

UNIVERZITA KARLOVA

Filozofická fakulta

Katedra psychologie



# BAKALÁŘSKÁ PRÁCE

Viktorie Havlíčková

**Zlepšení schopnosti tvorby  
pravděpodobnostních predikcí jako  
prostředek vedoucí k lepšímu  
rozhodování institucí**

**Improving the forecasting ability as a means to  
improve institutional decision-making**

Praha 2020

Vedoucí práce: Mgr. Ing. Marek Vranka

## Poděkování

Děkuji Markovi Vrankovi za vedení celé práce a cenné připomínky. Velký dík patří také Davidovi Janků a projektu Effective Thesis, který mi pomohl s celým procesem hledání a upřesňování tématu.

## Prohlášení

*Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.*

*V Praze dne 24.07.2020*

.....

*Viktorie Havlíčková*

## **Abstrakt**

Téma systematického zlepšování rozhodování, které se tvoří v institucích, je poměrně nová oblast psychologického zájmu s potenciálně velkým dopadem. Jedním z možných přístupů, jak toto téma řešit je snaha o zlepšování pravděpodobnostních predikcí, které rozhodovací subjekty uvnitř institucí dělají. Výzkumy ukazují, že jednou z cest, jak k tomu dojít je skrze trénink zaměřený na zlepšování úsudků a snižování kognitivních zkreslení. Tato práce se snaží shrnout dosavadní literaturu k tématu tvorby predikcí, kognitivních zkreslení a jejich zlepšování. Součástí práce je také návrh výzkumného projektu, jehož cílem je ověřit účinnost tréninku na přesnost predikcí jedince v prostředí reálných institucí.

## **Klíčová slova**

Forecasting, predikce, instituce, rozhodování, kognitivní zkreslení, debiasing, trénink

## **Abstract**

The topic of the systematic improvement of decision-making that takes place in institutions is a relatively new area of psychological interest with a potentially large impact. One possible approach to address this issue is to improve the probability predictions made by decision-makers within institutions. Research shows that one way to do this is through training focused on improving judgment and reducing cognitive bias. This work tries to summarize the existing literature on the topic of forecasting, cognitive bias and their improvement. Part of the work is also the design of a research project, which aims to verify the effectiveness of training on the accuracy of individual predictions in the environment of real institutions.

## **Keywords**

Forecasting, predictions, institutions, decision-making, cognitive biases, debiasing, training

## Obsah

Úvod.....	8
Literárně přehledová část.....	10
1. Forecasting.....	11
1.1. Superforecasting .....	14
1.2. Měření přesnosti forecastingu .....	15
1.3. Význam forecastingu v kontextu vylepšování rozhodování institucí.....	16
2. Teoretický základ nepřesnosti úsudků.....	18
2.1. Systém 1 a systém 2.....	18
2.2. Nadměrná sebedůvěra.....	18
2.3. Chyba zpětného pohledu .....	21
2.4. Necitlivost k rozsahu .....	21
2.5. Konfirmační zkreslení .....	21
2.6. Efekt kotvení.....	22
2.7. Chyba statusu quo.....	23
3. Možnosti zpřesňování pravděpodobnostních predikcí .....	25
3.1. Práce s pravděpodobnostmi .....	25
3.2. Kalibrační trénink .....	26
3.3. Debiasing.....	27
3.4. Pohled zevnitř vs. Pohled zvenčí.....	29
3.5. Fermiho odhadování .....	29
3.6. Záměrné procvičování .....	30
3.7. Agregování úsudků.....	30
Empirická část.....	33
4. Návrh výzkumu vlivu tréninku na forecastingu v prostředí organizací .....	34
4.1. Cíle výzkumu a hypotézy nebo výzkumné otázky .....	35
4.2. Design výzkumného projektu.....	35

4.3.	Výzkumný soubor.....	38
4.4.	Metody zpracování a analýzy dat .....	38
4.5.	Etika výzkumu .....	39
4.6.	Diskuse, limity a možný další vývoj .....	40
	Závěr .....	42
	Seznam použité literatury .....	43
	Seznam tabulek .....	49

## Úvod

V roce 2003 Spojené státy americké podnikly invazi Iráku. Odůvodnění pro tento akt vycházelo z odhadů zpravodajských služeb, které prohlásily, že Irák s velkou jistotou vlastní zbraně hromadného ničení. Ovšem s velkou jistotou může znamenat i 70 % nebo 80 %. Politici si tuto informaci ale vyložili tak, že Irán zbraně hromadného ničení vlastní na 100 % a rozhodli se zaútočit. O další dva roky později se ukázalo, že toto rozhodnutí bylo nejen mylné, ale stálo zemi také biliony dolarů a tisíce lidských životů (Chang et al., 2016). Jak zabránit, aby se podobné chyby nestaly i v budoucnosti?

Velké instituce, včetně státních organizací, dělají často rozhodnutí, která ovlivňují obrovské množství osob. Mají k tomu zdroje i sílu a svět tak mohou změnit k lepšímu, ale i k horšímu. Ovšem rozhodovací procesy, které k tomu využívají, často nejsou optimální. Pokud dokážeme zlepšit pravděpodobnostní úsudky a tím i rozhodování jednotlivců v daných organizacích, můžeme tím zlepšit i celková rozhodnutí, která organizace jako celek dělá.

Tato práce se primárně věnuje tématu tvorby pravděpodobnostních predikcí v kontextu organizací. Klade si za cíl shrnout současné poznatky o tématu pravděpodobnostních úsudků týkajících se budoucnosti, popsat některé empiricky prozkoumané nástroje, které mají za cíl tyto úsudky zpřesňovat. Práce se zároveň snaží ukázat některé aplikace poznatků o tématu, které fungují a v praxi se využívají. Návrh výzkumného projektu spočívá v prozkoumání vlivu tréninku zaměřeného na zlepšování usuzování a snižování kognitivních zkreslení na přesnost predikcí, kteří jedinci dělají v reálném prostředí organizací.

Nejdříve se práce zaměří na samotné téma tvorby předpovědí budoucích událostí (forecasting). V této kapitole bude vymezen a specifikován samotný pojem, budou v něm rozebrány osobnostní vlastnosti pojící se se schopností dělat dobré predikce, bude představena metoda hodnocení pravděpodobnostních úsudků týkajících se budoucích událostí. Poslední část této kapitoly bude věnována významu forecastingu v kontextu vylepšování rozhodování institucí. Druhá kapitola se bude věnovat samotnému teoretickému základu nepřesnosti úsudků. Budou v ní rozebrána některá kognitivní zkreslení negativně ovlivňující tvorbu predikcí se snahou ukázat implikace těchto chyb právě v kontextu



organizací. V třetí kapitole pak budou rozebrány některé možnosti toho, jak tyto zkreslení překonávat a zlepšovat tak samotné predikce.

Práce vychází primárně ze zahraničních odborných zdrojů, zejména z oblastí psychologie, ekonomie a politických věd, a to především kvůli multidisciplinární povaze tématu. Vzhledem k tomu, že se jedná o téma poměrně nové (hlavní výzkum v oblasti zlepšování intuitivních pravděpodobnostních úsudků se odehrává v posledních deseti letech) objevují se v práci i neoborné zdroje v podobě blogových příspěvků či webových stránek věnujících se dané oblasti. V práci je citováno podle normy APA (2010).

## **Literárně přehledová část**

## 1. Forecasting

Forecasting (ve smyslu předpovídání budoucích událostí) fascinoval lidi po tisíce let. V průběhu lidské historie byl oslavován jako znamení seslané od Boha i tvrdě postihován a trestán. Jak řekl prorok Izaiáš: „*o budoucnosti nám povězte, ať poznáme, že jste bohové!*“ (Izaiáš, 41:23). Zatímco dobré předpovědi nám mohou připadat téměř magické, důsledky špatných předpovědí můžou být často katastrofické, o to víc, když se týkají velkého množství osob (Hyndman & Athanasopoulos, 2018).

Forecasting je aktivita, kterou ve svých běžných životech děláme všichni. Na základě našich očekávání o budoucnosti rozhodujeme o tom, jakou práci si vybereme, kde budeme bydlet, jak budeme investovat atd. Ovšem tyto předpovědi neděláme jen sami pro sebe jako jednotlivce. Na jejich výsledcích stojí také spousta rozhodnutí, které dělají organizace, ať už se jedná o oblast politiky, byznysu nebo neziskového sektoru. A rozhodnutí organizací je pouze tak dobré, jak dobrá je předpověď, na které je rozhodnutí založeno (Good Judgement Project, n.d.). Pokud jako jednotlivci nebo organizace dokážeme zlepšit naši schopnost forecastingu, zlepšíme tím i rozhodnutí, která ve svém životě děláme (Tetlock et al., 2017).

Ve statistice a dalších příbuzných oborech se pod pojmem forecasting nejčastěji rozumí určité modelování pomocí statistických metod. Schopnost sestavovat a interpretovat forecastingové modely je důležitým elementem v mnoha doménách. Tato práce se primárně zaměřuje na takzvané úsudkové předpovědi, které jsou v mnoha oblastech jedinou možností forecastingu. Používají se nejčastěji v situacích s velkou mírou nejistoty, kde neexistují historická data, je jich málo nebo pokud se jedná o nový a unikátní typ předpovědi. Slouží například k odpovědím na otázky typu: Kolik uprchlíků opustí Sýrii v následujícím roce? Odpálí Severní Korea jadernou zbraň v následujících třech měsících? Kdo vyhraje další prezidentské volby v Rusku? apod. (Mellers, Stone, Murray, et al., 2015). Jejich použití můžeme často vidět při rozhodnutí týkajících se veřejného sektoru. Například v roce 2012 Australská vláda jako první zakázala používat firemní loga na krabičkách cigaret. Účinnost takové politiky pak musela být předpovídána skrze úsudkovou předpověď, protože neexistovali podobné precedenty, podle kterých by se mohla rozhodnout (Hyndman & Athanasopoulos, 2018).

Forecasting potřebujeme tam, kde panuje nejistota ohledně budoucnosti. Nepotřebujeme předpovídat, jestli ráno vyjde slunce. Stejně tak nepotřebujeme dělat

předpovědi u věcí, které můžeme přímo ovlivnit, jako je například teplota u nás doma. Forecasting je způsob, který může pomoci lidem a organizacím lépe plánovat budoucí události a dělat racionální rozhodnutí (Armstrong, 2001). Forecasting jako metoda spočívá na co nejpřesnějších předpovědích budoucnosti, s ohledem na všechny dostupné informace, včetně dat týkajících se minulosti a znalosti budoucích událostí, které by mohli mít na předpověď dopad (Hyndman & Athanasopoulos, 2018).

Při rozhodování týkajících se budoucích událostí se často spoléháme na předpovědi akademiků, politiků nebo expertů na dané téma. Ovšem výzkumy Philipa Tetlocka (Tetlock & Gardner, 2015), ve kterých hodnotil tisíce odborných předpovědí o ekonomice, akcích, volbách, válkách a dalších celosvětových tématech ukázaly že predikce průměrného odborníka byly zhruba tak úspěšné jako kdyby výsledky hádal úplně náhodně. Odborníci stejně jako ostatní lidé se při svých rozhodováních ohledně budoucích událostí dopouštějí množství zjednodušení a zkreslení, které ovlivňují jejich pravděpodobnostní úsudky (Lichtenstein et al., 1981; Tversky & Kahneman, 1974).

Na schopnost expertních předpovědí se zaměřili také DellaVigna & Pope (2016), kteří zkoumali, jako dobře dokáží experti predikovat výsledky experimentů (skrze předpovídání úsilí účastníků v patnácti rozdílných experimentálních podmínkách). Ukázalo se, že průměr expertních predikcí se blížil experimentálním výsledkům. Korelace mezi průměrnou předpovědí a aktuálním výsledkem experiment byla 0,77. Ovšem výsledky individuálních odpovědí expertů si už tak dobře nevedly a byli překonány průměrnými odpověďmi v 96 %. To poukazuje na problém, které organizace při svém rozhodování často zažívají. Při rozhodování vesměs nemají takový aparát, který by jim umožnil využití agregovaného expertního úsudku a musí se spoléhat na rady jednoho nebo několika expertů, což obecně vede k horším výsledkům, které může ještě ovlivňovat to, že samotní experti jsou ve svých úsudcích špatně zkalibrovaní a podléhají dalším zkreslením.

