

VILNIAUS UNIVERSITETAS

Adolfas Mackonis

GERIAUSIO PAAIŠKINIMO IŠVEDIMO PATIKIMUMO RIBOS

Daktaro disertacijos santrauka  
Humanitariniai mokslai, filosofija (01 H)

Vilnius, 2011

Disertacija rengta 2007-2011 metais Vilniaus universitete

Mokslinis vadovas:

prof. dr. Marius Povilas Šaulauskas (Vilniaus universitetas, humanitariniai mokslai, filosofija – 01 H)

Konsultantė:

doc. dr. Nijolė Radavičienė (Vilniaus universitetas, humanitariniai mokslai, filosofija – 01 H)

Disertacija ginama Vilniaus universiteto Filosofijos mokslo krypties taryboje:

Pirmininkas

prof. habil. dr. Evaldas Nekrašas (Vilniaus universitetas, humanitariniai mokslai, filosofija – 01 H)

Nariai:

prof. dr. Jūratė Rubavičienė (Vilniaus pedagoginis universitetas, humanitariniai mokslai, filosofija – 01 H)

doc. dr. Aldis Gedutis (Klaipėdos universitetas, humanitariniai mokslai, filosofija – 01 H)

dr. Algirdas Degutis (Lietuvos kultūros tyrimų institutas, humanitariniai mokslai, filosofija – 01 H)

doc. dr. Stanislovas Leonas Norgėla (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P)

Oponentai:

prof. dr. Peeter Mürsepp (Talino technologijos universitetas, Estija, humanitariniai mokslai, filosofija – 01 H)

doc. dr. Albinas Plėšnys (Vilniaus universitetas, humanitariniai mokslai, filosofija – 01 H)

Disertacija bus ginama viešame Filosofijos mokslo krypties tarybos posėdyje 2011 m. liepos 1 d. 16 val. Vilniaus universiteto Filosofijos fakulteto 204 auditorijoje

Adresas: Universiteto g. 9/1, LT-01513 Vilnius, Lietuva

Disertacijos santrauka išsiuntinėta 2011 m. gegužės 31 d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje

VILNIUS UNIVERSITY

Adolfas Mackonis

INFERENCE TO THE BEST EXPLANATION: THE LIMITS OF  
TRUTH-CONDUCTIVENESS

Summary of the Doctoral Dissertation  
Humanities, Philosophy (01 H)

Vilnius, 2011

The dissertation was prepared at Vilnius University during 2007-2011

Research supervisor:

Prof. Dr. Marius Povilas Šaulauskas (Vilnius University, Humanities, Philosophy – 01 H)

Counsellor:

Assoc. Prof. Dr. Nijolė Radavičienė (Vilnius University, Humanities, Philosophy – 01 H)

Dissertation is going to be defended at a public session of the Scientific Council of Philosophy, Vilnius University:

Chairman:

Prof. Dr. Evaldas Nekrašas (Vilnius University, Humanities, Philosophy – 01 H)

Members:

Prof. Dr. Jūratė Rubavičienė (Vilnius Pedagogical University, Humanities, Philosophy – 01 H)

Assoc. Prof. Dr. Aldis Gedutis (Klaipėda University, Humanities, Philosophy – 01 H)

Dr. Algirdas Degutis (Lithuanian Culture Research Institute, Humanities, Philosophy – 01 H)

Assoc. Prof. Dr. Stanislovas Leonas Norgėla (Vilnius University, Physical Sciences, Informatics – 09 P)

Opponents:

Prof. Dr. Peeter Mürsepp (Tallinn University of Technology, Humanities, Philosophy – 01 H)

Assoc. Prof. Dr. Albinas Plėšnys (Vilnius University, Humanities, Philosophy – 01 H)

The official defense of the dissertation will be held on 1st of July, 2011, at the Department of Philosophy, Vilnius University, room 204, at 4 p.m.

Address: Universiteto St 9/1, LT-01513 Vilnius, Lithuania

The summary of doctoral dissertation was circulated on 31th of May, 2011. The dissertation is available at the Vilnius University Library.

Disertacija analizuoja geriausio paaiškinimo išvedimą (GPI) – mokslinių teiginių išvedimo principą.

**Tyrimo aktualumas.** Mokslininkai taiko GPI ir kurdami mokslines teorijas, ir siekdami jas pagrįsti. Mokslininkai kuria tokias teorijas, kurios paaiškintų stebimus reiškinius, derėtų su jau pripažintu moksliniu žinojimu, būtų paprastos ar unifikuotų skirtingų reiškinių paaiškinimus; argumentai už teorijos pripažinimą paprastai skamba taip: teorija paaiškina stebimus reiškinius, dera su jau pripažintu moksliniu žinojimu, yra paprasta ar unifikuoja skirtingų reiškinių paaiškinimus. Bet koks argumentas už teorinės sąvokos, teorinio esinio ar teorijos apskritai teisingumą ar realumą yra GPI atvejis. Taigi, GPI yra fundamentalus teorinio mąstymo komponentas. GPI taikymas paplitęs taip pat ir medicininėje diagnostikoje, kriminaliniuose tyrimuose, teisiniame samprotavime, gedimų diagnostikoje, kasdieniame mąstyme ir bet kuriame kitame kontekste, kai reikia nustatyti kokių nors reiškinių priežastis ar jas pagrįsti.

GPI laikomas svarbiu dabartinės analitinės mokslo filosofijos ir epistemologijos tyrimų objektu. Atsakymas į klausimą, ar GPI veda į tiesą, ypatingai aktualus mokslinio realizmo problemai. Pagal mokslinį realizmą (Kitcher 1993; Leplin 1997; Niiniluoto 1999a; Psillos 1999), geriausias brandžių mokslinių teorijų empirinės sėkmės paaiškinimas yra jų teisingumas. Mokslinės teorijos yra GPI taikymo rezultatas. Taigi, jeigu mokslinis realizmas teisus, atrodytų, kad GPI turi vesti į tiesą. Kita vertus, jeigu GPI neveda į tiesą, tai leidžia stipriai suabejoti mokslinio realizmo teisingumu.

**Tyrinėjimų apžvalga.** GPI dažniausiai įvardinamas kaip C. S. Peirce'o pristatyto argumentacijos metodo – abdukcijos (Peirce 1932-1958) išplėtojimas, tačiau pačią sąvoką „geriausio paaiškinimo išvedimas“ įvedė Harmanas 1965 metais (Harman 1965). Svarbiausi GPI, kaip *vedančio į tiesą, (truth-conducive)* proponentai yra Liptonas (1993; 2001a; 2004) bei Psillosas (2002; 2007; 2009d). Už GPI, kaip vedantį į tiesą, taip pat pasisako Carruthersas (1992), Goldmanas (1990), Josephsonas (2001), Josephsonas ir Josephson (2003), Lycanas (1988), Thagardas (2007b) ir Tuomela (1985). Toks požiūris, pagal kurį aiškinamoji galia (*explanatory power*) tampa lemtingu kriterijumi sprendžiant apie teorijos teisingumą, vadinamas aiškinančiuoju požiūriu (*explanationist account*).

Kiti autoriai GPI interpretuoja nuosaikiau. Day'us ir Kincaidas (1994) ar Ben-Menahemas (1990) teigia, kad GPI veda į tiesą tik kaip bendra samprotavimo strategija, bet ne kaip konkreti samprotavimo forma. Niiniluoto (1999b; 2004) pasisako už tai, kad GPI kartu su empiriniais tyrimais yra geriausias mokslui prieinamas būdas ieškoti informatyvių tiesų. Kuipersas (2004) teigia, kad GPI turėtų daryti lyginamąsias (geriausias yra artimiausias tiesai), bet ne absoliučias (geriausias yra teisingas) išvadas.

