Nepřeji si být kontaktován(a) kvůli navazujícímu rozhovoru

## Příloha č. 2: Dotazník – naprogramované verze obou typů škál (obrázek)



16 %

A6. Do jaké míry byste doporučil(a) svým přátelům a známým následující společnost? Svůj názor vyjádřete prosím na jedenáctibodové škále, kde 0 znamená určitě nedoporučil(a) a 10 znamená naopak určitě doporučil(a).

*POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ*

**Air Bank**

- 0 Určitě nedoporučil(a)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 Určitě doporučil(a)

< Zpět

OK >

Copyright Ipsos 2013



24 %

A6. Do jaké míry byste doporučil(a) svým přátelům a známým následující společnost? Svůj názor vyjádřete prosím na pětibodové škále, kde 1 znamená určitě nedoporučil(a) a 5 znamená naopak určitě doporučil(a).

*POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ*

**Air Bank**

- 1 Určitě nedoporučil(a)
- 2
- 3
- 4
- 5 Určitě doporučil(a)

< Zpět

OK >

Copyright Ipsos 2013

### Příloha č. 3: Scénář rozhovoru (text)

Představení diplomové práce, jejího cíle, účelu rozhovorů + nastínění, jak bude rozhovor probíhat (1. nejdříve si respondent připomene dotazník, znovu vyplní tři otázky na doporučení – na své škále), bude trvat cca 30 min, bude nahráván, nahrávku si pak přepíše a okóduje; k doložení myšlenek v DP použiju i anonymizované citace.

Podpis informovaného souhlasu.

#### Předmět rozhovoru: hodnotící škály

- Výzkumníci v sociálních vědách i ve výzkumu trhu pracují s dotazníky, kde se hodnotí různé věci na stupnicích (spokojenost, atraktivita apod.). Dotazníky nejsou vždy ideální, mohou respondenty unavovat, být příliš náročné, že je potom například vyplní jen v letu, bez většího zamyšlení, něco do nich „naklikají“, aby udělali výzkumníkovi radost, že ho vyplnili. Různé takové případy jsou popsány i v literatuře, je to normální. Snažím se proto najít lepší cestu, jak získat kvalitní informace, a zároveň aby to bylo příjemnější pro respondenty.

#### 1. Mohl/a byste mi na tři otázky týkající se doporučení odpovědět znovu? (Budeme si povídat o Vašem dnešním hodnocení, abyste si to dobře pamatoval/a.)

- *Nejdříve otevřená otázka, bez hodnotící škály.*
- *Ukázat respondentovi otázku s jeho odpovědní škálou. Vždy ať řekne, jaká je jeho hlavní banka a pak odpoví na otázku týkající se hlavní banky atd. Pak již pracovat s těmito odpověďmi. Zkuste pro mě nyní přemýšlet nahlas, provádět mě tím, nad čím u hodnocení přemýšlíte.*

#### 2. Důvody hodnocení + porovnání hodnocení jednotlivých typů institucí

- Proč jste hodnotil/a takto? Proč jste dal/a u banky/pojišťovny/obchodu X? Do jaké míry jste si jistý/á svou odpovědí? Co by se muselo stát, abyste dala vyšší číslo? *Opakovat pro všechny 3 instituce.*
- Porovnání hodnocení jednotlivých institucí:
  - **Různé hodnocení:** Co vyjadřují rozdíly ve Vašem hodnocení? Jak to vnímáte? Porovnával/a jste instituce mezi sebou, když jste odpovídal/a?
  - **Stejně hodnocení:** Jak se to stalo, že jste odpověděl/a na všechny otázky stejně? Je to náhoda nebo to byl záměr, všechny tři instituce Vám připadají stejné?
- Vnímáte svou odpověď pozitivně/negativně/neutrálně? Jsou podle Vás některé body na škále pozitivní, neutrální, negativní...?
- Proč jste dal/a právě tuhle hodnotu a ne vedlejší (proč 8 a ne 7/9)? Jaký je rozdíl mezi těmito sousedními hodnotami? Liší se hodně?

#### 3. Porozumění otázce: Co podle Vás otázka vyjadřuje, na co se Vás ptá? Kdybych se Vás chtěla zeptat na Váš vztah k bance XXX, napadá Vás nějaká formulace, která by to lépe vyjadřovala?