Obecným problémem rozhodovacích subjektů při tvorbě predikcí je tendence hodnotit problémy jako jedinečné skrze přijetí vnitřního pohledu na daný problém, zanedbávat statistické údaje z minulosti a izolovat současné rozhodnutí od budoucích příležitostí. Důsledkem toho jsou přehnaně optimistické předpovědi, které jsou založené spíše na scénářích a plánech úspěchu než na výsledcích v minulosti (Kahneman & Lovallo, 1993). Schopnost tvorby predikcí ovšem není dar, který člověk buď má, nebo nemá. A i když její úspěšnost ovlivňují některé osobnostní rysy, jako je například inteligence či

kognitivní styly, je to stále schopnost, kterou je možné rozvíjet a zlepšovat (Tetlock & Gardner, 2015). Častým problémem, který se vyskytuje u odborných předpovědí je nejasná přesnost. Předpovědi odborníků v mnoha případech nemají ohraničený časový rámec, problémy mohou nastat i v samotných nejasných formulacích a takové předpovědi jsou pak neověřitelné (Tetlock & Gardner, 2015). Díky tréninku forecastingu člověk dělá predikce, které jsou zhodnotitelné. Tak dostává zpětnou vazbu, skrze kterou se může naučit tomu, jakým způsobem přemýšlí, jaké kognitivní zkreslení ho ovlivňují a jak pracuje s informacemi.

Forecasting, ve smyslu předpovídání budoucích událostí, je často brán jako ryze statistický problém, ale jak už bylo uvedeno výše, předpovědi mohou být zlepšeny i skrze behaviorální intervence. Psychologické intervence, jako je například trénink za účelem snížení kognitivních zkreslení, vytváření predikčních týmů a průběžné sledování výkonu včetně dávání zpětné vazby forecasterům dokáží rapidně zlepšit přesnost předpovědí a snížit chybovost při forecastingu. Rieber (2004) udává, že propojení statistických (například vážení předpovědí podle předchozího záznamu participanta) a psychologických (do nich se řadí například debiasing) poznatků dokázalo přinést ve forecastingovém turnaji nejlepší výsledky. I méně než hodinový trénink dokázal signifikantně zvýšit přesnost odhadů (Chang et al., 2016) a výhody pouze 45 minutového tréninku přetrvávali u forecasterů po dobu 8–10 měsíců. Forecasteri rozhodující v týmu podávají lepší výsledky než forecasteri, kteří se rozhodují na základě rozhodnutí davu a ti jsou přesnější než nezávislí forecasteri.

Největší výzkum forecastingu v posledních letech se děje mimo experimentální laboratoře. Velký vzhled do oblasti forecastingu přinesl projekt IARPA (Intelligence Advanced Research Projects Activity) - série predikčních turnajů sponzorovaná zpravodajskými službami USA. Ty měli za cíl prozkoumat nejlepší strategie pro vytváření co nejpřesnějších intuitivních pravděpodobnostních odhadů týkajících se geopolitických událostí. Turnaje probíhali mezi lety 2011–2014, přičemž každý turnaj trval v roce zhruba 9 měsíců. Účastníci z celého světa v něm udávali subjektivní pravděpodobnostní odhady výskytu různých geopolitických událostí. Celkový počet otázek, na které odpovídali byl v průběhu tří let trvání zhruba 350 (Mellers, Stone, Murray, et al., 2015). Vítězným týmem v této sérii byl [Good Judgement Project](#), který se tak stal primárním zdrojem k tématu forecastingu (nejen u) geopolitických událostí. Mezi čtyři klíčové elementy k přesnému forecastingu podle autorů projektu patří: kombinace tréninkových metod a záměrného procvičování, týmová práce, talent a agregování úsudků (Tetlock & Gardner, 2015).

## 1.1. Superforecasting

Philip Tetlock se svými kolegy dokázal v rámci Good Judgement Projektu identifikovat takzvané „superforecastery“ – lidi, kteří si ve svých předpovědích vedli v průběhu času signifikantně lépe než ostatní. Jejich výsledky nebyly pouhým dílem náhody. V průběhu dvou let, kdy predikční turnaj probíhal u nich nenastala předpokládaná regrese k průměru a jejich výsledky dokázaly být velmi přesné u několika stovek otázek a v široké škále témat (Mellers, Stone, Murray, et al., 2015).

Na základě Good Judgement project přišel Philip Tetlock na to, že superforecasteři se zdají být částečně imunní vůči kognitivním zkreslením. Jejich výsledky u Testu reflexivního myšlení (Cognitive Reflection Test) byly signifikantně lepší než u forecasterů, jejichž předpovědi byli méně přesné. Test reflexivního myšlení je složený ze tří otázek a slouží jako jednoduchý nástroj pro měření zapojení Systému 2 do rozhodování. (Frederick, 2005). Ačkoliv se zdá odpověď na otázky, na které se ptá, po přečtení okamžitě jasná, ke správné odpovědi, musí člověk vědomě odstoupit od svého automatického a intuitivního uvažování a kriticky ji zhodnotit. Superforecasteři také vykazovali menší míru necitlivosti k rozsahu. Ta se projevuje, když při řešení problému není brán v potaz rozsah problému, ale jen samotný problém. Dobří forecasteři tak dokázali přiřadit jiné pravděpodobnosti otázkám jako například „Jaká je pravděpodobnost, že v příštích **pěti** letech vypukne válka v Korei?“ a „Jaká je pravděpodobnost, že v příštích **patnácti** letech vypukne válka v Korei?“, zatímco forecasteři s horšími výsledky přiřazovali oběma událostem zhruba stejné pravděpodobnosti (Tetlock & Gardner, 2015). Lepší forecasteři také vykazují vyšší skóry v hodnotách fluidní a krystalické inteligence a open-mindedness (Mellers et al., 2014). Ještě důležitějším prediktorem než inteligence je pak odhodlání člověka aktualizovat svá přesvědčení a postupně se zlepšovat (Tetlock & Gardner, 2015).

Ohledně vztahu mezi přesností předpovědí a aktualizací se ukázalo, že nejpřesnějších předpovědí dosahovali ti forecasteři, kteří pravidelně aktualizovali své prognózy po malých částech, na základě nových informací, které v průběhu času vyvstávali. Naopak forecasteři, jejichž výsledky byli pravidelně méně přesné, se spíše zakotvili na svých prvotních úsudcích anebo aktualizovali nepravidelně s velkými rozdíly mezi jednotlivými odpověďmi (Atanasov et al., 2020). Záporný vztah byl také nalezen mezi přesností předpovědí a nadměrnou sebedůvěrou (Moore et al., 2017). Účinným prvkem na snížení nadměrné sebedůvěry pak byl trénink a týmové kolaborace.

Philip Tetlock ve své knize Superforecasting (Tetlock & Gardner, 2015) na základě výzkumu a rozhovoru se superforecastery došel k určitému portrétu, který by měl definovat dobré forecastery. Jedná se o lidi opatrné, kteří si jsou vědomi, že nic není jisté. Jsou skromní – uvědomují si, že realita je nekonečně složitá. Jejich pohled na svět je nedeterministický – ví, že to, co se stane v budoucnosti k tomu není předurčeno a nemusí se stát. Co se týče schopností a stylu myšlení dobří forecasteři jsou aktivně otevření. Věří, že domněnky jsou hypotézy k testování a ne poklady, které je potřeba chránit. Dobří forecasteři jsou intelektuálně zvědaví, užívají si hádanky a mentální výzvy, dokážou být introspektivní a sebekritičtí a nedělá jim problém pracovat s čísly.

Jejich metody forecastingu jsou:

- Pragmatické – nejsou zcela oddáni žádné myšlence nebo programu
- Analytické – jsou schopni ne dívat se na problém jen ze špičky svého nosu, ale dokáží zvážit další pohledy
- Vybavené „okem vážky“ – dokáží zhodnotit různé pohledy a slučovat je se svými
- Pravděpodobnostní – dokáží používat a posoudit různé „stupně možná“
- Uváživě aktualizující – pokud se změní fakta, změní svůj názor
- Dobří intuitivní psychologové – jsou si vědomi existence kognitivních a emočních zkreslení a ví, že si musí ověřovat, zda nějakému nepodlehli

Jejich pracovní etika spočívá v růstovém nastavení mysli (Dweck, 2016) – věří, že se dokáží zlepšovat a jsou vytrvalí – rozhodnutí vydržet tak dlouho, jak je třeba. Dobrá zpráva je, že spousta z těchto věcí se dá díky různým technikám natrénovat. Jaké možnosti zpřesňování úsudků a predikcí existují bude rozebráno v třetí kapitole.

## 1.2. Měření přesnosti forecastingu

K měření přesnosti pravděpodobnostních úsudků se nejčastěji používá Brierovo skóre, které vytvořil Glenn W. Brier (Brier, 1950). Původní použití sloužilo k měření miskalibrace u předpovědi počasí. Brierovo skóre měří vzdálenost mezi předpovědí a skutečným výsledkem. Nabývá hodnot mezi 0 a 2, přičemž 2 značí absolutní nepřesnost (nastalo by tedy v situaci, kdy by si byl jedinec výsledkem 100 % jistý a výsledek by nakonec nenastal) (Tetlock & Gardner, 2015). Čím nižšího čísla Brierovo skóre nabývá, tím jsou předpovědi přesnější. Brierovo skóre se vypočítá jako součet čtverců rozdílu mezi výskytem události (událost nastala = 1; událost nenastala = 0) a odpovědí forecastera. (Chang et al., 2016). Skrze umocnění pak Brierovo skóre odměňuje rozhodnou správnost, naopak trestá nadměrnou sebedůvěru (Tetlock et al., 2017).

Výpočet je uveden na tomto případu: Forecaster předpověděl, že existuje 80 % šance, že průměr Dow Jonesova indexu (ukazatel vývoje na akciovém trhu) bude na konci roku vyšší než 20 000 (obráceně tedy předpověděl, že je 20 % šance, že se tak nestane). Protože na konci roku index nebyl vyšší než 20 000, forecasterovo skóre se vypočítá jako  $(0 - 0,8)^2 + (1 - 0,2)^2 = 1.28$  (Tetlock et al., 2017).

### **1.3. Význam forecastingu v kontextu vylepšování rozhodování institucí**

Systematické zlepšování toho, jak se organizace v současné době rozhodují je poměrně zanedbaná oblast. Vylepšování rozhodovacích procesů uvnitř organizací může mít velký dopad na celou společnost. Naopak špatně vytvořené úsudky, které podléhají spoustě kognitivních chyb a zkreslení, pak často můžou vést až ke katastrofickým scénářům. Potřeba zlepšit procesy, na základě kterých se lidé rozhodují, je v dnešní době naléhavější než kdy dříve. V minulosti neexistovala možnost, že by jedno špatné rozhodnutí mohlo zavinit vyhynutí celé lidské rasy. To se ovšem změnilo s objevem atomových zbraní. A můžeme předpokládat, že množství objevů, které budou pro lidstvo představovat existenční riziko se bude stále zvyšovat.

Samotný pojem vylepšování rozhodování v institucionálním kontextu je ovšem velmi široký a vágní na to, aby se dal nějak konkrétněji uchytit. Jess Whittlestone (2017) na svém blogu přináší dva pohledy na to, jak lze ho lze chápat. Jako první způsob uvádí poměrně dobře prozkoumanou a posledních letech populární cestu využití poznatků behaviorálních věd, především ve formě takzvaných „pošťouchnutí“ (Thaler & Sunstein, 2009). Tato pošťouchnutí neslouží přímo k tomu, aby se zlepšili rozhodovací procesy uvnitř organizací, ale spíše aby skrze různé designy, které si organizace osvojí, lidé obecně dělali lepší rozhodnutí. Příkladem implementace těchto poznatků může být například vytvoření Behavioural Insight týmu ve Spojeném království. Ten využívá současných poznatků behaviorálních věd při tvorbě politik, které „pošťouchují“ občany k lepším volbám – k tomu, aby lépe spořili, aby pravidelně odváděli daně, aby žili zdravější život atd. Skrze spoustu malých rozhodnutí tak lze poměrně dobře zlepšovat životy ostatních.