Žymiausias GPI, kaip vedančio į tiesą, kritikas – Van Fraassenas, jo darbuose (1980; 1989) pateikti svariausi argumentai prieš GPI ir jo taikymą. Van Fraasseno teigimu, GPI ne tik neveda link tiesos, bet ir verčia būti nenuosekliais. Ši kritika sukėlė plačias diskusijas: Psillosas (1996; 1999) gina GPI nuo van Fraasseno argumentų; Ladymanas su kolegomis (1997) kritikuoja Psilloso gynybą; Douvenas (1999; 2002) gina GPI nuo van Frasseno kritikos, tačiau neteigia, kad GPI veda į tiesą; Barnes'o (1995), Bartelbortho (2005), Salmono (2001a) ir Schurzo (2008) tyrimai atstovauja skeptišką poziciją GPI, kaip vedančio į tiesą, atžvilgiu.

Tyrinėtojus taip pat domina GPI ir aiškinamosios galios apibrėžimai. Thagardas (1978; 1993) ar Ylikoski ir Kuorikoski (2010) aptaria paaiškinimų vertinimo kriterijus. Glassas (2007; 2010), McGrew (2003) bei Schupbachas ir Sprengeris (2011) siūlo tikimybinis aiškinamosios galios apibrėžimus.

Bajesizmas šiuo metu dominuoja analitinėje mokslo filosofijoje, epistemologijoje ir sprendimų teorijoje. Dėl šios priežasties suderinamumas tarp GPI ir bajesizmo taip pat yra aktuali tyrimų tema. Liptonas (2001a), Niiniluoto (1999b), Okasha (2000), Salmonas (2001b) ir Weisbergas (2009) aptaria galimus GPI ir bajesizmo suderinimo būdus. Tuo tarpu Psillosas (2004; 2009b) abejoja, ar šis suderinimas reikalingas ir prasmingas.

GPI artima abdukcijos tema svarbiausius dabartinius tyrimus atliko Aliseda (2006), Gabbay'us ir Woodsas (2005), Paavola (2006b) ir Schurzas (2008). Kai kurie tyrinėtojai neskiria GPI nuo abdukcijos ir tai, kas šiame tyrime įvardinama GPI vardu, vadina abdukcija (Barnes 1995; Carruthers 2006; Douven 2011; Fodor 2000; Josephson and Josephson 2003; Niiniluoto 1999b; 2004; Plečkaitis 2006; Psillos 2000; 2002; 2004). Kiti tyrinėtojai aiškina, kad GPI ir abdukcija skiriasi (Campos 2009; Minnameier 2004; Hintikka 1998; McKaughan 2008). Dar kiti tyrinėtojai mano, kad GPI ir abdukcija tam

tikru mastu persidengia (Aliseda 2006; Gabbay and Woods 2005; Kuipers 2004; Magnani 2001; Paavola 2006a; Schurz 2008; Thagard 2007b). Šiame tyrime atstovaujamas požiūris, kad GPI ir abdukcija, nors ir itin tarpiai susiję, yra konceptualiai skirtingi dalykai.

**Tyrimo problema.** Disertacija analizuoja GPI, kaip vedančio į tiesą, problemą – ar GPI pobūdis laiduoja, kad GPI veda į tiesą? Ši problema turi du aspektus: pirma, kuris paaiškinimas iš potencialių paaiškinimų aibės geriausias; antra, jeigu nustatėme geriausią paaiškinimą, ar galime būti tikri, kad šis paaiškinimas teisingas. Kitaip tariant, pirma, tyrimas eksplikuoja, kas yra GPI: kokie kriterijai apsprendžia, kas yra potencialių paaiškinimų aibė; kokie kriterijai lemia, kuris paaiškinimas iš potencialių paaiškinimų aibės yra geriausias; ar visada galima nustatyti geriausią paaiškinimą? Antra, tyrimas siekia įvertinti GPI pretenzijų į tiesą apginamumą – ar teisinga hipotezė, jeigu ji laikoma geriausiu tam tikro reiškinio paaiškinimu?

**Ginami teiginiai.** Disertacijoje ginami šie teiginiai:

1. GPI yra materialaus samprotavimo forma, kuri priskiria teisingumą toms hipotezėms, kurios tarp savo varžovių pasižymi aukščiausiu aiškinamųjų veiksnių laipsniu: mažiausiai prieštarauja pripažintam žinojimui, unifikuoja svarbiausius reiškinius, yra giliausias paaiškinimas ir yra paprasčiausias paaiškinimas. Šis aiškinamasis požiūris į GPI fundamentalesnis nei tikimybinis požiūris į GPI, o koherencija neturėtų būti laikoma vienu iš aiškinamųjų veiksnių.

2. Šiuo metu literatūroje galima išskirti keturis GPI, kaip vedančio į tiesą, pagrindimo būdus: patikimumo-koherentinis, evoliucinis, tikimybinis bei empirinis-istorinis.

3. Nė vienas iš išskirtų GPI pagrindimo būdų nėra pakankamas pagrįsti, kad GPI veda į tiesą. Juos pakerta prastos atrankos argumentas (*bad lot argument, argument from underconsideration*), pesimistinės indukcijos argumentas (*pessimistic induction*), argumentas „atsarga gėdos nedarą“ (*better safe than sorry*) bei tarpusavyje prieštaraujančių aiškinamųjų veiksnių rikiuočių galimybė. Dėl šių veiksnių būti geriausiu paaiškinimu nereiškia būti tiesa.

4. GPI yra pagrįstas pragmatiškai, t.y. GPI plačiai taikomas, nėra geresnio žinojimą plečiančio samprotavimo būdo, veikla, taikant GPI, yra sėkminga.

**Tyrimo metodologija.** Tyrimas remiasi korespondentine tiesos teorija, kadangi tik šios tiesos teorijos kontekste prasminga kalbėti apie tai, ar GPI veda į tiesą. Tyrimas priima natūralizuotos epistemologijos perspektyvą, pagal kurią filosofiniai klausimai analizuojami remiantis ir gamtos ar kitų mokslų atradimais. Nors GPI, kaip vedančio į tiesą, kritika silpnina mokslinį realizmą, mokslinis realizmas gali būti teisingas netgi, jeigu GPI neveda į tiesą. Todėl tyrimas neanalizuos mokslinio realizmo ir klausimo, ar šis požiūris teisingas. Tyrimas pabrėžia skirtumą tarp GPI ir abdukcijos. Nors kartais argumentai ar citatos, mininčios abdukciją, panaudojami kalbant apie GPI, taip atsitinka tik tuo atveju, jeigu argumento ar citatos autorius neskiria GPI nuo abdukcijos ir abi šias sąvokas vartoja kaip sinonimiškas. Sąvoka „veda į tiesą“ vartojama samprotavimo būdo patikimumui žymėti, t.y. GPI veda į tiesą, jei ir tik jei GPI išvados būna teisingos ar artimos tiesai.