#### 4. Je pro Vás náročné na tuto otázku odpovědět? Nebo je to snadné?

**5. Jaká je Vaše strategie, když na takovou otázku odpovídáte?**

- Jak uvažujete, když si přečtete otázku a podíváte se na možné odpovědi?
- Jakým způsobem si odpověď vyberete?
- Je za tím nějaká logika? Představíte si instituci a podle nějakého klíče jí přiřadíte hodnocení? *Nebo může chtít jen uspokojit výzkumníka, udělat to pro odměnu za dotazník apod.*

**6. Pokud bych Vám položila stejné otázky (tj. do jaké míry byste doporučil/a svou hlavní banku/pojišťovnu/obchod), ale dala Vám tyhle možnosti odpovědí, jaké by to bylo? Ukázat dotazník s alternativní délkou škály.**

- Jak byste odpověděl/a? Kam byste se zařadil/a? A proč?
- Proč jste zvolil/a právě tuhle hodnotu a ne vedlejší (proč 4 a ne 3/5)? Jaký je rozdíl mezi jednotlivými hodnotami? Které vnímáte jako pozitivní, neutrální, negativní?
- Jak se Vám tato stupnice jeví jako alternativa? Je lepší nebo horší? A proč?
- Která je snadnější pro odpovídání?

**7. Porovnání top a bottom boxů:** Porovnejte prosím krajní body obou škál, tj. desítku s pětkou a nulu s jedničkou. Vyjadřují to samé nebo je vnímáte rozdílně? Pokud ano, jak se liší?



**Příloha č. 4: Další materiály využívané při rozhovoru (text)****Pětibodová škála**

**A6.** Do jaké míry byste doporučil(a) svým přátelům a známým následující společnost?

1	<BANKA>
2	<POJIŠŤOVNA>
3	<OBCHOD>

Svůj názor vyjádřete prosím na pětibodové škále, kde 1 znamená určitě nedoporučil(a) a 5 znamená naopak určitě doporučil(a).

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

1	1 Určitě nedoporučil(a)
2	2
3	3
4	4
5	5 Určitě doporučil(a)

**Jedenáctibodová škála**

**A6.** Do jaké míry byste doporučil(a) svým přátelům a známým následující společnost?

1	<BANKA>
2	<POJIŠŤOVNA>
3	<OBCHOD>

Svůj názor vyjádřete prosím na jedenáctibodové škále, kde 0 znamená určitě nedoporučil(a) a 10 znamená naopak určitě doporučil(a).

**POKYN: POUZE JEDNA MOŽNÁ ODPOVĚĎ**

0	0 Určitě nedoporučil(a)
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10 Určitě doporučil(a)

**Příloha č. 5: Souhlas k využití dat agentury IPSOS (obrázek)****SOUHLAS S VYUŽITÍM DAT**

Já, Radek Jalůvka, jednatel společnosti IPSOS s. r. o. se sídlem Topolská 1591, Černošice, svým podpisem níže stvrzuji, že **uděluji své zaměstnankyni Vendule Bláhové, zároveň studentce Sociologie na Institutu sociologických studií na FSV UK<sup>1</sup>, souhlas s využitím níže uvedených dat pro účely její diplomové práce<sup>2</sup>:**

- Data z kontinuálního výzkumu „Vnímání a dopady krize na populaci ČR“, vlna září 2013 (interní číslo projektu 13-006957-01), která jsou majetkem společnosti IPSOS s. r. o., nikoliv dalších subjektů.

Výše uvedená data budou použita jen pro účely výzkumu Venduly Bláhové, jehož výsledky budou publikovány v diplomové práci, případně v odborném článku.

Tento souhlas je vyhotoven ve dvou stejnopisech, každý s platností originálu, z nichž jeden obdrží moje osoba a druhý řešitelka diplomové práce.