Mimo tento přístup, kde dochází k vylepšování institucionálních struktur, která ovlivňují spoustu malých rozhodnutí a která vedou k dobrému rozhodování můžeme využít poznatky z psychologie i při vylepšování procesů podle kterých jsou rozhodnutí v institucích přijímána. Rozdíl oproti předchozímu přístupu je, že se úsilí nezaměřuje na vylepšování spousty malých rozhodnutí, ale na několik vysoce důležitých. Ty jsou mnohem komplexnější



a složitější na provedení, ale v dlouhodobém horizontu by mohly být nesmírně užitečné a jejich dopad by mohl být mnohem větší než u prvního typu (Whittlestone, 2017b). Do tohoto přístupu můžeme řadit například využívání rozhodování založeného na důkazech (Evidence-Based Decision Making), systematické snižování kognitivních zkreslení u důležitých rozhodovacích subjektů, a také zlepšování predikcí, které organizace dělají.

Využití poznatků z oblasti výzkumu forecastingu v prostředí organizací je velmi široké. Může se týkat zlepšování predikcí samotných zaměstnanců, čemuž se věnuje empirická část této práce. Poznatky o tomto tématu by se také daly využít například tak, že by si organizace nebo vláda sama vytvořila svou jednotku superforecasterů, na které by se mohla obracet v situacích s velkou mírou nejistoty. Jako příklad může sloužit současná situace s onemocněním Covid-19, kdy na začátku propuknutí epidemie byla zásadní rozhodnutí dělána často na základě názoru jednotlivých politických poradců a expertů. A jak výzkumy ukázaly, takováto rozhodnutí se ukazují jako naprosto náhodná (Tetlock, 2017b). Pokud by vláda měla takovýto nástroj k dispozici hned na začátku epidemie, mohla by tvořit rozhodnutí na základě mnohem přesnějších predikcí, než jaké ji mohli poskytnout jednotliví experti.

### **1.3.1. Překážky důležitosti řešit toto téma**

Testovat přesnost svých predikcí by mělo být v zájmu každé organizace. Kdo by nechtěl mít přesnější úsudky a s tím spojené i lepší rozhodnutí a výsledky? Ovšem organizace jsou složeny z lidí, kteří jednají ve vlastním zájmu, který často souvisí se zůstáním v pohodlném statusu quo (Tetlock & Gardner, 2015). Rozhodování institucí je velmi komplexní systém. A tak i kdyby se podařilo identifikovat klíčové metody zlepšující jejich rozhodování, uvedení do praxe by mohlo být téměř nemožné kvůli různým, převážně byrokratickým překážkám. Velká většina velkých organizací má také již aktivní vnitřní procesy a jejich celofiremní změna by pro ně mohla představovat takový náklad, který by nebyli ochotni podstoupit (Whittlestone, 2017a). Jako efektivnější by tedy v některých případech mohlo být soustředit se spíše na jednotlivá témata, které jsou potřeba řešit, a ne na celou systematickou změnu toho, jak organizace svá rozhodnutí dělají.

## 2. Teoretický základ nepřesnosti úsudků

Tato práce by byla bezpředmětná, pokud bychom se všichni chovali podle standardního ekonomického modelu. Měli všechny dostupné informace, uměli posoudit všechny varianty a chovali se zcela racionálně. Tak to ovšem není. Jako lidé podléháme spoustě kognitivních chyb a zkreslení a naše zdroje jsou značně omezené. Tyto chyby mohou vést k relativně malým problémům, jako je špatné investiční rozhodnutí, koupě nesprávného produktu nebo přijetí neodpovídajícího zaměstnance. Projevují se ovšem i ve velkých rozhodnutích jako vládní neefektivnost, bankrot nebo sociální nespravedlnosti (Bazerman & Moore, 2013).

### 2.1. Systém 1 a systém 2

Při přemýšlení a rozhodování využíváme dva rozdílné modely, které Daniel Kahneman (2003) popisuje jako Systém 1 a Systém 2. Systém 1 reprezentuje ty procesy naší mysli, které jsou rychlé, automatické, nevyžadují zvláštní námahu a jsou spíše emočně zaměřené. Pro Systém 1 je kvůli rychlosti jeho zpracování charakteristická určitá necitlivost vůči důkazům. Naopak Systém 2 reprezentuje procesy pomalé, kontrolované, vyžadující kognitivní úsilí a řízené pravidly. Systém 1 je intuitivní, naopak Systém 2 slouží k úvahám a usuzování. Systém 2 slouží ke kontrole aktivit Systému 1. Při velkém množství rozhodnutí nám stačí využívat Systém 1. Bylo by nepraktické uvažovat nad spoustou každodenních aktivit, jako je například ranní hygiena nebo nákup potravin. Ovšem je to Systém 2, který by měl ovlivňovat důležitá rozhodnutí našeho života (Bazerman & Moore, 2013). Chyby, kterých se v rozhodování dopouštíme se častěji objevují, když využíváme náš Systém 1 v situacích, kdy bychom měli využít spíše Systému 2.

### 2.2. Nadměrná sebedůvěra

Daniel Kahneman (2011) nadměrnou sebedůvěru (ang.: *overconfidence*) ve své knize *Myšlení rychlé a pomalé* označil za „nejvýznamnější ze všech kognitivních zkreslení“. Bazerman & Moore (2013) nadměrnou sebedůvěru označují jako matku všech zkreslení. Nadměrná sebedůvěra vede k přeceňování schopností jedince a naopak k podceňování oponenta, složitosti úkolu nebo možného rizika (Johnson & Fowler, 2011). Nadměrná sebedůvěra byla zkoumána v mnoha odvětvích. Velké pozornosti se jí dostalo například v oblasti financí (Malmendier & Tate, 2005; Russo & Schoemaker, 1992; Scheinkman & Xiong, 2003) nebo medicíny (Berner & Graber, 2008; Croskerry & Norman, 2008). Dle De

Bondt a Thaler (1995) je nadměrná sebedůvěra jedním z nejrobustnějších a nejlépe zdokumentovaných zjištění v psychologii. A nejsou ji ušetřeni ani odborníci v různých oblastech (Tetlock, 2017a). U tvůrců politik je tento problém stejný jako u jiných profesionálů. Lidé nadhodnocují predikce týkající se jejich pole expertízy (De Bondt & Thaler, 1995).

### **2.2.1. Typy nadměrné sebedůvěry**

Nadměrná sebedůvěra se dá rozdělit do tří typů - overestimation, overplacement a overprecision (Bazerman & Moore, 2013; Moore & Healy, 2008).

#### **Overprecision**

Overprecision je tendence být si příliš jistý přesností svých úsudků a rozhodnutí, nezajímat se o testování předpokladů a odmítat důkazy naznačující, že by se člověk mohl mýlit. Projevuje se tendencí stanovovat si úzké intervaly spolehlivosti a přílišnou jistotou toho, že má jedinec pravdu. Naše úsudky jsou kvůli overprecision špatně zkalibrované. Přehnaná víra v kvalitu našeho úsudku nás vede k tomu, že se na něho příliš spoléháme, navzdory všem jeho nedostatkům. Jsme si jisti, že známe pravdu. Overprecision se považuje na nejrobustnější formu overconfidence. Overprecision nemusí nutně souviset s pozitivitou nebo optimismem. Lidé mohou být například skálopevně přesvědčeni, že se stanou špatné věci. Také ti, kteří spáchají sebevraždu jsou vesměs jisti, že jejich život se už nezlepší (Moore et al., 2015).

Mezi příčiny vzniku overprecision se řadí například touha zbavit se vnitřní disonance nebo stavu napětí týkající se výběru správného rozhodnutí nebo postupu. Lidé se ve stavu napětí cítí motivováni k úlevě od této disonance, i když to často vyžaduje, aby změnilo, čemu věří. Dále je to určitá tendence věřit spíše lidem, kteří ve svých úsudcích projevují velkou míru jistoty, a tak pokud chceme, aby nám lidé věřili je výhodnější strategií zaujmout tuto cestu. Lidé, kteří projevují větší míru overprecision bývají také častěji povýšeni (Bazerman & Moore, 2013). Na příkladu overprecision jde také vidět, jak spolu jednotlivá zkreslení často souvisí. Míra overprecision může stoupat díky konfirmačnímu zkreslení – tendenci vyhledávat si jen takové informace, kterým jsme věřili již předtím. Kvůli konfirmačnímu zkreslení si tak neustále potvrzujeme představy, které o sobě máme a nekonfrontujeme je s opakem.

#### **Overestimation**

Overestimation je tendence myslet si, že jsme lepší, chytřejší, rychlejší, schopnější, atraktivnější, populárnější atd. než ve skutečnosti jsme. V důsledku toho přeceňujeme, kolik toho v omezeném časovém horizontu dosáhneme nebo věříme, že máme větší kontrolu nad věcmi, než ve skutečnosti máme. Overestimation se projevuje například skrze self-enhancement, iluzi kontroly, chybu plánování (ang: *planning fallacy*) nebo optimistické zkreslení (Bazerman & Moore, 2013). Self-enhancement spočívá v přehnané motivaci vnímat se více pozitivně, než by bylo přesné. Iluze kontroly znamená předpoklad pravděpodobnosti osobního úspěchu nepřiměřeně vyšší než by se jevilo skrze objektivní pravděpodobnost (Langer, 1975). K pocitu kontroly nad nekontrolovatelnými událostmi lidem obecně stačí i povrchní známky kontroly, jako je například možnost zadat si vlastní čísla v loterii.

Chyba plánování se projevuje když odhadujeme, jaký čas budeme potřebovat na splnění určitého úkolu, přičemž dojde téměř vždy k jeho podcenění (Buehler et al., 2002). Chyba plánování se ukazuje také při plánování velkých a komplexních projektů, které jsou náchylné ke komplikacím. Děje se tak také proto, že často nedokážeme předvídat pravděpodobnost výskytu komplikací nebo co vše bude potřeba ke zvládnutí projektu. Toto zkreslení se projevuje jak v našich osobních životech, kdy si například špatně rozvrhneme svůj volný čas, tak v pracovních (Bazerman & Moore, 2013). S chybou plánování souvisí i optimistické zkreslení. Jeho důsledky můžeme vidět v předpovídání toho, co se nám stane v budoucnosti, přičemž přeceňujeme pravděpodobnost toho, že se nám stanou pozitivní události a podceňujeme výskyt špatných událostí. Například můžeme podceňovat možnost, že se rozvedeme nebo vážně onemocníme, ale přeceňujeme věk, kterého se dožijeme, náš úspěch na trhu práce nebo víru v talent našich dětí (Sharot, 2011).

### **Overplacement**

Overplacement je tendence věřit, že jsme určitými způsoby lepší než ostatní, i když ve skutečnosti nejsme (Bazerman & Moore, 2013). Hlavní rozdíl oproti overestimation tedy spočívá v porovnávání s ostatními osobami. Příkladem overplacement je například lepší-než-průměr efekt. Ten se projevuje jako tendence hodnotit se příznivěji, než jako průměrná osoba referenční skupiny se kterou se porovnávám (Alicke & Govorun, 2005). Ve známé studii se výzkumníci ptali řidičů v USA, aby posoudili svoji řidičskou kompetenci oproti ostatním řidičům. 93 % z nich uvedlo, že se považují za lepší než průměrné řidiče (Svenson, 1981).

### 2.3. Chyba zpětného pohledu

Chyba zpětného pohledu (ang: *hindsight bias*) bylo poprvé popsáno v roce 1975 v přelomovém článku Barucha Fischhoffa (Fischhoff, 1975). Ten ukázal, že lidé považují výsledky událostí za pravděpodobnější, pokud se na ně dívají ze zpětného pohledu, když mají k dispozici reálný výsledek, než když mají predikovat stejný výsledek bez znalosti skutečného výstupu. Chyba zpětného pohledu je přesvědčení, že událost je lépe předpověditelná po tom, co se stane její výsledek známým než předtím, když její výsledek ještě známý nebyl. Pokud tedy po tom, co událost nastala lidé udávají, jak dobře by dokázali predikovat událost v době, kdy informaci o výsledku ještě neměli, jejich odhad je zkreslený a ve skutečnosti by dané události s takovou jistotou predikovat nedokázali. Chyba zpětného pohledu nám přináší pocit, že jsme to přeci „věděli celou dobu“ a my si tak myslíme, že svět je více předvídatelný, než ve skutečnosti je. Jeho důsledkem pak může být nadměrná sebedůvěra i v našich dalších úsudcích (Roese & Vohs, 2012a).