**Tyrimo naujumas.** Disertacija rekonstruoja GPI bei ištiria gautų rezultatų vidinį neprieštaringumą, konceptualią koherenciją bei santykį su empiriniais ir istoriniais duomenimis. Tyrimo naujumas skleidžiasi dviem aspektais. Pirmą, tyrimas pateikia originalią GPI sampratą bei aiškina, kaip išvengti galimų prieštaravimų dabartinėse GPI sampratose. Pasisakoma už aiškinamąjį požiūrį į GPI (Lipton 2004; Psillos 2002) ir kritikuojamas tikimybinis požiūris į GPI (Glass 2007; Schupbach and Sprenger 2011). Antra, tyrimas išskiria keturis būdus, kuriais filosofai siekia pagrįsti, kad GPI veda į tiesą, bei teigia, kad šie būdai tam nėra pakankami. Ginamas bei išvystomas van Fraasseno (1989) prastos atrankos argumentas bei Sticho (1990) argumentas apie „atsarga gėdos nedarą“ įsitikinimus. GPI problematikai pritaikomas Laudano (1981) pesimistinės indukcijos argumentas bei pristatomas argumentas apie galimą aiškinamųjų veiksmų nebendramatiškumą.

**Tyrimo struktūra.** Disertaciją sudaro du pagrindiniai skyriai, kurie atitinka anksčiau išskirtas dvi pagrindines tyrimo problemas. Pirmajame skyriuje rekonstruojamas GPI bei pasisakoma už aiškinamąjį požiūrį į GPI. GPI dažnai aprašomas kaip susidedantis iš dviejų žingsnių: pirmajame žingsnyje surenkama



potencialių paaiškinimų aibė, o antrajame iš šios aibės išrenkamas geriausias paaiškinimas. Dėl šios priežasties, skirsnis 1.2 pateikia abdukcijos kaip pirmojo GPI žingsnio sampratą, o likę du pirmojo skyriaus skirsniai pristato dvi alternatyvias antrojo GPI žingsnio sampratas: aiškinamąjį požiūrį (1.3) ir jo gynybą (1.4.2), pagal kurią, hipotezės aiškinamoji galia lygi aiškinamųjų veiksmų, kuriais pasižymi hipotezė, laipsniui; tikimybinį požiūrį ir jo kritiką (1.4), pagal kurią, hipotezės aiškinamoji galia yra hipotezės pradinės tikimybės ir duomenų tikėtumo tam tikro derinio rezultatas. Antrajame skyriuje įvertinama, ar GPI veda į tiesą. Išvardinami keturi būdai, kuriais grindžiamas GPI, kaip vedantis į tiesą, teigiama, kad šie būdai to nepagrindžia, bei skyrius užbaigiamas skirsniu (2.6), apibendrinančiu argumentus, kurie neigia, kad GPI veda į tiesą. Skirsniai apie psichologinį GPI adekvatumą (2.3.1) bei ontologinius įsipareigojimus (2.5.1) yra relevantiški kai kuriems GPI pagrindimo būdams ir pateikti atitinkamose antrojo skyriaus dalyse.

### **Kas yra GPI?**

**Abduktvyus GPI mechanizmas (1.2).** GPI labai dažnai siejamas su abdukcija. Abdukcijos sąvoka, savo ruožtu, dažniausiai siejama su C. S. Peirce'o vardu. Peirce'as pateikė klasikinę abdukcijos apibrėžimą:

Stebimas neįprastas faktas.

Tačiau, jei  $A$  būtų teisingas, tai  $C$  būtų savaime suprantamas.

Todėl turime pagrindą manyti, kad  $A$  yra teisingas.

GPI skiriasi nuo abdukcijos dviem esminiais ypatumais. Pirma, GPI prielaidas sudaro abi abduktvyaus samprotavimo prielaidos kartu su viena papildoma prielaida, kad nėra geresnio paaiškinimo už analizuojamąjį. Antra, šios papildomos prielaidos įtraukimas į samprotavimą lemia, kad GPI išvada vertinama ne kaip galimai teisinga, bet kaip teisinga. Abdukcija teikia tik galimai teisingą išvadą, o GPI teikia tariamai teisingą ar bent labai artimą tiesai išvadą:

Stebimas neįprastas faktas.

Tačiau, jei  $A$  būtų teisingas, tai  $C$  būtų savaime suprantamas.

Jokia kita hipotezė negali paaiškinti  $C$  taip gerai kaip  $A$ .

Todėl  $A$  yra tiesa.

Poreikis abdukcijai išskyla, kai susiduriama su abdukcijos paleidikliu (*abductive trigger*): konkrečiu stebimu reiškiniu, kuris neseka iš turimo žinojimo arba kuris prieštarauja turimam žinojimui. Abdukcijos tikslas yra pasiūlyti tokį turimo žinojimo papildymą arba modifikaciją, kuri leistų abdukcijos paleidikliui logiškai plaukti iš turimo žinojimo. Jeigu taip įvyksta, tai reiškia, kad abdukcijos paleidiklis buvo paaiškintas. GPI teorijos nereikalauja kokios nors konkrečios aiškinimo sampratos. GPI proponentai teigia, kad GPI turi būti suderinamas su bet kuria aiškinimo samprata. Vis dėlto, kai kurie GPI proponentai palankesni priežastinei-mechaninei aiškinimo sampratai. Santykį tarp GPI ir abdukcijos apibendrina skirtis tarp potencialaus ir aktualaus paaiškinimo. Abdukcijos išvada yra potencialus paaiškinimas – galimas, bet neaišku, ar teisingas paaiškinimas. GPI išvada, manoma, yra aktualus paaiškinimas – tas vienas iš potencialių paaiškinimų (viena iš abdukcijos išvadų), kuris yra teisingas. Todėl abdukcija laikytina pirmuoju GPI žingsniu.

**Aiškinamoji galia ir aiškinamieji veiksniai (1.3).** Kai surenkama aibė potencialių paaiškinimų, išskyla poreikis nustatyti, kuris iš potencialių paaiškinimų yra geriausias paaiškinimas, t.y. tariamai teisingas paaiškinimas. Aiškinamoji galia matuojama pagal paaiškinimų demonstruojamą aiškinamųjų veiksnių laipsnį. Taigi, aiškinamųjų veiksnių laipsnio svėrimas yra antrasis GPI žingsnis. Tyrimai, analizuojantys GPI, abdukciją, aiškinimą ar teorijų pasirinkimą, mini daug įvairių aiškinamųjų veiksnių. Šį daugį galima suskirstyti į grupes. Aiškinamieji veiksniai būna susiję su aiškinimo koherencija, apimtimi (unifikacija), gyliu, paprastumu ar empiriniu adekvatumu.

Koherencija. Koherencija yra dažniausiai minimas aiškinamasis veiksnys. Jis žymi koherenciją tarp paaiškinančios hipotezės ir turimo žinojimo, t.y. tokio žinojimo, kuris laikomas teisingu. Hipotezė yra tuo geresnis paaiškinimas, kuo ji labiau koherentiška. Kartais koherencija vertinama tik kaip eilinis aiškinamasis veiksnys, tačiau kai kurie tyrinėtojai mano, kad koherencija yra svarbiausias aiškinamasis veiksnys. Koherencija laikoma svarbiausiu aiškinamuoju veiksmu, nes ji ne tik būtina, bet kartais ir pakankama geriausio paaiškinimo nustatymui. Kita vertus, GPI teorijos retai apibrėžia, kas turima omeny vartojant koherencijos sąvoką. Disertacijoje parodoma, jog jeigu atsižvelgsime į tai, kaip koherenciją eksplikuoja filosofiniai tyrimai, GPI teorijoms kyla grėsmė patekti į

ydingą ratą. Įvairiuose koherencijos tyrimuose pastebimas dėsningumas – koherencija apibrėžiama kaip susidedanti iš dviejų komponentų: loginio neprieštaravimo bei išvedimo-paaiškinimo ryšių. Loginis neprieštaravimas būtinas, bet nepakankamas koherencijos dėmuo. Likęs dėmuo yra išvedimo-paaiškinimo ryšiai, t.y. teiginiai ar teiginių sistemos yra koherentiškos, jeigu jos ne tik neprieštarauja viena kitai, bet ir paaiškina viena kitą. Dėl pastarojo koherencijos apibrėžimo ir kyla grėsmė GPI teorijoms pakliūti į ydingą ratą – geriausias paaiškinimas tas, kuris yra koherentiškiausias su turimu žinojimu, o koherentiškiausias su turimu žinojimu paaiškinimas tas, kurį su turimu žinojimu sieja stipriausi paaiškinimo ryšiai.