V Praze dne 20.1.2014

**IPSOS s.r.o.** ③  
Národní 6, 110 00 Praha 1  
Tel.: +420 226 513 111  
DIČ: CZ26738902 IČ: 26738902

  
Radek Jalůvka

  
Vendula Bláhová

<sup>1</sup> Pro více informací viz <http://www.fsv.cuni.cz>, <http://iss.fsv.cuni.cz>

<sup>2</sup> Vedoucí diplomové práce: Jitř Remr

**Příloha č. 6: Vzor informovaného souhlasu (text)****SOUHLAS S ÚČASTÍ NA ROZHOVORU**

Já, \_\_\_\_\_, níže svým podpisem stvrzuji, že dobrovolně souhlasím:

- s účastí na výzkumu (rozhovoru) pro diplomovou práci Bc. Venduly Bláhové<sup>44</sup>, studentky Sociologie na Institutu sociologických studií na FSV UK<sup>45</sup> (vedoucí diplomové práce Ing. Mgr. Jiří Remr, Ph.D., MBA);
- s nahráváním celého průběhu rozhovoru;
- s tím, že všechny získané údaje budou použity jen pro účely výzkumu;
- s tím, že výsledky výzkumu mohou být publikovány (v diplomové práci, odborném článku atd.), a to včetně anonymních citací z rozhovoru.

Zároveň prohlašuji, že jsem byl/a informována o všech náležitostech týkajících se rozhovoru, tj. o jeho podstatě, cíli, metodách a postupech, které budou při rozhovoru využívány, stejně jako o dalších způsobech zpracování získaných informací.

Tento souhlas je vyhotoven ve dvou stejnopisech, každý s platností originálu, z nichž jeden obdrží moje osoba a druhý řešitelka diplomové práce.

V Praze dne \_\_\_\_\_

Podpis řešitelky diplomové práce: \_\_\_\_\_

Podpis účastníka/účastnice rozhovoru: \_\_\_\_\_

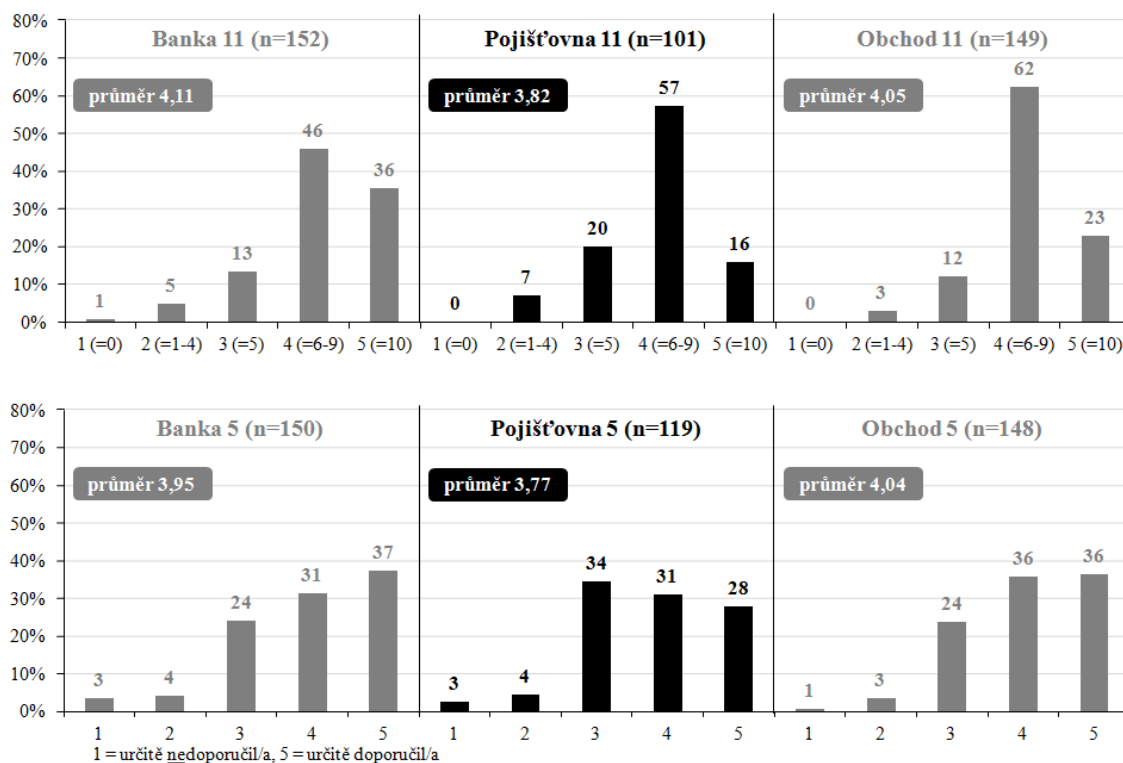
---

<sup>44</sup> Kontakt: [e-mailová adresa a telefonní číslo autorky]