### 2.4. Necitlivost k rozsahu

Toto zkreslení popsal ve svých výzkumech Daniel Kahneman (Kahneman, 2003). Typickým příkladem v takovém výzkumu byli otázky, při kterých se participantů ptali na stejné věci pouze s jiným rozsahem. Respondenti například udávali, kolik by byli ochotni zaplatit za prevenci utopení migrujících ptáků, přičemž bylo respondentům udáno jedno ze tří čísel kolik ptáků by mohlo být zachráněno. Pro počet 2 000, 20 000 a 200 000 ptáků se množství, které byli lidé ochotni zaplatit téměř nezměnilo. Lidé nereagovali na množství, které jim bylo předloženo, ale na to, co se jim vybavilo v mysli – obrázek utopeného ptáka. Toto zkreslení se často projevuje v ochotě platit za veřejné statky a ve veřejné politice obecně. Například v oblasti záchrany lidských životů se necitlivost k rozsahu projevuje nelineárním zájmem o individuální životy s rostoucím počtem zasažených osob (Dickert et al., 2015).

Schopnost odolat tomuto zkreslení je pro předpovídání budoucích událostí velmi relevantní. Projevuje se například i ve schopnosti brát do úvahy časový rámeček, který by měl reflektovat určitý „rozsah“ předpovědi (Tetlock & Gardner, 2015).

### 2.5. Konfirmační zkreslení

Konfirmační zkreslení je v literatuře dobře a hojně popsané zkreslení, kterému bylo věnováno mnoho pozornosti. Jak píše Nickerson (1998) hned na úvod svého článku věnovanému tomuto zkreslení: „*Pokud by se někdo pokusil identifikovat jeden*

*problematický aspekt lidského uvažování, který si zaslouží pozornost nad ostatními, konfirmační zkreslení by muselo být mezi kandidáty“.* Konformační zkreslení popisuje tendenci vyhledávat a pracovat pouze s takovými informacemi, které již potvrzují naše předchozí domněnky či názory a také ignorovat ty, které tomu neodpovídají. Výzkumy ukazují, že jakmile si člověk jednou vytvoří názor na určitou věc má tendenci se ho pak držet a obhajovat ho (Nickerson, 1998).

Ve výzkumu věnující se vlivu zkreslení při selektivnímu vystavení politickým informacím se konfirmační zkreslení projevovalo ve větším výběru takových politických zpráv, které byly v souladu s již existujícími politickými postoji jedince než zpráv, které politické postoje jedince spíše zpochybňovaly. Konfirmační zkreslení může u jedinců sloužit k tomu, aby se vyhnuli kognitivní disonanci a tím i psychologickému nepohodlí pramenícího z protichůdných informací (Knobloch-Westerwick et al., 2020). V oblasti politiky se může konfirmační zkreslení projevovat také jako neracionální ospravedlňování politik, ke kterým se daná vláda zavázala. Tuchman (2011) tuto tendenci popisuje na příkladu USA a války ve Vietnamu, kdy USA 16 let setrvalo ve Vietnamu i přes četné důkazy, že toto rozhodnutí není správné.

Konfirmační zkreslení se projevuje také v samotné tvorbě a testování našich hypotéz. Wason (1960) ve svém výzkumu účastníkům zadal úkol, který spočíval v objevení pravidla, podle kterého se řídí sekvence tří čísel 2-4-6. Participanti mohli na papír psát další sekvence tří čísel, které podle nich respektují daní pravidlo a výzkumník jim řekl, zda je sekvence správná nebo ne. Participanti měli tendenci potvrzovat si první jednoduché pravidlo, které je napadlo (například čísla, která stoupají po dvou nebo obecně čísla mezi nimiž je stejný rozdíl). Wasonovo pravidlo bylo přitom pouze sekvence jakýchkoliv tří stoupajících čísel. Jeho závěrem pak bylo, že dobrat se ke správnému výsledku vyžaduje ochotu být otevřený falzifikaci našich vlastních hypotéz a jít tak proti pocitu jistoty, který nám primární hypotézy často přináší.

Konfirmační zkreslení se může projevit také v takzvaném asymetrickém optimismu, což je zkreslení, při kterém jednotlivci více reagují na dobré zprávy než na špatné (Vivaldi & Coville, 2020).

## **2.6. Efekt kotvení**

Daniel Kahneman a Amos Tversky (1974) kotvení popisují jako zkreslení plynoucí z toho, že lidé dospívají k cílovým hodnotám na základě počáteční hodnoty, která je jim

prezentována. Lidé přikládají počáteční hodnotě nepřiměřenou váhu a ta se pak projeví i ve finální odpovědi. Efekt kotvení se projevuje i tam, kde je to zdánlivě iracionální. Kahneman a Tversky ve svém experimentu požádali vysokoškolské studenty, aby se snažili odhadnout, jaké procento afrických zemí je zastoupeno v OSN. Před samotnou odpovědí bylo před studenty roztočeno kolo štěstí, na které padlo číslo od 0 do 100. Ukázalo se, že finální odpověď studentů byla silně ovlivněná předchozí hodnotou, i když se samotné otázky vůbec netýkala. Kotvení bylo od té doby zkoumáno v mnoha dalších studiích a jeho efekt se jeví v oblasti lidského jednání jako extrémně robustní (Furnham & Boo, 2011).

Efekt kotvení je další nežádoucí vlastností, která se při forecastingu objevuje. Projevuje se v tendenci předpovědí blížit se k počátečnímu známému referenčnímu bodu. Častým případem je brát jako referenční bod poslední pozorovanou hodnotu. Forecaster je pak nepatříčně ovlivňován na základě předchozí informace, a proto jí dává v procesu tvorby předpovědi i větší váhu. Efekt kotvení může vést ke většímu konzervatismu a podceňování nových a aktuálnějších informací, a tím vytvářet systematické zkreslení (Hyndman & Athanasopoulos, 2018).

## **2.7. Chyba statusu quo**

Většina rozhodnutí, která ve svých životech děláme má alternativu v podobě statusu quo, tedy nedělat nic nebo pokračovat v předchozím rozhodnutí (Samuelson & Zeckhauser, 1988). Ovšem podlehnout tomuto zkreslení může mít množství negativních dopadů na jedince i společnost (lidé si například nezačnou včas spořit na důchod, pokračují ve splácení nevýhodných půjček, zůstávají ve stejných pracovních pozicích...). Kahneman et al. (1991) chybu statusu quo vnímají jako jednu z implikací averze ke ztrátě, protože případné nevýhody z opuštění statusu quo se jeví jako větší než případné výhody. Důsledek tohoto fenoménu můžeme často vidět například při politických rozhodnutích. Pokud by politický subjekt zavedl intervenci za cílem zlepšit oblast, která relativně dobře funguje vystavil by se tak velkému riziku v případě, že by intervence celkové výsledky zlepšila, naopak potenciální přínosy v takovém případě nejsou tak velké.

Status quo zarámování má signifikantní efekt na finální rozhodnutí jedince (Samuelson & Zeckhauser, 1988). V tomto výzkumu byla participantům dán úkol vybrat si jednu ze dvou nabízených možností, pro jednu skupinu bez zarámování, pro druhou se zarámováním status quo. Participanti, kterým byla při rozhodování jedna možnost ze dvou prezentována jako status quo si tuto možnost vybírali mnohem častěji než participanti,

kterým byli obě dvě možnosti prezentovány bez status quo zarámování. Vliv statusu quo také stoupá spolu s množstvím možností, které jsou subjektům předkládány.

Efekt statusu quo také ovlivňuje policymaking firem, ať už se jedná o soukromé nebo veřejné. Jakmile jsou opatření zavedena, přetrvávají ve své podobě i do budoucna a jejich dopad je málokdy hodnocen (Samuelson & Zeckhauser, 1988). Výzkum Hambrick et al. (1993) ukázal, že se statusem quo pozitivně souvisí i výkonnost firmy. Pokud firma podává dobré výkony mají i její manažeři větší tendenci upínat se ke statusu quo a nezvažovat další možné alternativy.



### 3. Možnosti zpřesňování pravděpodobnostních predikcí

I když je v literatuře hlavní část věnována samotným kognitivním zkreslením, existuje i množství zdrojů věnující se tomu, jak naše úsudky a pravděpodobnostní odhady mohou být zlepšeny. Výzkum (Chang et al., 2016), který experimentálně testoval účinnost různých nástrojů vedoucích ke zlepšování úsudků ukázal, že méně než hodinový trénink dokázal zlepšit přesnost úsudků (měřenou Brierovým skórem) o 6 až 11 % v porovnání s kontrolní skupinou. Většina organizací má možnost signifikantně zlepšit přesnost a efektivnost rozhodovacích procesů právě díky naší znalosti specifických a systematických chyb, kterých se v rozhodování dopouštíme (Bazerman & Moore, 2013). V této kapitole budou rozebrány některé typy, strategie a metody sloužící k překonání kognitivních zkreslení a zlepšení lidského rozhodování.

#### 3.1. Práce s pravděpodobnostmi

V hovorové mluvě často využíváme slova jako pravděpodobně ano, pravděpodobně ne, s malou šancí... Ovšem toto slovní vyjádření si může každý vykládat jinak, jak bylo ukázáno na prvním příkladu v úvodu této práce. Řešením této nejednoznačnosti v našich slovních výpovědích by mohlo být vyjadřovat místo toho pravděpodobnost v číslech. To se ovšem v praxi neseťká s úspěchem (Tetlock & Gardner, 2015). Tetlock tak na základě návrhu zpravodaje CIA Shermana Kenta (Davis, 2002) doporučuje mít jasně stanovené číselné hodnoty pravděpodobnosti, které se budou vázat k jednotlivým slovním výpovědím viz tabulka níže.

**Tabulka 1.** Přiřazení slovních výpovědí k číselně vyjádřeným pravděpodobnostem

MÍRA JISTOTY	OBEČNÁ PRAVDĚPODOBNOST
100 %	Jistý
93 % (+/- 6 %)	Téměř jistý
75 % (+/- 12 %)	Pravděpodobný
50 % (+/- 10 %)	Vyrovnané šance
30 % (+/- 10 %)	Pravděpodobně ne
7 % (+/- 5 %)	Téměř jistě ne
0 %	Nemožný

Zdroj: (Tetlock & Gardner, 2015)

V praxi by zavedení takového standartu v jednotlivých institucích mohlo vést k efektivnějším a více transparentním rozhodnutím. Pokud by například finanční ředitel

dostal od svých analytiků informaci, že se daný scénář finančního plánu téměř jistě nestane, věděl by, že to znamená zhruba 7 % šanci a mohl by tomu přizpůsobit svá rozhodnutí.

### 3.2. Kalibrační trénink

Validita našich pravděpodobnostních úsudků může být měřena skrze kalibraci. Být dobře zkalibrovaný je kritickou vlastností v rozhodovacím procesu a při rozvoji technik pro podporu rozhodování (Lichtenstein et al., 1981). Kalibrace je důležitá při všech pravděpodobnostních úsudcích, ať už se týkají minulosti nebo budoucnosti (Rieber, 2004). Jednotlivec je dobře zkalibrovaný pokud v dlouhodobém horizontu platí, že u tvrzení, kterým přiřadil určitou pravděpodobnost, dojde k naplnění s udanou pravděpodobností. Například polovina tvrzení s přiřazenou pravděpodobností 0.5 by měla být pravda, stejně tak jako 60 % s přiřazenou pravděpodobností 0.6 a všechna tam, kde je si člověk 100 % jistý (Fischhoff et al., 1977). Čili tvrzení, které člověk udělá s 60 % jistotou se v 60 % případech opravdu stane.

Výzkumy ukazují, že sebejistota je velmi nedokonale korelována s přesností (Lichtenstein & Fischhoff, 1980). Dokonalá kalibrace znamená dokonalou korelaci mezi sebejistotou a přesností. Místo toho se stává, že sebejistota není dostatečně citlivá na změnu přesnosti (Moore et al., 2015). Bez tréninku lidé obecně ve svých tvrzeních vykazují nadměrnou sebedůvěru. Zveličují, do jaké míry je to, co vědí, správné. Jejich úsudky ve skutečnosti vypadají tak, že v momentě, kdy by jejich úsudky měly být správné na 70 % jejich skutečný výsledek je pouze 60 % a když si jsou jisti na 90 % ve skutečnosti jsou jejich výsledky správné pouze na 75 % (Fischhoff et al., 1977). To může mít vážné důsledky. Například v akademickém prostředí vede k tomu, že studenti dělají špatná studijní rozhodnutí (podcení čas potřebný k napsání semestrální práce, zapíší si nadměrné množství předmětů...), která se pak stávají překážkami v procesu učení (Dunlosky & Rawson, 2012).