Aiškinimo plotis (unifikacija). Hipotezė yra tuo geresnis paaiškinimas, kuo ji labiau suvienija, unifikuoja įvairių reiškinių paaiškinimą. Tam, kad hipotezė būtų laikoma unifikuojanti, jai keliami du reikalavimai. Pirma, hipotezė unifikuoja, jeigu ji paaiškina bent du skirtingus reiškinius. Antra, hipotezė unifikuoja, jeigu jos teorinių postulatų skaičius mažesnis už hipotezės sekmenų skaičių. Myrvoldas (2003), McGrew (2003) ar Hitchcockas (2007) siūlo unifikaciją apibrėžti kaip priklausančią nuo pozityvaus relevantiškumo (*positive relevance*) laipsnio, t.y. hipotezė tuo labiau unifikuoja, kuo labiau pozityviai relevantiškus ji padaro savo aiškinamus reiškinius (MMH matas). Tačiau toks unifikacijos matas gali nustatyti unifikaciją, net jeigu nėra aiškinamojo ryšio. Taip pat, šis matas gali nenustatyti unifikacijos, net jeigu hipotezė paaiškina ir unifikuoja kelis reiškinius. Reichenbacho bendros priežasties principą (CCP matas) taip pat galima laikyti unifikacijos modeliu. Tačiau MMH matas ir CCP matas prieštarauja vienas kitam. Pagal MMH matą, unifikacija įvyksta, jeigu tikimybiškai nepriklausomi duomenys tampa pozityviai relevantiški. Tuo tarpu pagal CCP matą, unifikuojanti hipotezė yra ta, kuri pozityviai relevantiškus duomenis padaro tikimybiškai nepriklausomais. Be to, abu matai – tiek MMH, tiek CCP – turi dar vieną bendrą labai svarbų trūkumą. Jei iš vienos hipotezės deduktyviai plauks tik du sekmenys, o iš kitos deduktyviai plauks nors ir dešimt skirtingų sekmenų, tiek MMH, tiek CCP abi hipotezes neįvertins kaip unifikuojančias. Taip atsitinka todėl, kad tiek MMH, tiek CCP hipotezes, kurios paaiškina duomenis deduktyviai, nevertina kaip unifikuojančias. Todėl tam tikros hipotezės unifikacijos verte turėtų būti laikoma skaičius skirtingų faktų, kuriuos

paaishkina (deduktyviai paaishkina, padaro pozityviai relevantiškais ar, sąlygojant šia hipoteze, padaro tikimybiškai nepriklausomais) ši hipotezė.

Aiškinimo gylis. Paaishkinimas yra tuo geresnis, kuo jis gilesnis. Paaishkinimo gylis  $m$  yra aibės  $\{M_i \mid 1 \leq i \leq m\}$  kardinalumas, kur  $M_i$  yra paaishkinimo  $M_{i-1}$  paaishkinimas.

Paprastumas. Paaishkinimas yra tuo geresnis, kuo jis paprastesnis. Išskirtinas semantinis ir sintaksinis paprastumas. Semantinis paprastumas yra ontologinis paprastumas ir priklauso nuo postuluojamų esinių – priešasčių, dėsnių, objektų, principų, savybių ir pan. – skaičiaus. Sintaksinis paprastumas yra paaishkinimo kalbos struktūrinis paprastumas ir priklauso nuo paaishkinimo kalbos elementų – simbolių, žodyno, reguliuojamų parametrų ir pan. – skaičiaus. Taip pat išskirtinas paprastumas kaip postuluojamų esinių skaičius ar sintaksinis paaishkinimo ilgis ir paprastumas kaip taupumas. Taupumas apibrėžtinai kaip santykis  $n/k$ , kur  $n$  yra reiškinijų, kuriuos hipotezė paaishkina, skaičius, o  $k$  yra hipotezės teigiamų esinių skaičius ar sintaksinis paaishkinimo ilgis.

Empirinis adekvatumas. Empirinis adekvatumas yra visų mokslinių teorijų siekis, dėl to atrodytų, kad empirinis adekvatumas turėtų būti labiausiai siekiamas paaishkinimų ypatumas. Tačiau GPI teorijų perspektyvoje, jeigu GPI veda į tiesą, tai empirinis adekvatumas turėtų būti ne tiek aiškinamasis veiksnys, kiek visų likusių aiškinamųjų veiksmių pasekmė.

Ryšiai tarp aiškinamųjų veiksmių. Literatūroje minimų aiškinamųjų veiksmių daugis disertacijoje sumažinamas juos suskirstant į koherencijos, unifikacijos, gylio, paprastumo ir empirinio adekvatumo veiksmius. Šis daugis disertacijoje dar labiau sumažinamas išskleidžiant ryšius tarp išskirtų aiškinamųjų veiksmių. Pats ryškiausias ryšys yra tarp unifikacijos ir paprastumo (taupumo): jie abudu panašiai apibrėžiami ir naudojami. Be to, ryšių tarp aiškinamųjų veiksmių išskleidimas leidžia išvengti ydingo rato, kurį sukelia aiškinamosios galios apibrėžimas remiantis koherencija. Disertacijoje teigiama, kad koherencija tapatintina su aiškinamąja galia ir abu, savo ruožtu, apibrėžtini per likusius aiškinamuosius veiksmius: unifikaciją, gylį bei paprastumą, t.y. koherenciją tarp hipotezės ir turimo žinojimo sudaro neprieštarinumas ir tarpusavio paaishkinimo ryšiai. Tarpusavio paaishkinimo ryšiai, savo ruožtu, tuo stipresni, pirma, kuo daugiau

skirtingų reiškinių ir kuo giliau hipotezė paaiškina, antra, kuo detalesnę informaciją (duomenis ar kitus paaiškinimus, jei hipotezė tokius numato), paremiančią hipotezės teisingumą, teikia turimas žinojimas. Paprastumo rūšis, kuri savo struktūra panaši į unifikaciją (taupumas), įeina į šį apibrėžimą, bet paprastumas kaip sintaksinis ilgis ar teigiamų esinių skaičius į šį apibrėžimą neįeina.