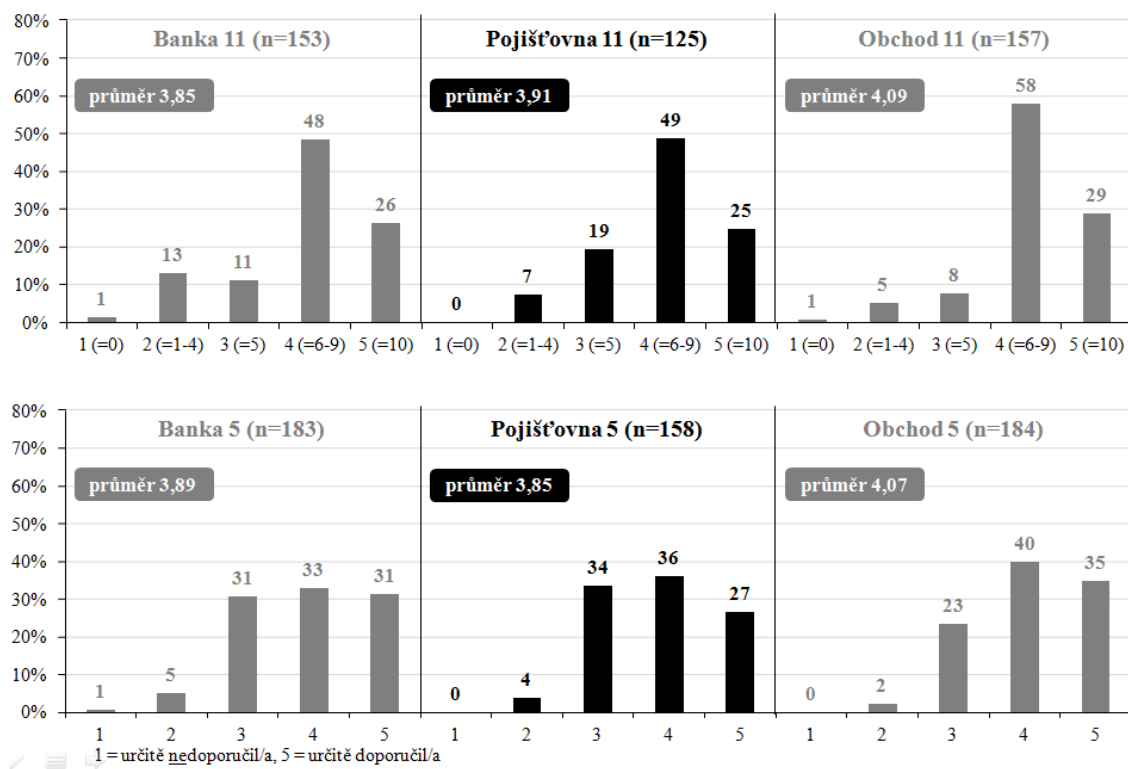
<sup>45</sup> Pro více informací viz <http://www.fsv.cuni.cz>, <http://iss.fsv.cuni.cz>