Jednou z metod, kterou lze využít při hodnocení přesnosti úsudků je využití konfidenčních intervalů. (Alpert & Raiffa (1982 v Moore et al., 2015) ve svém experimentu nechali studenty odpovídat na faktické otázky na různých intervalech spolehlivosti. Například měli odhadnout, kolik vajíček bylo v USA vyprodukováno v roce 1965. Ukázalo se, že studenti používali taková rozdělení pravděpodobnosti, které zdaleka nepokrývala správné odpovědi – intervaly, které si stanovili byli příliš úzké. 98 % intervaly spolehlivosti tak obsahovali správnou odpověď pouze v 60 % případech. 50 % intervaly spolehlivosti pak správnou odpověď obsahovali jen v jedné třetině případech.

Při rozhodování by se lidé měli snažit být dobře zkalibrování – domněnky jednotlivce by měli co nejlépe odpovídat realitě, což je předpoklad, který není tak lehké dodržet. Svět totiž většinou vidíme jen ze své perspektivy, která obsahuje zjednodušené předpoklady, copingové mechanismy a zkreslení (Bazerman & Moore, 2013). V reálném světě je trénovat si přesnost našich pravděpodobnostních úsudků těžké. Asi nikdo nemá tabulku, do které by si zapisoval vyřčené úsudky a jejich pravděpodobnosti a díky jejich hodnocení pak zjistil, zda je správně zkalibrováný. A když se na naše rozhodnutí podíváme zpětně, často podléháme chybě zpětného pohledu – zdá se nám, že to co se stalo jsme také předpovídali nebo že to bylo jasné (Roese & Vohs, 2012b). Ovšem mnohé zdroje ukazují, že lidská kalibrace jde zlepšit a natrénovat (Lichtenstein & Fischhoff, 1980; Rieber, 2004)

Právě k tomu slouží kalibrační trénink, který spočívá v přiřazování pravděpodobností k jednotlivým otázkám, následným vyhodnocením a zpětné vazbě a sledováním výkonnosti jedince v průběhu. Hubbard Decision Research, v kterém prošlo skrze kalibrační trénink přes 1 000 jednotlivců ukázal, že perfektní kalibrace po několika hodinách tréninku dosáhlo přes 80 % účastníků (Hubbard & Seiersen, 2016) Limitací tohoto výsledku je, že se nejedná o vědeckou studii (účastníci se na trénink sami přihlásili, neproběhlo tedy rozdělení do výzkumné a kontrolní skupiny, další limitace spočívá například i v předpokládané vysoké motivaci účastníků se v kalibraci cíleně zlepšit).

V současné době existuje na internetu několik online nástrojů, díky kterým se dá kalibrace natrénovat. Jedním z nepropracovanějších je aplikace [Calibrate Your Judgement](#) od Open Philanthropy (Muehlhauser, 2018). Funguje tak, že pokládá otázky, u kterých je odpověď známá a poté se ptá, jak je si člověk svojí odpovědí jistý. Po každé odpovědi dostane člověk zpětnou vazbu na to, jak si vedl a také jsou mu přiřazeny body za to, zda měl otázku správně a také jak jistý si jí byl. Po několika kolech otázek pak aplikace vyhodnotí, jak dobře je člověk zkalibrováný – zda nenadhodnocuje své znalosti nebo se naopak ve svých odpovědích systematicky podhodnocuje<sup>1</sup>.

### 3.3. Debiasing

Debiasing (tedy snižování zkreslení – ang: *bias*) je jednou ze strategií, jak zlepšit naše rozhodování nejen v osobních životech, ale také v oblastech jako je politika, byznys nebo zdravotnictví (Soll et al., 2014). Je určen k použití tam, kde rozhodnutí porušují ekonomický předpoklad racionality – například v situacích, kdy podléháme heuristikám a

---

<sup>1</sup> Podrobný popis skórování odpovědí je uveden v (Greenberg, 2018).

kognitivním zkreslením. Debiasing může být vhodný pro ty, kteří opakovaně podhodnocují nebo nadhodnocují pravděpodobnosti. Mezi metody debiasingu patří například osvojení si strategií jako zvaž protipól nebo zvaž alternativu (Morewedge et al., 2015).

(Chang et al., 2016) dělí výzkum technik týkajících se debiasingu do čtyř skupin – založené na informacích, procesu, feedbacku a formátu. Při debiasingu založeném na informacích (didaktickém) se účastníci pouze dozví o chybách v úsudcích, například co je to overconfidence. Tento typ se však jeví jako nejméně účinný. Debiasing založený na procesu spočívá v ověření, jestli jedinec sám zkreslení nepodléhá, například skrze strategii brát do úvahy i opačné názory. Při technice založené na zpětné vazbě jedinec dostává informace o jeho jednotlivých rozhodnutích a o strategiích k jejich lepšímu zvládnutí. Poskytnutí zpětné vazby při úkolech s pravděpodobnostními úsudky dokáže zlepšit kalibraci i Bayesiánské usuzování (Lichtenstein & Fischhoff, 1980). Kalibrační trénink zmíněný výše tak může zároveň patřit do formy debiasingu. Technika založená na formátu slouží ke zlepšení přesnosti úsudků díky lepšímu strukturování problému, který je pak lépe pochopitelný a zpracovatelný.

Podle (Soll et al., 2014) se debiasing rozděluje podle toho, zda se zkreslení přisuzují buď osobám nebo úkolům. Jako nejlepší způsob debiasingu v prvním případě udávají možnost poskytnout lidem kombinaci tréninku, znalostí a dalších nástrojů, které jim pomohou překonat jejich limity. Debiasing v druhém případě spočívá v modifikaci prostředí tak, aby sám podněcoval lidi k lepším rozhodnutím. Typickým příkladem druhého typu debiasingu mohou být takzvané „nudges“ = pošťouchnutí. Ty značí typ intervence, který určitým způsobem navede osoby k lepšímu rozhodování aniž by jim přitom omezila další výběr (Thaler & Sunstein, 2009).

Morewedge et al. (2015) ve svém výzkumu testovali, zda dokáže jediná tréninková intervence zaměřená na debiasing efektivně zlepšit rozhodování, a to jak okamžitě po tréninku, tak i v dlouhodobém horizontu. Účastníci ve výzkumu buď hráli tréninkovou video hru nebo zhlédli instruktážní video, oboje zaměřené na snížení kognitivních zkreslení. Obě intervence měly na jedince okamžitý efekt (snížení zkreslení o 32 a 19 %), efekt byl pozorován i dva měsíce po absolvování experimentu (24 a 19 %). Podobně jako u kalibračního tréninku zde tedy platí že i jednorázová krátká intervence může mít dlouhodobý dopad na přesnost rozhodování jedince.

### **3.4. Pohled zevnitř vs. Pohled zvenčí**

Při rozhodování o problému můžeme využít pohledu zevnitř nebo zvenčí. Pohled zevnitř spočívá ve zaměření se na daný problém, posouzením jeho plánu a překážek, které by při něm mohly nastat a konstrukci jeho budoucího vývoje. Pohled zvenčí se nezaměřuje na detaily problému, místo toho se věnuje statistikám obdobných případů a porovnáním s referenční třídou (Kahneman & Lovallo, 1993). Osvojení si schopnosti pohledu zvenčí může snížit optimistické zkreslení a pomoci při hodnocení rizik problému. Při intuitivních rozhodováních ovšem lidé využívají primárně pohled zevnitř. Je pro nás totiž přirozené přemýšlet o problému jako jedinečném a zaměřovat se jen na jeho unikátní prvky. Preference pohledu zevnitř je způsobená overconfidence a přehnaným optimismem (Bazerman & Moore, 2013).

Pokud při forecastu instituce využívá pohledu zevnitř zaměřuje se na znalosti a informace o daném případě, na podrobnosti existujícího plánu a na jejich představy o některých překážkách a možnosti jejich překonání. Oproti tomu pohled zvenčí je čistě komparativní a nezaměřuje se na přesné předpovědi plynoucí z informací, které jsou o případě dostupné (Kahneman & Lovallo, 1993). Výsledky z experimentu Kahnemana a Lovalla ukazují, že přesnějších výsledků je dosahováno skrze pohled zvenčí. Typickou otázkou zaměřenou na pohled zvenčí je pak: „Jak často se takové věci dějí v podobných situacích?“ (Tetlock & Gardner, 2015).

Bazerman & Moore (2013) přináší návrh na strategii, kterou je dobré si osvojit, pokud jedinec dlouhodobě podléhá zkreslení pohledu zevnitř a dělá mu problém osvojit si schopnost pohledu zvenčí a tou je přizvat si k rozhodování člověka zvnějšku, který není v problému zainteresovaný. Pokud například manažer tvoří časový a finanční plán projektu, může se sám podívat na trvání a cenu podobných projektů ve firmě nebo si může přizvat kolegu, který na podobných projektech pracoval, aby mu dal potřebné informace.

### **3.5. Fermiho odhadování**

Pojem Fermiho odhad je spojen s fyzikem Enricem Fermi, který mimo jiné proslul tím, že dokázal dělat velmi přesné odhady u problémů, ke kterým neměl mnoho dat a které se často mohly zdát jako neřešitelné. Hlavním pravidlem Fermiho odhadů je rozložit zdánlivě neřešitelný problém na sledovatelné dílčí problémy. Fermiho odhadování spočívá v několika dalších pravidlech a to: rozebrat problém na zjistitelné a nezjistitelné části, přiznat

si nevědomost, odhalit a prozkoumat své domněnky a nebát se splést (Tetlock & Gardner, 2015).

Klasickou otázkou, na které se rozložení na dílčí problémy trénuje (a která byla i nějakou dobu součástí přijímacího pohovoru do Google) je otázka, kterou vymyslel přímo Enrico Fermi: Kolik ladičů piana je v Chicagu? Otázka se na první pohled jeví, jako že k jejímu vyřešení půjde použít nic víc než náhodný odhad, ovšem po rozložení na dílčí otázky, jako je například: Kolik obyvatel má Chicago? Jaké procento domácností může vlastnit piano? Kolik pian dokáže v průměru naladit ladič za den? Jak často je potřeba piana ladit? apod. Rozložením na menší subotázky a následnému výpočtu tak dokáže člověk přijít k mnohem přesnějšímu výsledku než kdyby se jen snažil odhadnout výsledek hlavní otázky samostatně.

### **3.6. Záměrné procvičování**

Záměrné procvičování (z ang: *deliberate practice*) je v literatuře chápáno jako trénink zaměřený na procvičování a zlepšování konkrétních úkonů. Definuje se jako soustředěné a postupné zlepšování dílčích úkolů vyžadující úsilí a přispívající k celkové dovednosti (Ericsson, 2006). Jeho součástí je poskytnutí okamžité zpětné vazby, čas na řešení problému a jeho vyhodnocení a možnost opakování řešení problému za účelem zdokonalení chování. Záměrné procvičování se, oproti klasickému procvičování, nezaměřuje jen na věci, které již víme, jak dělat. Spočívá více v konkrétním a trvalém úsilí dělat to, co dobře nedokážeme nebo co nedokážeme vůbec (Ericsson et al., 2007). Bylo dokázáno, že záměrné procvičování dokáže zlepšovat výkon expertů a jejich aktivní zapojení (Ericsson, 2008).

### **3.7. Agregování úsudků**

O tom, že průměr úsudků mnoha jednotlivců dokáže poskytnout přesnější odpověď na otázku, než jedinci samotní, i když odborníci na dané téma, se přesvědčil již v roce 1906 Francis Galton. Ten ukázal, že přes 800 návštěvníků trhu dokázalo dohromady téměř přesně odhadnout váhu jednoho z volů, kteří se tam nacházeli. A to dokonce lépe než experti na dané téma jako farmáři nebo řezníci (Solomon, 2006).