**Tikimybinis požiūris į aiškinamąją galią (1.4).** Pateiktas aiškinamasis požiūris nėra vienintelis literatūroje randamas aiškinamosios galios apibrėžimo būdas. Esama ir tikimybinų (bajesiškų) aiškinamosios galios apibrėžimų. Dažnai aiškinamoji galia siejama su duomenų tikėtinumu (ML), t.y. su abdukcijos paleidiklio tikimybe, jeigu paaiškinimas būtų teisingas. Tikimybinį aiškinamosios galios matą, besiremiantį duomenų tikėtinumu, yra pateikę McGrew (2003) bei Schupbachas ir Sprengeris (2011). Aiškinamoji galia taip pat siejama su duomenų tikėtinumu ir pradinės tikimybės deriniu (ML & MPRIOR). Už tokį požiūrį pasisako van Frassenas (1980) ir Okasha (2000). Aiškinamoji galia taip pat siejama su duomenų tikėtinumu ir baigtinės tikimybės deriniu (ML & MPOST). Glassas (2007) pasisako už aiškinamosios galios matą paremtą šiuo deriniu. Disertacijoje teigiama, kad aiškinamajam požiūriui turėtų būti teikiama pirmenybė prieš tikimybinį požiūrį. Pirmia, GPI yra žinojimą plečiantis samprotavimo būdas, o bajesizmas aprašo tik įsitikinimų tikimybės kitimą, bet ne žinojimo plėtimą. Bajesizmas negali paaiškinti kaip pasiūlomos naujos hipotezės. Antra, esama hipotezių, kurios teikia labiau pageidaujamas, bet mažiau tikėtinas paaiškinimus. Trečia, pagal aiškinamąjį požiūrį, GPI leidžia išskirti empiriškai ekvivalenčias hipotezes, bajesizmas to padaryti negali. Ketvirta, jeigu geriausią paaiškinimą sutapatinam su tikimybiniais samprotavimais, tai GPI savoka tampa nereikalinga: arba aiškinamieji veiksniai yra kažkas plačiau nei tik tikimybiniai samprotavimai, arba GPI žymi tą patį dalyką (tikimybinį samprotavimą) ir dėl to yra perteklinė. Kita vertus, bajesizmas gali taikyti Olandiškų lažybų argumentą ir teigti, kad aiškinamasis požiūris veda į nenuoseklumą. Van Frassenas (1989) teigia, kad taikydamas GPI asmuo gali sutikti su tokiomis lažybotomis, kuriose, kokia bebūtų jų baigtis, asmuo patirs nuostolį. Disertacijoje parodoma, kad van Fraasseno Olandiškų lažybų argumente naudojama nekorektiška GPI samprata, o aiškinamojo požiūrio atstovai aprašo būdus, kuriais galima suderinti GPI ir

bajesizmą taip, kad GPI negrėstų Olandiškų lažybų argumentas. Aiškinamojo požiūrio teigimu, GPI galima naudoti tam, kad būtų nustatytas tikimybių pasiskirstymas, t.y. GPI gali padėti bajesizmui įvertinti hipotezių pradinę tikimybę bei duomenų tikėtinumą.

### **Ar GPI veda į tiesą?**

**Pretenzijos į deduktyvų pagrįstumą (2.1).** GPI siekia būti panaši į dedukciją. Samprotavimas yra deduktyviai pagrįstas, jeigu, jo prielaidoms esant teisingoms, jo išvada yra būtinai teisinga. GPI būtų šia prasme pagrįstas, jeigu GPI prielaidoms esant teisingoms, GPI išvada būtų būtinai teisinga, t.y. GPI būtų pagrįstas, jeigu hipotezės gebėjimas paaiškinti abdukcijos paleidiklį bei hipotezės buvimas geriausiu abdukcijos paleidiklio paaiškinimu reikštų, kad hipotezė yra teisinga. Toks teiginys ir yra būtent tai, ką teigia GPI teorijos. Kita vertus, šis teisingumas priklauso ne nuo samprotavimo loginės formos, kaip tai yra dedukcijos atveju, bet nuo samprotavimo turinio: loginių, ontologinių, nomologinių ir pan. ryšių, kurie glūdi turimame žinojime ir analizuojamoje hipotezėje. Tai reiškia, kad GPI yra materialus išvedimas, o ne formalus. Nepaisant tam tikrų GPI ir dedukcijos panašumų, GPI proponentai suvokia, kad GPI nėra ir negali būti deduktyviai pagrįsta – garantuoti savo išvadų teisingumo. Tačiau GPI teorijos vis vien teigia, kad GPI veda į tiesą, kad tai patikimas samprotavimo būdas ar kad tai gera strategija ieškoti tiesų. Filosofinėje literatūroje randami keturi būdai, kuriais bandoma pagrįsti šias kone deduktyvias pretenzijas.

**Patikimumo-koherentinis pagrindimas (2.2).** Patikimumo-koherentinis pagrindimas yra pagrindinė GPI, kaip vedančio į tiesą, pagrindimo strategija. GPI išvada tariamai yra pagrįsta, kadangi ji yra GPI produktas, o GPI yra patikimas įsitikinimų formavimo procesas. Savo ruožtu GPI yra patikimas įsitikinimų formavimo procesas, kadangi jo taikymas užtikrina, kad nauji įsitikinimai derės su jau turimu žinojimu, kad išlaikoma naujų ir jau turimų įsitikinimų koherencija (Lipton 2004, Psillos 2002). Tačiau visada esama loginės galimybės, kad tikrasis paaiškinimas nebuvo įtrauktas į potencialių paaiškinimų aibę. Van Fraassenas (1989) pateikė šiuo mąstymu paremtą prastos atrankos argumentą – niekada negalime žinoti, ar mūsų pasirinkimas nėra geriausias iš blogiausių. Jeigu visgi teigiama, kad GPI veda į tiesą, tai GPI proponentai privalo daryti prielaidą, kad jie turi privilegijuotą priėjimą prie tiesos. Peirce'as iš tiesų kalbėjo apie tokią

privilegiją. Liptono manymu šią privilegiją garantuoja IBE patikimumas, esą prastos atrankos argumentas remiasi dvejomis prielaidomis, kurios tarpusavyje nesuderinamos. Tačiau pastarasis Liptono argumentas taip pat remiasi dvejomis prielaidomis: prielaida, kad aiškinamosios galios vertinimas yra patikimas, bei prielaida, kad turimas žinojimas yra teisingas. Mokslo istorija rodo, kad turimas žinojimas nebūtinai būna teisingas, o aiškinamųjų veiksnių nebendramatiškumo galimybė rodo, kad aiškinamosios galios vertinimas nebūtinai yra patikimas. Taigi, disertacijoje parodoma, kad pačio Liptono kontraargumento prieš prastos atrankos argumentą prielaidos nėra nepriekaištingos. Tačiau esama ir kitos strategijos atsakyti į prastos atrankos argumentą. Siūloma į potencialių paaiškinimų aibę įtraukti visus turimus paaiškinimus neigiančią (kontradiktorinę) hipotezę, t.y. hipotezes ar hipotezę teigiančią, kad visos iki šiol turimos hipotezės yra klaidingos (Lipton 2004; Minnameier 2004; Niiniluoto 1999b). Taip esą konkuruojančių paaiškinimų aibė taps išsami ir negrės, kad liks nepastebėtas teisingas paaiškinimas. Disertacijoje teigiama, jog toks potencialių paaiškinimų aibės papildymas iki išsamos aibės turi rimtą trūkumą – kontradiktorinė hipotezė bus ypatingai prastas paaiškinimas. Dar blogiau, ji nesugebės paaiškinti reikiamo abdukcijos paleidiklio, o jos vienintelis sekmuo bus tas, kad abdukcijos paleidiklis neturėjo įvykti. Kadangi abdukcija yra pirmasis GPI žingsnis ir kadangi abdukcijos išvada gali būti tik tokia hipotezė, kuri paaiškina abdukcijos paleidiklį, tai kontradiktorinė hipotezė negalės būti abdukcijos išvada, negalės būti įtraukta į potencialių paaiškinimų aibę bei, savo ruožtu, negalės būti vertinama GPI proceso metu. Patikimumo-koherentinis pagrindimas daro pernelyg daug induktyvių apibendrinimų ar kvestionuotinių prielaidų, todėl labai mažai tikėtina, kad jis pagrindžia, jog GPI veda į tiesą.