## Příloha č. 7: Rozložení odpovědí dle věku a vzdělání (grafy)

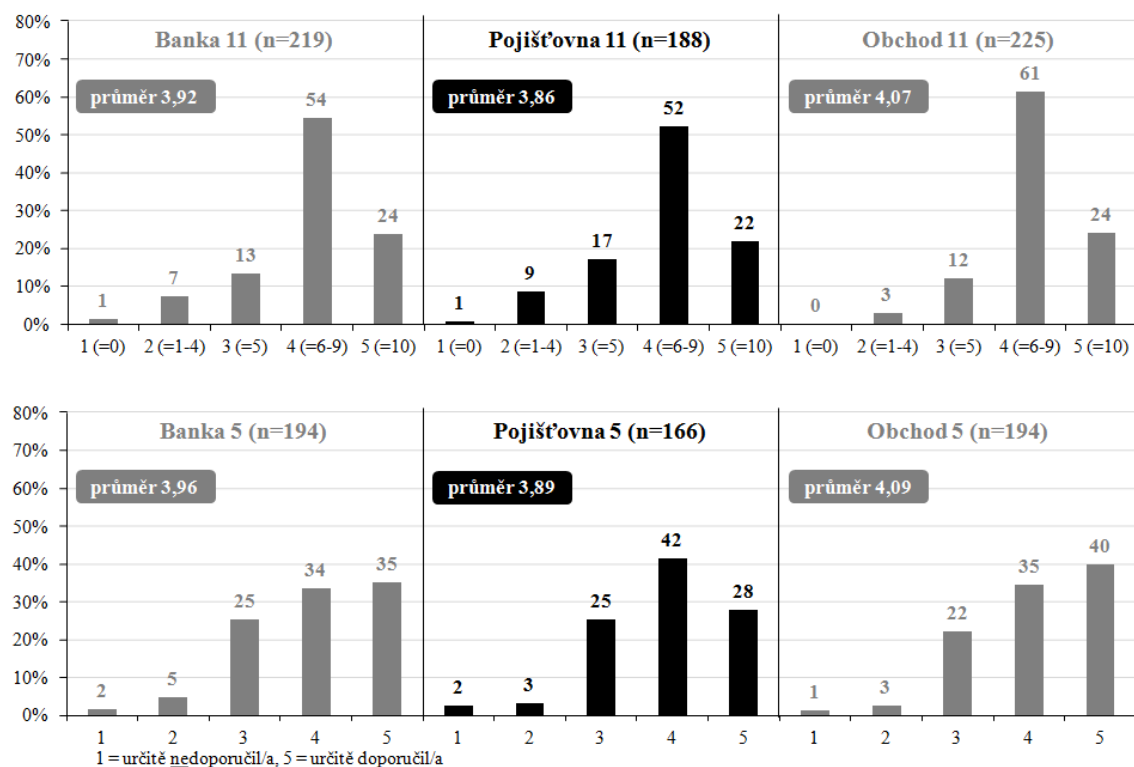
### Graf č. 1.1: Rozložení odpovědí – lidé ve věku do 30 let (relativní četnosti)



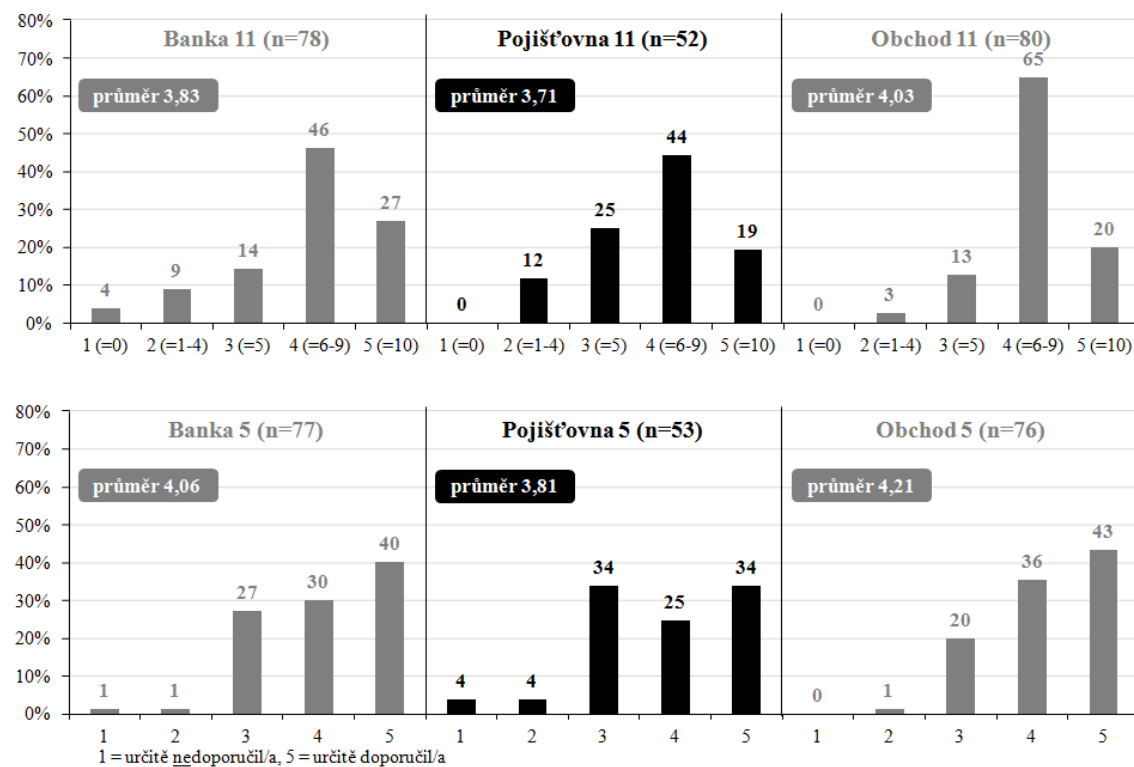
### Graf č. 1.2: Rozložení odpovědí – lidé ve věku 31 – 45 let (relativní četnosti)