#### **3.7.1. Predikční trhy a predikční turnaje**

Stejným způsobem jako si v sázkových kancelářích člověk může vsadit na výsledek zápasu může si na predikčním trhu vsadit například na to, kdo se stane našim dalším premiérem nebo jaký bude dopad určité politické intervence. Princip predikčních trhů

spočívá v tom, že jednotlivci sázejí peníze na očekávané výstupy jednotlivých událostí v budoucnu a podle přesnosti jejich úsudků je jim pak vyplácena odměna. Existují důkazy, že agregované výsledky z geopolitických predikčních trhů mohou být přesnější reálnému výsledku než jiné metody tvorby prognóz (Arrow et al., 2008). Predikční trhy se používají také například v predikcích korporátního sektoru (Cowgill & Zitzewitz, 2015) nebo ve vědě pro odhad replikovatelnosti výzkumů (Dreber et al., 2015), kde výsledky predikčního trhu překonaly dotazník individuálních předpovědí. V ideálním případě jsou účastníci na predikčním trhu motivováni skrze tržní zisky k nákupu a prodeji podílu o budoucích událostech. Pokud má některý z účastníků relevantní informaci a dané události, tato informace se brzy propíše i do predikčního trhu. Informace o budoucích událostech jsou průběžně aktualizovány a tržní ceny se pak stávají dobrým odhadem k výsledkům budoucích událostí (Atanasov et al., 2017). Otázkou zůstává, proč se predikční trhy při jejich relativní nenáročnosti nestaly obecně používaným nástrojem pro agregování úsudků. Jako jedna z možností se nabízí problém s právními předpisy některých států, kde je online sázení zakázáno (Arrow et al., 2008). Největším predikčním trhem v současné době je nejspíš [PredictIt](#), který nabízí předpovědi zejména politických a finančních událostí. Účast v některém z predikčních trhů může člověku také pomoci v dosažení lepší kalibrace při forecastingu. Hlavní výhodou je reálný trénink na budoucích událostech, a ne na otázkách, u kterých jsou odpovědi již známy (jako je tomu u klasického kalibračního tréninku).

Na téma predikčních trhů navazuje i studie (Dana et al., 2019), která zjistila, že k vysoce přesným odhadům pravděpodobnosti u geopolitických událostí může místo predikčních trhů posloužit i průzkum osob, pokud si dá ovšem člověk pozor na to, jak agreguje jejich odpovědi. Na úroveň predikčních trhů se dá dostat průzkumem osob v případě, pokud se po zodpovězení přiřadí odpovědím váha podle předešlých predikčních výsledků účastníka.

Predikční turnaje se používají k odhalení, jací jednotlivci, týmy nebo algoritmy přicházejí s přesnějšími pravděpodobnostními odhady na jaká témata (Tetlock et al., 2014). Tetlock navrhuje využívat predikční turnaje například pro depolarizaci zbytečně polarizovaných politických debat a k tomu, abychom dokázali být „kolektivně chytrější“ (Tetlock & Gardner, 2015). Největší výzkum predikčních turnajů a jejich efektivnosti byl sponzorován skrze projekt IARPA (Intelligence Advanced Research Project Activity) mezi roky 2011 až 2014 (Mellers, Stone, Murray, et al., 2015). Oproti Tetlockovému dřívějšímu výzkumu ohledně expertních politických úsudků (Tetlock, 2017), který pouze hodnotil

expertní predikce byl tento projekt skutečně koncipován jako turnaj, který se soustředil na to, jak lze generovat co nejpřesnější pravděpodobnostní úsudky, které byli opět hodnoceny podle Brierova skóre. Skrze tento turnaj byli Tetlock a jeho tým schopni odhalit 4 hlavní psychologické hnací síly přesnosti odpovědí, které jim zaručily lepší výsledky než zbylým týmům, které se projektu IARPA také účastnily. Byly to: a) nábor a udržení lepších forecasterů, b) trénink na snížení kognitivních zkreslení, c) poutavější pracovní prostředí (například ve formě kolaborativní týmové práce) a d) lepší statistické metody pro získání „moudrosti davu“ (Tetlock et al., 2014).



## **Empirická část**

## 4. Návrh výzkumu vlivu tréninku na forecastingu v prostředí organizací

V literatuře je popsáno množství technik, u nichž existují důkazy, že dokáží zlepšit rozhodovací procesy. Jen málo z nich bylo reálně testováno v prostředí organizací. Stejně je tomu u forecastingu. Good Judgement Project ukázal, že trénink ke zlepšení usuzování a snížení kognitivních zkreslení se jeví jako účinná metoda zlepšení pravděpodobnostních úsudků týkajících se budoucnosti. Ovšem výběr do tohoto experimentu probíhal na základě self-selekce a účastníci v tomto projektu byli dobrovolníci, kteří vykazovali velkou míru vnitřní motivace a jednotlivým úkolům věnovali spoustu času. Sedmdesát procent z nich za sebou mělo nějakou formu postgraduálního tréninku, což je v porovnání s běžnou populací řadí mezi vysoce vzdělané (Atanasov et al., 2017). Účastníci vykazovali v průměru v porovnání s běžnou populací také nadprůměrnou inteligenci a vyšší znalost politiky (Mellers et al., 2015). Bylo by tedy zajímavé dále prozkoumat poznatky, které primárně Good Judgement Project přinesl, i z vnitřního pohledu reálných institucí.

Většina predikcí, které jsou v organizacích dělány nejsou tvořeny čistě na základě čísel a dat. Jsou ovlivněny porozuměním nejen základním statistickým argumentům, náchylností ke kognitivním zkreslením, ale také touhou ovlivnit myšlení druhých nebo obavami o svou pověst. Predikce se tak stávají často vágními, aby se z nich jedinec mohl „vykroutit“ pokud by se ukázaly jako špatné. Dobrou zprávou, kterou ukazují výzkumy, je že skrze trénink zaměřený na uvažování a snižování kognitivních zkreslení může být tvorba predikcí ve firmě zlepšena (Schoemaker & Tetlock, 2016). V Good Judgement Project dokázal méně než hodinový trénink zlepšit přesnost úsudků o 6 až 11 % v porovnání s kontrolní skupinou (Chang et al., 2016).

Ne každá instituce má prostředky na to, aby při svých rozhodováních o budoucích událostech využívala složité způsoby získávání přesných predikcí. Ovšem může si osvojit některé nenáročné techniky, jako je právě třeba trénink za účelem snížení kognitivních zkreslení, které mohou zlepšit přesnost úsudků jejich rozhodovacích subjektů a tím potažmo i celé organizace. I malé zlepšení schopnosti organizace dělat přesnější predikce jí může přinést výhodu oproti její konkurenci. Organizace, jejíž predikce budou správné ve třech případech z pěti bude mít stále výhodu oproti té, jejíž predikce budou správné jen ve dvou případech z pěti (Schoemaker & Tetlock, 2016).

#### **4.1. Cíle výzkumu a hypotézy nebo výzkumné otázky**

Výsledky z Good Judgement Project ukázaly, že nejpřesnějších predikcí bylo dosaženo pomocí tréninku, týmování a agregování úsudků. Tento návrh výzkumu se soustředí na první z nich – trénink. Ten se jeví jako poměrně dobře popsáná a nenáročná intervence s potenciálním velkým dopadem. Cílem výzkumu je tedy provést experiment v prostředí fungující organizace, který by ověřoval poznatky týkající se vlivu tréninku na schopnost forecastingu, které přinesl Good Judgement Project. Konkrétně by šlo o význam tréninku na forecasting událostí jednotlivých subjektů v organizaci.

Výzkumnou otázkou je tedy jaký bude mít vliv trénink na schopnost forecastingu v prostředí fungujících organizací?

#### **4.2. Design výzkumného projektu**

Výzkumná data budou získávána při některém z pravidelných školení zaměstnanců nebo bude pro účel výzkumu vytvořeno školení nové. Tak jako tak se budou účastníci účastnit výzkumu ve své pracovní době. Participanti by po příchodu na školení byli náhodně rozdělení do experimentální a kontrolní skupiny. Úkolem experimentální skupiny by bylo nejdříve projít tréninkem zaměřeného na snížení kognitivních zkreslení, poté na dvou až tři měsíční časové ose předpovídat pravděpodobnost výskytu uvedených událostí. Kontrolní skupina by pouze předpovídala pravděpodobnost výskytu daných událostí.

Dopředu je přitom důležité věnovat velkou pozornost samotnému výběru a formulaci otázek. Otázky je nutné pokládat tak, aby se na ně dalo jasně odpovědět a nebyl v nich prostor pro spekulace. Philip Tetlock udává, že každá otázka by měla projít takzvaným „jasnovideckým testem“ (Clairvoyance Test) (P. Tetlock, 2012). Takovou otázkou bychom mohli vzít k jasnovidci a on by na ni mohl dát přesnou odpověď bez doptávání co přesně je danými pojmy v otázce myšleno. Důraz je také potřeba dát na přesné uvedení zdroje, podle kterého bude posuzována správná odpověď. Ten bude uveden vždy u dané otázky (Například pokud by se otázka týkala odhadu kurzu koruny vůči euru v určitému datu bude i ní uvedeno i že se hodnotí podle dat České národní banky, v daném datu a čase).

Samotné otázky budou kombinací všeobecných otázek (například „Získá Pirátská strana ve volbách do Poslanecké sněmovny v roce 2021 více než 15 % hlasů?“, „Bude průměrná teplota v lednu 2021 větší než v lednu 2020?“) a otázek relevantních pro chod firmy. Ty se budou týkat přímo jejich cílů a úkolů souvisejícími s danými pozicemi, a také událostí s přímým dopadem na daný obor. Tyto otázky budou formulovány na základě

konzultace s vyšším manažerem v dané firmě. Otázky budou kladeny na dvou až šestiměsíční časové škále (tj. jejich vyhodnocení bude možné udělat za dva až šest měsíců po experimentu).

Motivace participantů bude zajištěna výši odměny přímo související s přesností jednotlivých predikcí. Například pokud bude Brierovo skóre participanta u dané otázky bude v rozmezí  $<1,5;2>$  dostane za danou otázku 0 Kč, pokud  $<1;1,5>$  dostane 30 Kč, při Brierově skóre  $<0,3;1>$  50 Kč a 100 Kč při skóre v rozmezí  $<0;0,3>$ . Tři nejpřesnější z nich navíc dostanou extra bonus v hodnotě 1500 Kč, 1000 Kč a 500 Kč. Případně se dá použít i sociální motivační faktor ve formě např. oznámení nejlepších forecasterů ve vnitropodnikovém newsletteru apod.

Participantů také stejně jako při tréninku Good Judgement Projektu absolvují Test reflexivního myšlení (Cognitive Reflection Test) (Frederick, 2005). Ten měří schopnost odstoupit od odpovědi svého intuitivního Systému 1, která je zavádějící, a použít Systém 2, díky kterému se jedinec může dobrat správné odpovědi. Vzhledem k tomu, že 3 položky používané v původní verzi testu jsou v současné době poměrně zprofanované a nemusely by měřit daný typ kognitivní schopnosti budou při tréninku použity další 4 položky z rozšířené sedmpoložkové varianty testu (Toplak et al., 2014).

Příklady položek v Testu reflexivního myšlení:

1. Pokud John dokáže vypít jeden barel vody za 6 dní a Mary dokáže vypít jeden barel vody za 12 dní, jak dlouho by jim trvalo, než vypijí jeden barel vody dohromady?  
Dohromady by jim to trvalo \_\_\_\_\_ dny
2. Jerry obdržel jak 15. nejvyšší, tak 15. nejnižší známku ve třídě. Kolik studentů je ve třídě?  
Ve třídě je \_\_\_\_\_ studentů

Dalším dotazníkem, který budou účastníci vyplňovat Racionální versus zkušenostní inventář (Rational versus Experiential Inventory) v jeho zkrácené verzi REI-10 (Norris et al., 1998). Dotazník se skládá ze škál Need for Cognition, která měří schopnost myslet logicky a analyticky a tendenci člověka zapojovat se do a užívat si přemýšlení (Cacioppo & Petty, 1982) a Faith in Intuition, která měří do jaké míry se jedinec spoléhá na svoji vlastní intuici. Participantů budou odpovídat na pětistupňové Likertově škále od naprosto nesouhlasím po naprosto souhlasím.

Příklad položek v Racionálním versus zkušenostním inventáři:

1. Dávám přednost složitějším úkolům před jednoduššími.
2. Mám rád intelektuální výzvy.
3. Moje počáteční dojmy z lidí jsou téměř vždy pravdivé.