**Psichologinis GPI adekvatumas (2.3.1).** Epistemologijos ir mokslo filosofijos tyrimai retai klausia, ar mąstymas pagal GPI yra būdingas žmonėms, ar net atvirai tuo abejoja. Kita vertus, kai kurie filosofai teigia, kad žmonės vadovaujasi GPI, tačiau nepateikia platesnio tokio teiginio pagrindimo. Disertacijoje parodoma, jog esama psichologijos tyrimų, kuriais remiantis teigtina, kad žmonės vadovaujasi GPI (šis teiginys disertacijoje vadinamas psichologine hipoteze). Žmonės daug dažniau, nei tą leistų logika, mano, kad abdukcija yra pagrįstas samprotavimo būdas. Žmonės daug

labiau tiki tais įsitikinimais, kuriems turėjo sugalvoti paaiškinimą, kodėl šie įsitikinimai galėtų būti teisingi. Gebėjimą paaiškinti žmonės vertina kaip įrodymą, kad įsitikinimas teisingas. Hipotezės, kurios pasižymi didesniu aiškinamųjų veiksmių laipsniu, vertinamos kaip labiau tikėtinos.

**Pragmatizmas ir GPI (2.3.2).** Peirce'ui abdukcija yra pagrįsta dėl to, kad tai vienintelis būdas pristatyti naujas idėjas, kurias galima patikrinti. Šis pagrindimo būdas pragmatistams yra episteminis pagrindimas. Tam tikras įsitikinimas pragmatistui yra epistemiškai pagrįstas, jeigu, pirma, tuo laiko momentu nėra žinomas joks kitas įsitikinimas, kuris būtų labiau pagrįstas, bei, antra, šio įsitikinimo priėmimas turėtų daugiau naudingų pasekmių nei jo neigimas ar ignoravimas. GPI kontekste pirmasis reikalavimas atitinka reikalavimą pasirinkti geriausią iš prieinamų paaiškinimų, o antrasis reikalavimas atitinka reikalavimą, kad GPI taikymas turėtų turėti asmeniui naudingų pasekmių. GPI vertintinas kaip naudingas dėl kelių priežasčių. Pirma, geriausią paaiškinimą paprasčiau suprasti ar juo operuoti. Antra, taikyti GPI ekonomiškai. Trečia, jeigu psichologinė hipotezė yra teisinga, tai evoliucinė psichologija ją aiškintų tuo, kad GPI taikymas didina išgyvenimo ar reprodukcijos galimybes. Tačiau tokio pobūdžio pragmatinis pagrįstumas nereiškia, kad GPI veda į tiesą. Sąvoka „episteminis“ pragmatizmo kontekste neturi nieko bendra su tiesa ar artumu tiesai.

**Evoliucinis pagrindimas (2.3.3).** Kai kurie filosofai mano (Carruthers 1992; Goldman 1990), kad evoliucinės psichologijos požiūriu interpretuojant faktą, jog GPI taikymas paplitęs tarp žmonių, reiškia, jog GPI išvados yra teisingos. Esą teisingų įsitikinimų turėjimas padidina galimybes išgyventi. Organizmai, kurie geba formuoti teisingus įsitikinimus, išgyvena dažniau ir turi didesnę galimybę tokį gebėjimą perduoti savo palikuonims. GPI taikymas yra paplitęs tarp žmonių. Šis taikymas turėjo paplisti evoliucijos eigoje, o paplisti evoliucijos eigoje jis galėjo tik tuo atveju, jei padėdavo prisitaikyti prie aplinkos, t.y. teikė teisingus įsitikinimus apie aplinką ir pan. Nors pagrindinis natūralios atrankos tikslas yra išlikimas, bet teisingi įsitikinimai esą labiau nei klaidingi padeda išgyventi. Šis argumentas gali būti paneigtas, jeigu esama klaidingų įsitikinimų, kurie padeda išgyventi. Disertacijoje parodoma, jog tokių klaidingų, bet išgyventi padedančių įsitikinimų pavyzdys yra „atsarga gėdos nedaro“ įsitikinimai –



geriau sureaguoti į klaidingą įsitikinimą apie pavojų, nei vėliau gailėtis. Tokie klaidingi įsitikinimai tuo dažnesni, kuo mažiau kainuoja reagavimas į klaidingą pavojų bei kuo labiau žalingas yra klaidingas įsitikinimas, kad pavojaus nėra. Taigi, išgyvenimui nebūtinai vien teisingi įsitikinimai ir todėl, net jeigu GPI padeda išgyventi ar prisitaikyti, nereiškia, kad GPI veda į tiesą.

**Tikimybinis pagrindimas (2.4).** Viena vertus, Kleinas ir Warfieldas (1994) parodė, kad labiau koherentiška teiginių aibė negali būti labiau tikėtina. Kita vertus, galima įvesti tam tikras sąlygas (analizuoti ne teiginių aibes, bet įsitikinimų aibes; daryti prielaidą, kad liudijimai apie įsitikinimus yra nepriklausomi ir tik dalinai patikimi) ir tada analizuoti, ar koherencija veda į tiesą. Bovensas ir Hartmannas (2003) bei Olssonas (2005a; 2005b) įrodė, kad ir priėmus šias sąlygas koherencija neveda į tiesą. Tolimesnės pastangos įrodyti, kad koherencija veda į tiesą, apima bandymus modifikuoti sąlygas, kurioms esant koherencija turėtų vesti į tiesą. Vis dėlto, nei Kleino ir Warfieldo, nei Bovenso ir Hartmanno, nei Olsono rezultatai netaikytini GPI kontekste, kadangi įvairios šių filosofų daromos prielaidos nėra suderinamos su GPI teorijomis. Glassas (2010) ištyrė ar koherencija veda į tiesą kitokiu būdu. Jis atliko kompiuterinę simuliaciją, kaip įvairūs aiškinamosios galios matai sugeba atspėti geriausią paaiškinimą, kuris buvo numatytas kaip geriausias prieš simuliaciją. Glasso simuliacija parodė, kad jo siūlomas aiškinamosios galios matas (Glasso ar Olsono (2002) siūlomas kaip koherencijos matas), geriausiai atspėdavo geriausią paaiškinimą. Disertacijoje teigiama, kad šis rezultatas nereiškia, kad GPI veda į tiesą. Tikimybinis aiškinamosios galios matas galima taikyti tik turint tikimybių pasiskirstymą. Jeigu šis tikimybių pasiskirstymas teisingas, tai ir iš anksto numatytas geriausias paaiškinimas bus lengvai atspėjamas. Tačiau, jeigu šis tikimybių pasiskirstymas nėra teisingas ar artimas teisingam, tai tikimybiniai aiškinamosios galios matai daug rečiau atspės teisingą hipotezę. Taigi, tikimybiniai matai yra veiksmingi tik po to, kai visas hipotezių įvertinimas jau būna atliktas, kai remiantis aiškinamaisiais veiksniais jau buvo įvertintas tikimybių pasiskirstymas. Tikimybinė unifikacijos ar paprastumo analizė taip pat neparodo, kad unifikacija ar paprastumas veda į tiesą. Tikimybinis GPI, kaip vedančio į tiesą, pagrindimas reikštų, kad GPI yra deduktyvus samprotavimas, tačiau tai prieštarauja

pačiai GPI teorijai. Todėl, jeigu nenori prieštarauti pačios sau, tikimybinės GPI teorijos turi at mesti tikimybinių pagrindimo galimybę.