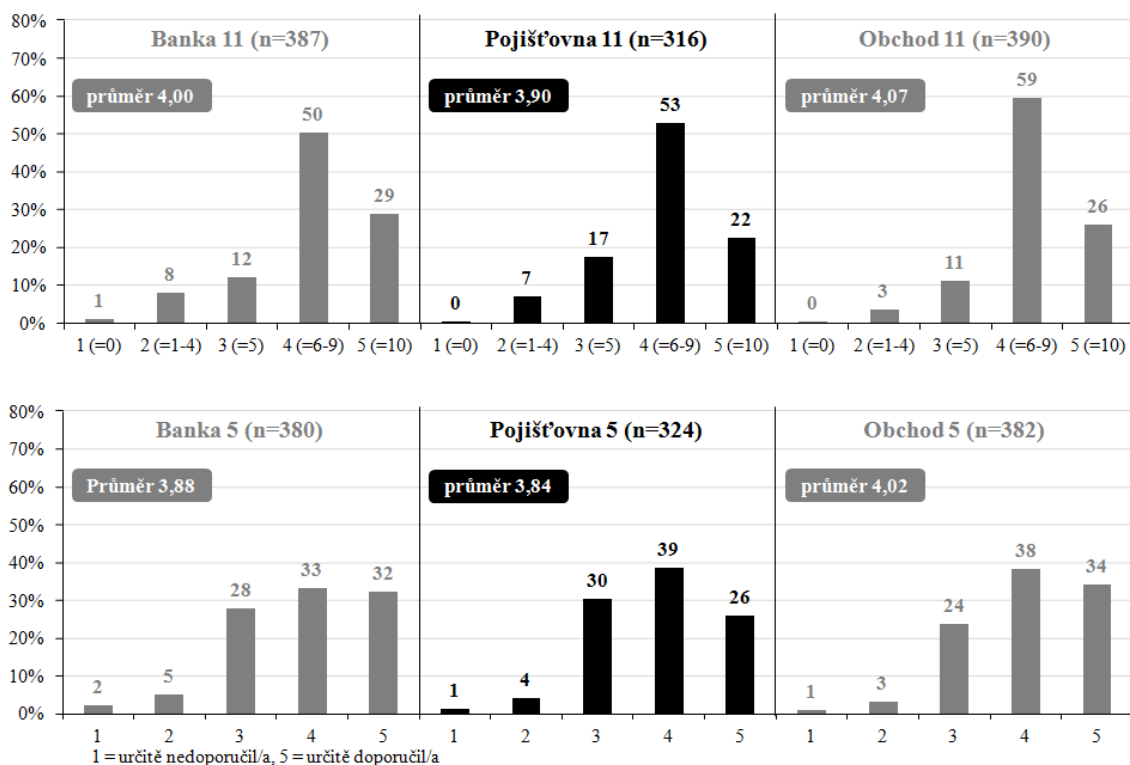
Graf č. 1.3: Rozložení odpovědí – lidé ve věku nad 45 let (relativní četnosti)



Graf č. 2.1: Rozložení odpovědí – lidé se základním vzděláním (relativní četnosti)



**Graf č. 2.2: Rozložení odpovědí – lidé se středním vzděláním a vyučení (relativní četnosti)**



**Graf č. 2.3: Rozložení odpovědí – lidé s vysokoškolským a vyšším odborným vzděláním (relativní četnosti)**