Výsledky obou dvou testů budou poté korelovány s průměrným Brierovým skórem jednotlivých účastníků. Součástí této části bude i základní demografický dotazník.

Samotný trénink by spočíval nejdříve v teoretickém úvodu, který bude kopírovat trénink popsany v Moore et al. (2017). V tom by byly participantům představeny některé z metod popsanych v kapitole 3, které by si mohli i vyzkoušet. Například technika Fermiho odhadů by byla představena na otázce podobné této: Kolik ladičů klavírů je v Chicagu?, kde by si vyzkoušeli jak rozložit problém na dílčí části. Dále dostanou informace o zapojení pohledu zvenčí (Kahneman & Lovallo, 1993) do svého uvažování při tvorbě predikcí. Díky tomu by neměli problémy brát jako jedinečné, ale měli by být schopni je porovnat s relevantními událostmi, které se odehráli v minulosti. Po tomto představení by následoval trénink kalibrace na počítačích vycházející z Lichtenstein & Fischhoff (1980) v podobě uvedené v příloze Moore et al. (2017) a mírně upravené do kontextu České republiky. Participantům bude postupně předkládáno 30 otázek na které budou udávat, zda si myslí, že je to pravda nebo lež a následně jak si jsou svojí odpovědí jisti (50 – 100 % s intervaly po desíti %). Při tomto cvičení nebude participantům dovoleno hledat si odpověď na internetu (oproti samotnému forecasting, kde jsou v tom naopak podporováni).

Příklad položek tréninku kalibrace:

1. Příčinou ročních období je vzdálenost země od Slunce.

Pravda/lež

Jak jistá/ý si jste ve správnosti své odpovědi? 50 % ,60 %, 70 %, 80 %, 90 %, 100 %

Po zodpovězení každé otázky bude participantům ukázán správný výsledek. Očekává se, že při prvních otázkách budou participant vykazovat vyšší míru přehnané sebedůvěry, která by se měla postupem tréninku snižovat. Ke konci tréninku by měl být člověk dobře zkalibrovaný – například otázky, které označil jako pravdivé a byl si u nich jistý na 60 % by měli být pravdivé v 60 % případů.

Po tréninku u experimentální skupiny bude následovat samotný forecasting. Kontrolní skupině bude dán rovnou bez tréninku. Design bude podobný jako u tréninku kalibrace. Participanti dostanou otázky týkající se budoucích událostí a budou udávat na kolik % si myslí, že daná událost nastane. Součástí každé odpovědi bude i stručná odpověď s vysvětlením, jak k danému odhadu dospěli. Ta bude sloužit jako dodatečná informace pro výzkumníky, ale také jako kontrola, aby participanti jen bezmyšlenkovitě neudali první hodnotu, která je napadne, ale museli se nad odpovědí hlouběji zamyslet a snažit se k ní dohledat nějaké zdroje.

Analýza pak bude provedena až po uplynutí doby potřebné na zjištění, jaký je výsledek otázky. Vzhledem k jednorázovému charakteru a omezenému času tréninku (předpokládaná doba tréninku a tvorby předpovědi je cca 3 hodiny) výzkum nebude zkoumat roli skupinové interakce na přesnost předpovědi.

### **4.3. Výzkumný soubor**

Odhad potřebné velikosti vzorku stavím zejména na power analýze provedené v programu G\*Power (Faul et al., 2007). Za předpokladu splnění podmínek použití t-testu, hladiny významnosti 0.05 a požadované statistické síly 0.80 (standardně doporučené hodnoty v psychologii) bude pro detekci efektu alespoň střední velikosti ( $d=0.5$ ) nutné získat data způsobilá k analýze od 210 participantů. Dále předpokládám, že data některých participantů nebude z různých důvodů možno využít, nebo naměřené proměnné nemusí splňovat podmínky předpokládaného testu, což by oboje vedlo k poklesu statistické síly, tudíž zamýšlený vzorek preventivně navyšuji na 240. Výzkumným souborem by byli zaměstnanci jedné ze čtyř velkých pražských konzultačních firem (Ernst & Young, Deloitte, KPMG nebo PwC) na pozicích konzultant a manažer. Každá z těchto firem má na podobných pozicích několik stovek zaměstnanců, díky čemuž bude získán dostatečně velký vzorek, což umožňuje rekrutovat adekvátní vzorek v českém prostředí.

### **4.4. Metody zpracování a analýzy dat**

Hodnocení přesnosti předpovědi by bylo hodnoceno skrze Brierovo skóre (Brier, 1950) v upravené verzi pro binární výstupy (Chang et al., 2016). Brierovo skóre může nabývat hodnot od 0 do 2, přičemž menší číslo značí větší přesnost v odpovědi. Vypočítá se jako součet čtverců rozdílů mezi výskytem událostí a odpovědí forecastera. Události, která se vyskytla je přiřazeno číslo 1, události, které se nevyskytla číslo 0. Pokud se tedy například

zeptáme na otázku, zda zítra bude pršet a dostaneme odpověď, že ano, na 90 % a pokud o den později bude opravdu pršet vypočítáme Brierovo skóre takto:

**Tabulka 2.** Příklad výpočtu Brierova skóre

Pravděpodobnost, že bude pršet	Pravděpodobnost, že pršet nebude	Výstup - prší	Brierovo skóre
0.9	0.1	1	$(1 - 0.9)^2 + (0 - 0.1)^2 = 0.02$

Brierovo skóre pro každého participanta by se pak vypočítalo jako průměr Brierových skórů ze všech jeho predikcí po uplynutí doby potřebné k získání jejich odpovědí. Následně by se použil t-test pro porovnání výsledků experimentální a kontrolní skupiny. T-test by byl použit i pro porovnání výsledků v Testu kognitivní reflexivity a REI-10. Dále bude měřen vztah mezi Brierovým skórem a výsledky v Testu kognitivní reflexivity a REI-10 a to pomocí lineární regrese.

Získaná data by mohla být dále mohla sloužit k hlubší exploraci. Zajímavé by mohlo být například pozorovat, zda bude možné ze získaných dat identifikovat osoby, které by se výsledky blížili Tetlockových superforecasterům. Ovšem při interpretaci by musel být dán zřetel na možné dílo náhody, vzhledem k tomu, že participanta budou předpovídat relativně malý počet událostí. Můžeme se také podívat na to, jak by na tom byli výsledky jednotlivců v porovnání s jejich agregovaným úsudkem a jak se agregovaný úsudek bude lišit od nejpřesnějších forecasterů.

#### 4.5. Etika výzkumu

Vzhledem k tomu, že participanta absolvují výzkum u počítačů v rámci běžného školení ve své pracovní době nevyvstává zde žádná závažná etická otázka. Jejich účast na experimentu bude dobrovolná a budou ho moct kdykoliv ukončit.

Jedním z možných problémů by ovšem mohlo být to, že polovině zaměstnanců bude trénink nabídnut a druhé polovině ne. Ta tak bude ochuzena o možnost snížit některá z kognitivních zkreslení a případně zlepšit své pravděpodobnostní úsudky týkající se budoucnosti. Toto rozdělení je ovšem nutné abychom mohli zjistit, zda daný trénink v reálném prostředí opravdu funguje. Jako řešení se jeví nabídnout kontrolní skupině trénink ve stejné podobě jako u experimentální potom, co ukončí samotné predikce.

#### 4.6. Diskuse, limity a možný další vývoj

Návrh výzkumu vychází z prozkoumaných poznatků o vlivu kognitivních zkreslení na lidský úsudek a tréninku zaměřeného na snížení těchto zkreslení na schopnost tvorby pravděpodobnostních predikcí. Snaží se získané poznatky prozkoumat z hlediska aplikace do prostředí reálných organizací a zároveň ověřit, zda jejich výsledný efekt bude podobného rázu. Využití tréninku za účelem snížení kognitivních zkreslení a tím i zlepšení pravděpodobnostních úsudků jedinců se jeví jako poměrně nenáročná intervence. Pokud by se v prostředí organizací projevil její pozitivní vliv její implementace z hlediska technické a časové náročnosti do prostředí dalších firem relativně jednoduchá. Návrh neřeší otázku, zda by na takovou intervenci samotné firmy přistoupili. Jednoduchost implementace se tak ztrácí v kontextu byrokracie a toho, jak těžké je prosadit změny v organizaci, což by mohlo být tématem samostatné práce. I kdyby měl tréninkový modul pozitivní vliv na přesnost pravděpodobnostních předpovědí nelze předpokládat, že by tato intervence mohla vyřešit celkový problém týkající institucionálního rozhodování. Ovšem může sloužit k hlubšímu prozkoumání tématu, zejména díky aplikaci do reálného prostředí organizací.

Většina teoretických poznatků o forecastingu, z kterých vychází i návrh tohoto výzkumu vznikla v rámci jednoho, i když obrovského, projektu. Není tedy dobré brát jeho výstupy jako obecně platné zákonitosti. Je možné, že mi při zpracování této práce unikly důležité výzkumy a poznatky týkající se tématu, které můžou mít na tvorbu predikcí také velký vliv. Další z limitací návrhu spočívá v tom, že jde o jednorázovou a krátkou intervenci. Výzkum neměří dlouhodobý efekt kalibračního tréninku, vzhledem k tomu, že participanté dělají predikce hned po jeho absolvování. Možným řešením by bylo dát participantům další set otázek po určité časové době po absolvování tréninku. Problémem by také mohla být nízká motivace participantů, kterou by nezvýšila peněžní ani sociální incentiva. Zde by řešením mohlo být navrhnout participantům vyšší odměnu. Velkým rizikem samotného spuštění výzkumu může být i nezájem organizace zapojit se do něj. Mohlo by totiž dojít ke zjištění, že predikce jejich zaměstnanců jsou velmi nepřesné a nespolehlivé, což je informace, kterou by firma nemusela být ochotná přiznat.

Otázkou zůstává případná reálná možnost aplikace poznatků pro prostředí organizací, která by si také zasloužila pozornost. Pokud by se vliv tréninku na schopnost forecastingu ukázal a instituce by byli ochotné takovýto trénink svým zaměstnancům poskytovat mohl by se stát například součástí pravidelného vzdělávání, které je ve většině organizací již zakotveno.



Dalšími prvky, které zlepšují přesnost predikcí, jsou dle Good Judgement Project týmování a agregování úsudků. To by bylo možné prozkoumat dalším rozšířením navrhnutého výzkumu, který by byl mnohem více časově náročný, ale mohl by přinést další zajímavé poznatky o reálném rozhodování v organizacích.

## Závěr

Důsledky rozhodnutí institucí po celém světě mají velký vliv na množství osob. Rozhodovací procesy, ze kterých vychází, jsou ovšem ovlivněné množstvím kognitivních chyb a zkreslení. Jejich snížení u rozhodovacích subjektů v daných organizacích by mělo vést k lepšímu rozhodování organizace jako celku. Tato práce představila jeden z prostředků, jak zlepšit rozhodovací procesy uvnitř institucí a rozhodnutí, které instituce jako celek dělají. Tím je zlepšení schopnosti tvorby pravděpodobnostních predikcí (forecastingu) skrze trénink zaměřený na zlepšení lidského usuzování a snížení kognitivních zkreslení.

Ke špatným úsudkům často dochází, když využíváme intuitivní myšlení v situacích, kdy by bylo potřeba využít spíše analytické myšlení. Výzkumy ukazují, že mezi kognitivní zkreslení a heuristiky nejvíce ovlivňující přesnost predikcí patří nadměrná sebedůvěra, chyba zpětného pohledu, necitlivost k rozsahu, konfirmační zkreslení, efekt kotvení a chyba statusu quo. Sílu těchto systematických chyb a zkreslení jde ovšem snižovat, a to například skrze trénink kalibrace, debiasing, osvojení si pohledu zvenčí apod.

Součástí práce je i návrh výzkumného projektu, který se snaží ověřit efektivitu tréninku na přesnost forecastingu v prostředí reálných institucí. První část tréninku spočívá v podání teoretických informací o chybách v usuzování a kognitivních zkresleních s představením strategií, jak je lépe překonávat. Druhá část je více praktická a zaměřuje se na kalibrační trénink, pomocí kterého by měla být snížena nadměrná sebedůvěra. Samotný forecasting je pak koncipován jako souhrn všeobecných otázek a otázek relevantních pro danou instituci, které se následně vyhodnocují pomocí Brierova skóre.