**GPI ontologiniai įsipareigojimai (2.5.1).** Tam tikra GPI išvada gali būti teisinga tik tam tikrais ypatumais pasižyminčiame pasaulyje. GPI yra ne formali, bet materialiai samprotavimo forma, jos išvadų teisingumas priklauso nuo to, kokia dalykų padėtis teigiama tam tikro GPI atvejo prielaidose bei turimame žinojime. Kitaip tariant, disertacijoje teigiama, kad GPI veda į tiesą, jei ir tik jei pasaulis pasižymi tam tikra ontologine struktūra. Išorinio pasaulio realumas yra esminė GPI prielaida. Jeigu nebūtų išorinio pasaulio, negalėtų įvykti abdukcijos paleidikliai, nebūtų to, ką reikėtų paaiškinti. Kita GPI ontologinė prielaida ta, kad pasaulis yra vienodas ir reguliarus. Be šios prielaidos tie patys paaiškinimai negalėtų būti taikomi tiems patiems reiškiniams paaiškinti. Trečia GPI būtina prielaida ta, kad turimas žinojimas, kuriuo remiamasi tam tikru GPI atveju, turi būti teisingas, t.y. dalykų padėtis privalo būti tokia, kokia aprašyta tame turimame žinojime. Paskutinė prielaida, kuri reikalinga tam, kad GPI būtų vedantis į tiesą, yra ta, kad aiškinamieji veiksniai turi vesti į tiesą. Kitaip tariant, jeigu paaiškinimai, kurie yra neprieštaringi, labiausiai unifikuojantys, giliausi ir paprasčiausi, turėtų būti teisingi, tai ir pasaulis turi būti neprieštaringas, unifikuotas ir paprastas. Taigi, GPI veda į tiesą, jei ir tik jei esama išorinio pasaulio, kuris yra vienodas ir pasižymi reguliarumais, turimas žinojimas, aprašantis šį išorinį pasaulį, yra teisingas, o aiškinamieji veiksniai šiame pasaulyje veda į tiesą. Šios ontologinės prielaidos privalo būti teisingos tam, kad GPI galėtų būti vedantis į tiesą.

**Empirinis-istorinis pagrindimas (2.5.2).** Empirinė hipotezės sėkmė yra prieinamiausias hipotezės teisingumo požymis. Todėl, jeigu GPI veda į tiesą, geriausias paaiškinimas turi būti labiau empiriškai sėkmingas negu alternatyvūs paaiškinimai. Tačiau mokslo istorija rodo, kad esama savo laiku geriausiais paaiškinimais pripažintų teorijų, kurios, vėliau paaiškėjo, buvo klaidingos. Laudano (1981) pesimistinės indukcijos argumentas pateikia tokių teorijų pavyzdžius. Kiti pavyzdžiai rodo, kad unifikuojančios ar paprastesnės teorijos nebūtinai yra teisingos – aiškinamieji veiksniai nebūtinai veda į tiesą. Taigi, disertacijoje parodoma, kad mokslo istorijoje randame klaidingų GPI išvadų. Jeigu būtų atsakyta, kad klaidingų GPI išvadų atvejais

aiškinamosios galios kriterijai nebuvo pakankamai griežti ir kad nėra paneigtų griežtesnius kriterijus atitinkančių GPI išvadų atveju, tai tik sustiprintų argumentą, kad esama prieštaraujančių viena kitai aiškinamosios galios rikiuočių.

**GPI neveda į tiesą (2.6).** Induktyvus GPI pobūdis neleidžia GPI būti vedančiam į tiesą. GPI neatitinka pagrįsto samprotavimo taisyklių. Antrasis GPI žingsnis, savo forma panašus į disjunktivyvų silogizmą, irgi nepadaro GPI pagrįstu, nes potencialių paaiškinimų aibė negali būti išsami. Pirma, kontradiktorinių hipotezių naudojimas šią aibę padarytų išsamia, bet GPI pobūdis neleidžia jų naudoti. Antra, potencialių paaiškinimų aibė yra potencialiai begalinė. Šie du argumentai bei pesimistinės indukcijos argumentas pritaikytas GPI kontekstui sustiprina prastos atrankos argumentą. GPI teorijas toliau pakerta nesuderinamų aiškinamosios galios rikiuočių galimybė ir „atsarga gėdos nedaro“ argumentas. Todėl, nors GPI ir keliami griežtesni reikalavimai nei abdukcijai, kad GPI teikia labiau tikėtinas hipotezes nei abdukcija – dabartiniai GPI, kaip vedančio į tiesą, pagrindimo būdai yra nesėkmingi. Tačiau platus GPI naudojimas yra sėkmingas ir ši sėkmė yra abdukcijos paleidiklis, kurį reikėtų paaiškinti. Šiam paaiškinimui rasti tikslingas tolimesnis materialaus išvedimo sąvokos tyrimas.

**Išvados.** 1. GPI yra materialaus samprotavimo forma priskirianti tiesą hipotezei, kuri tarp savo varžovių pasižymi didžiausiu aiškinamųjų veiksnių laipsniu: yra labiausiai suderinama su turimu žinojimu, yra labiausiai unifikuojanti, teikia giliausią paaiškinimą bei yra paprasčiausia. Jeigu hipotezė prieštarauja turimam žinojimui ar neištaiso prieštaravimų turimame žinojime, tai ši hipotezė negali būti abdukcijos išvada ir, savo ruožtu, negali būti įtraukta į potencialių paaiškinimų aibę. Be neprieštaravimo, unifikacija kaip kelių skirtingų reiškinių paaiškinimas, paaiškinimo gylis kaip paaiškinimas, kodėl siūlomas paaiškinimas turėtų būti teisingas, ir paprastumas kaip teigiamų aiškinamųjų esinių skaičiaus mažumas ar kalbinės išraiškos, kurioje išreikštas paaiškinimas, trumpumas yra ypatumai, kurie priskiriami geriems ir priėmimo vertiems paaiškinimams. Dėl šios priežasties, jeigu pastaraisiais ypatumais pasižymintys paaiškinimai yra geri, tai paaiškinimai pasižymintys aukščiausiu neprieštaravimo, unifikacijos, gilumo ir paprastumo laipsniu yra geriausi paaiškinimai. Tačiau GPI teorijose atrandami du nenuoseklumai:

(a) Teorijos, kurios koherenciją vertina kaip aiškinamąjį veiksnį, patenka į ydingą ratą, kadangi pačios koherencijos laipsnis priklauso nuo neprieštaravimo plius paaiškinimo ryšių. Dėl šios priežasties, aiškinamoji galia ir koherencija turėtų būti vertinamos kaip žyminčios tą patį dalyką ir eksplikuojamos kaip išvedamos iš likusių aiškinamųjų veiksmių.

(b) Tikimybiniai aiškinamosios galios matai yra pertekliniai. Jeigu tikimybių pasiskirstymas žinomas, tai baigtinės tikimybės paskaičiavimo formulė yra teisingas būdas nustatyti labiausiai tikėtiną hipotezę ir GPI taikymas nereikalingas. Kita vertus, jeigu tikimybių pasiskirstymas nėra žinomas, tai nėra galimybės taikyti tikimybių aiškinamosios galios matų ir belieka remtis aiškinamuoju požiūriu. Abiem atvejais nėra jokio poreikio tikimybiniam aiškinamosios galios matams. Aiškinamasis požiūris yra pakankamas geriausio paaiškinimo nustatymui.