Nástroje na snížení kognitivních zkreslení a zlepšení úsudků jsou v literatuře poměrně dobře popsány. Zkoumání vlivu jejich tréninku na schopnost forecastingu je však téma poměrně nové, ve kterém je stále ještě velký prostor pro další bádání. Stejně tak je to u tématu systematického vylepšování rozhodnutí, které organizace dělají. Vzhledem k tomu, že se jedná o velmi interdisciplinární a široké téma je jasné, že tato práce postihuje pouze jeho malou část. Je to také mladé, rozvíjející se téma a je tak možné, že současné poznatky o něm budou brzy překonány. Budoucí výzkum věnující se tomuto tématu, by se tak mohl zaměřit na další průzkum a testování technik, které již existují a které se pro toto téma jeví jako slibné. Další cesta může spočívat i v objevování nových technik, které by mohli zlepšit institucionální rozhodování.

## Seznam použité literatury

APA (2010). *Publication Manual of the American Psychological Association* (6th Ed.). Washington, DC: American Psychological Association.

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The Self in Social Judgment*, 1, 85–106.

Alpert, M., & Raiffa, H. (1982). A Progress Report on the Training of Probability Assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 294–305). Cambridge University Press.

Anders Ericsson, K. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine*, 15(11), 988–994.

Armstrong, J. S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Springer Science & Business Media.

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., Neumann, G. R., Ottaviani, M., Schelling, T. C., Shiller, R. J., Smith, V. L., Snowberg, E., Sunstein, C. R., Tetlock, P. C., Tetlock, P. E., ... Zitzewitz, E. (2008). The Promise of Prediction Markets. *Science*, 320(5878), 877–878. <https://doi.org/10.1126/science.1157679>

Atanasov, P., Rescober, P., Stone, E., Swift, S. A., Servan-Schreiber, E., Tetlock, P., Ungar, L., & Mellers, B. (2017). Distilling the wisdom of crowds: Prediction markets vs. prediction polls. *Management Science*, 63(3), 691–706.

Atanasov, P., Witkowski, J., Ungar, L., Mellers, B., & Tetlock, P. (2020). Small steps to accuracy: Incremental belief updaters are better forecasters. *Organizational Behavior and Human Decision Processes*, 160, 19–35.

Bazerman, M. A., & Moore, D. A. (2013). *Judgment in Managerial Decision Making*, 8th Edition | Wiley. Wiley.Com. <https://www.wiley.com/en-us/Judgment+in+Managerial+Decision+Making%2C+8th+Edition-p-9781118065709>

Berner, E. S., & Graber, M. L. (2008). Overconfidence as a cause of diagnostic error in medicine. *The American Journal of Medicine*, 121(5), S2–S23.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1), 1–3.

Buehler, R., Griffin, D., & Ross, M. (2002). *Inside the planning fallacy: The causes and consequences of optimistic time predictions*.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116.

- Chang, W., Chen, E., Mellers, B., & Tetlock, P. (2016). Developing expert political judgment: The impact of training and practice on judgmental accuracy in geopolitical forecasting. *Judgment and Decision Making*, *11*(5), 509–526.
- Cowgill, B., & Zitzewitz, E. (2015). Corporate prediction markets: Evidence from google, ford, and firm x. *The Review of Economic Studies*, *82*(4), 1309–1341.
- Croskerry, P., & Norman, G. (2008). Overconfidence in clinical decision making. *The American Journal of Medicine*, *121*(5), S24–S29.
- Dana, J., Atanasov, P., Tetlock, P., & Mellers, B. (2019). Are markets more accurate than polls? The surprising informational value of “just asking.” *Judgment and Decision Making; Tallahassee*, *14*(2), 135–147.
- Davis, J. (2002). *Sherman Kent and the Profession of Intelligence Analysis*. 17.
- De Bondt, W. F. M., & Thaler, R. H. (1995). Chapter 13 Financial decision-making in markets and firms: A behavioral perspective. In *Handbooks in Operations Research and Management Science* (Vol. 9, pp. 385–410). Elsevier. [https://doi.org/10.1016/S0927-0507\(05\)80057-X](https://doi.org/10.1016/S0927-0507(05)80057-X)
- DellaVigna, S., & Pope, D. (2016). *Predicting Experimental Results: Which Experts Know What*. Working paper.
- Desvousges, W. H., Johnson, F. R., Dunford, R. W., Boyle, K. J., Hudson, S. P., & Wilson, K. N. (1993). Measuring natural resource damages with contingent valuation: Tests of validity and reliability. *Contingent Valuation: A Critical Assessment*, *91*, 91–93.
- Dickert, S., Västfjäll, D., Kleber, J., & Slovic, P. (2015). Scope insensitivity: The limits of intuitive valuation of human lives in public policy. *Journal of Applied Research in Memory and Cognition*, *4*(3), 248–255.
- Dreber, A., Pfeiffer, T., Almenberg, J., Isaksson, S., Wilson, B., Chen, Y., Nosek, B. A., & Johannesson, M. (2015). Using prediction markets to estimate the reproducibility of scientific research. *Proceedings of the National Academy of Sciences*, *112*(50), 15343–15347.
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students’ learning and retention. *Learning and Instruction*, *22*(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Dweck, C. (2016). What having a “growth mindset” actually means. *Harvard Business Review*, *13*, 213–226.
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge Handbook of Expertise and Expert Performance*, *38*, 685–705.
- Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, *85*(7/8), 114.

- Evidence-Based Forecasting Techniques For Private & Public Sectors*. (n.d.). Good Judgment Project. Retrieved July 9, 2020, from <https://goodjudgment.com/who-we-serve/>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). *G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences*. *Behavior Research Methods*, 39, 175-191.
- Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 552–564. <https://doi.org/10.1037/0096-1523.3.4.552>
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42.
- Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1), 35–42.
- Greenberg, S. (2018). Calibration Scoring Rules for Practical Prediction Training. *ArXiv:1808.07501 [Stat]*. <http://arxiv.org/abs/1808.07501>
- Hambrick, D. C., Geletkanycz, M. A., & Fredrickson, J. W. (1993). Top executive commitment to the status quo: Some tests of its determinants. *Strategic Management Journal*, 14(6), 401–418.
- Hubbard, D. W., & Seiersen, R. (2016). *How to measure anything in cybersecurity risk*. John Wiley & Sons.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
- Johnson, D. D., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, 477(7364), 317–320.
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review*, 93(5), 1449–1475. <https://doi.org/10.1257/000282803322655392>
- Kahneman, D. (2011). *Thinking, fast and slow* (p. 499). Farrar, Straus and Giroux.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1), 193–206.
- Kahneman, D., & Lovallo, D. (1993). Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. *Management Science*, 39(1), 17–31. <https://doi.org/10.1287/mnsc.39.1.17>

- Knobloch-Westerwick, S., Mothes, C., & Polavin, N. (2020). Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication Research*, 47(1), 104–124.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311.
- Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior and Human Performance*, 26(2), 149–171. [https://doi.org/10.1016/0030-5073\(80\)90052-5](https://doi.org/10.1016/0030-5073(80)90052-5)
- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1981). *Calibration of Probabilities: The State of the Art to 1980* (PTR-1092-81-6). DECISION RESEARCH EUGENE OR. <https://apps.dtic.mil/docs/citations/ADA101986>
- Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance*, 60(6), 2661–2700.
- Mellers, B., Stone, E., Atanasov, P., Rohrbaugh, N., Metz, S. E., Ungar, L., Bishop, M. M., Horowitz, M., Merkle, E., & Tetlock, P. (2015). The psychology of intelligence analysis: Drivers of prediction accuracy in world politics. *Journal of Experimental Psychology: Applied*, 21(1), 1.
- Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., Chen, E., Baker, J., Hou, Y., & Horowitz, M. (2015). Identifying and cultivating superforecasters as a method of improving probabilistic predictions. *Perspectives on Psychological Science*, 10(3), 267–281.
- Mellers, B., Ungar, L., Baron, J., Ramos, J., Gurcay, B., Fincher, K., Scott, S. E., Moore, D., Atanasov, P., Swift, S. A., Murray, T., Stone, E., & Tetlock, P. E. (2014). Psychological Strategies for Winning a Geopolitical Forecasting Tournament. *Psychological Science*, 25(5), 1106–1115. <https://doi.org/10.1177/0956797614524255>
- Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2), 502–517. <https://doi.org/10.1037/0033-295X.115.2.502>
- Moore, D. A., Swift, S. A., Minster, A., Mellers, B., Ungar, L., Tetlock, P., Yang, H. H., & Tenney, E. R. (2017). Confidence calibration in a multiyear geopolitical forecasting competition. *Management Science*, 63(11), 3552–3565.
- Moore, D. A., Tenney, E. R., & Haran, U. (2015). Overprecision in Judgment. In *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 182–209). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118468333.ch6>
- Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 129–140.
- Muehlhauser, L. (2018, December 14). *New web app for calibration training*. Open Philanthropy. <https://www.openphilanthropy.org/blog/new-web-app-calibration-training>

- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.
- Norris, P., Pacini, R., & Epstein, S. (1998). The rational-experiential inventory, short form. *Unpublished Inventory*. University of Massachusetts at Amherst.
- Rieber, S. (2004). Intelligence Analysis and Judgmental Calibration. *International Journal of Intelligence and CounterIntelligence*, 17(1), 97–112.  
<https://doi.org/10.1080/08850600490273431>
- Roese, N. J., & Vohs, K. D. (2012a). Hindsight bias. *Perspectives on Psychological Science*, 7(5), 411–426. <https://doi.org/10.1177/1745691612454303>
- Russo, J. E., & Schoemaker, P. J. (1992). Managing overconfidence. *Sloan Management Review*, 33(2), 7–17.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7–59.
- Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6), 1183–1220.
- Schoemaker, P. J. H., & Tetlock, P. E. (2016, May 1). Superforecasting: How to Upgrade Your Company’s Judgment. *Harvard Business Review*, May 2016.  
<https://hbr.org/2016/05/superforecasting-how-to-upgrade-your-companys-judgment>
- Schoemaker, P. J., & Tetlock, P. E. (2016). Superforecasting: How to upgrade your company’s judgment. *Harvard Business Review*, 94, 72–78.
- Sharot, T. (2011). The optimism bias. *Current Biology*, 21(23), R941–R945.
- Soll, J. B., Milkman, K. L., & Payne, J. W. (2014). *A User’s Guide to Debiasing* (SSRN Scholarly Paper ID 2455986). Social Science Research Network.  
<https://papers.ssrn.com/abstract=2455986>
- Solomon, M. (2006). Groupthink versus the wisdom of crowds: The social epistemology of deliberation and dissent. *The Southern Journal of Philosophy*, 44(S1), 28–42.
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2), 143–148.
- Tetlock, P. (2012, December 6). *How to Win at Forecasting* | *Edge.org* [Interview].  
<https://www.edge.org/conversation/win-at-forecasting>
- Tetlock, P. E. (2017). *Expert political judgment: How good is it? How can we know?-New edition*. Princeton University Press.
- Tetlock, P. E., Mellers, B. A., Rohrbaugh, N., & Chen, E. (2014). Forecasting tournaments: Tools for increasing transparency and improving the quality of debate. *Current Directions in Psychological Science*, 23(4), 290–295.

- Tetlock, P. E., Mellers, B. A., & Scoblic, J. P. (2017). Bringing probability judgments into policy debates via forecasting tournaments. *Science*, 355(6324), 481–483.
- Tetlock, P., & Gardner, D. (2015). *Superforecasting: The Art and Science of Prediction* (01 edition). Cornerstone Digital.
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20(2), 147–168.
- Tuchman, B. W. (2011). *The march of folly: From Troy to Vietnam*. Random House.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Vivalt, E., & Coville, A. (2020). *How Do Policymakers Update Their Beliefs?* 47.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12(3), 129–140.
- Whittlestone, J. (2017a, August 25). *Improving institutional decision-making*. 80,000 Hours. <https://80000hours.org/problem-profiles/improving-institutional-decision-making/>
- Whittlestone, J. (2017b, September 30). *Two ways of “improving decision-making.”* <https://jesswhittlestone.com/blog/2017/9/30/two-ways-of-improving-decision-making>



## **Seznam tabulek**

Tabulka 1. Přirazení slovních výpovědí k číselně vyjádřeným pravděpodobnostem.....	25
Tabulka 2. Příklad výpočtu Brierova skóre .....	39