2. Dabartinėse GPI teorijose galima išskirti keturis pagrindinius būdus, kaip siekiama pagrįsti, kad GPI veda į tiesą:

(a) Patikimumo-koherentinis pagrindimas teigia, kad GPI veda į tiesą, nes GPI yra patikimas, t.y. linkęs teikti teisingus, o ne klaidingus įsitikinimus, o patikimas GPI yra todėl, kad GPI stiprina bendrą viso žinojimo koherenciją.

(b) Evoliucinis pagrindimas teigia, kad GPI veda į tiesą, nes mąstymas remiantis GPI yra išgyvenimo galimybes didinantis žmogaus kognityvinis gebėjimas, o geriausias paaiškinimas, kodėl GPI didina galimybes išgyventi tas, kad GPI teikia teisingus įsitikinimus.

(c) Tikimybinis pagrindimas teigia, kad GPI gali būti išreikštas tikimybiškai ir kad GPI veda į tiesą, nes hipotezės, pasižyminčios aiškinamųjų veiksmių aukščiausiu laipsniu, pasižymi ir aukščiausia baigtine tikimybe.

(d) Empirinis-istorinis pagrindimas teigia, kad GPI veda į tiesą, nes hipotezės, pasižyminčios aiškinamųjų veiksmių aukščiausiu laipsniu, yra empiriškai patvirtinamos, o jų buvimas teisingomis ar bent artimomis tiesai yra geriausias paaiškinimas, kodėl jos yra empiriškai patvirtinamos.

3. Nei vienas iš išskirtų keturių pagrindimo būdų nepagrindžia, kad GPI veda į tiesą:

(a) „Atsarga gėdos nedaro“ argumentas paneigia evoliucinį pagrindimo būdą – esama klaidingų, bet išgyvenimo galimybes didinančių įsitikinimų.

(b) Pesimistinė indukcija paneigia empirinį-istorinį pagrindimo būdą – esama geriausio paaiškinimo atvejų, kurie buvo laikomi teisingais, bet vėliau paaiškėjo esantys klaidingais.

(c) Aiškinamųjų veiksnių nebendramatiškumo galimybė pakerta patikimumo-koherentinį pagrindimą – niekada negalima būti tikram, kad nėra jokios alternatyvios ir adekvatesnės aiškinamosios galios rikiuotės.

(d) Prastos atrankos argumentas silpnina patikimumo-koherentinį pagrindimą – potencialių paaiškinimų aibė niekada negali būti išsami ir aibės papildymas potencialiems paaiškinimams kontradiktorinėmis hipotezėmis negali to ištaisyti. Pesimistinė indukcija stiprina prastos atrankos argumentą – kiekvienas paneigtos teorijos atvejis, kurį pateikia pesimistinės indukcijos argumentas, yra prastos atrankos argumento iliustracija, t.y. geresnė teorija, kuri pakeitė paneigtą teoriją, nebuvo įvertinta tuo metu, kai buvo priimama paneigtoji teorija. Aiškinamųjų veiksnių nebendramatiškumo galimybė stiprina prastos atrankos argumentą – net jeigu pasirinkome tam tikrą hipotezę kaip geriausią paaiškinimą, kita hipotezė, kuri yra geresnis paaiškinimas pagal kokius nors kitus aiškinamuosius veiksnus, gali būti tikrasis aktualus paaiškinimas.

(e) Materialus GPI pobūdis prieštarauja tikimybinių būdų taikymui pagrįsti, kad GPI veda į tiesą. Tikimybinis pagrindimas rodytų, kad GPI yra deduktyviai, t.y. formaliai, pagrįstas. Tačiau GPI yra nededuktyvus samprotavimo būdas, jo tariamas pagrįstumas ne formalus, bet materialus, priklausomas nuo substancialių aplinkybių.

(f) Tam, kad galima būtų įvertinti, ar esama kokio nors kito būdo pagrįsti, kad GPI veda į tiesą, reikalingas tolimesnis materialaus samprotavimo sąvokos tyrimas.

4. Nors negalima teigti, kad GPI yra pagrįstas, GPI yra plačiai taikoma tam, kad pateiktų hipotezes, kurias vėliau galima empiriškai tikrinti. Po sėkmingų empirinių patikrinimų GPI išvadas galima priimti, bet ne anksčiau. Šiam laiko momentui GPI, jei ir yra pagrįstas, tai tik pragmatiškai: nėra geresnio žinojimą plečiančio samprotavimo būdo, o veikla, taikant GPI, yra sėkminga. Tačiau tokio pobūdžio pragmatinis pagrįstumas nereiškia, kad GPI veda į tiesą.

### **Mokslinių publikacijų disertacijos tema sąrašas:**

Psychological Adequacy and Ontological Commitments of Inference to the Best Explanation, *Problemos* 79, 41-54 psl., 2011.

Geriausio paaiškinimo išvedimas. Tarp dedukcijos, indukcijos ir abdukcijos, *Problemos* 76, 150-161 psl., 2009.

### **Pranešimai konferencijose:**

What Does the Mind Tell About the Inference to the Best Explanation?, Ketvirtoji Olandijos-Flandrijos analitinės filosofijos asociacijos konferencija (VAF IV), Leuveno katalikiškasis universitetas, 2010 m. vasario mėn.

How to Measure Epistemic Justification of Inference to the Best Explanation?, Centrinės Europos universiteto vasaros universitetas, kursas "Probabilistic Causality", 2008 m. liepos mėn.

### **Išsilavinimas:**

Specializacijos studijos Leuveno katalikiškojo universiteto Filosofijos institute, Belgijoje (2009 m. spalio mėn. - 2010 m. liepos mėn.)

Filosofijos doktorantūros studijos Vilniaus universitete, Filosofijos istorijos ir logikos katedroje (nuo 2007 m. spalio mėn.).

Filosofijos magistro kvalifikacinis laipsnis (2007 m.). Magistro darbas „Mokslinis realizmas šiuolaikinėje filosofijoje“.

Filosofijos bakalauro kvalifikacinis laipsnis (2005 m.). Bakalauro darbas „Semantinis komputacinių proto teorijų aspektas“.

Studijos pagal Erasmus mainų programą Skövde universitete, Švedijoje (2003 m. rugsėjo - gruodžio mėn.).

The dissertation analyses inference to the best explanation and claims about its truth-conduciveness.

**Relevance of the thesis.** The thesis explicates and analyses the inference to the best explanation (IBE): a meta-theoretical principle of inference of scientific propositions. Scientists apply IBE both in the context of discovery and in the context of justification of scientific theories. More particularly, scientists seek such theories that would explain observable phenomena, cohere with already accepted scientific knowledge, be simple, or unify explanations of different kinds of phenomena, and scientists argue for the acceptance of their theories on the grounds that they do indeed explain observable phenomena, cohere with already accepted scientific knowledge, are simple, or unify explanations of different kinds of phenomena. Any argument for the truth or reality of a theoretical term, concept, entity or theory in general is an instance of IBE. IBE is a fundamental component of theoretical reasoning in general and of scientific practice in particular.

IBE is not bound to the context of science. It is also prevalent in medical diagnosis, criminal investigation, judicial argumentation, technical troubleshooting, common sense reasoning and other contexts where one wants to find out the causes of some phenomena. For example, IBE is so prevalent in medical diagnosis or criminal investigation that it even receives a substantial amount of attention in fictional representations of these practices, e.g., the differential diagnoses in the TV-series “House” or the deduction method used by Sherlock Holmes in Arthur Conan Doyle’s stories.

Because of its role in the scientific and commonsense reasoning contexts, IBE is also an important topic for philosophical research in present analytical philosophy of science and epistemology. The truth-conduciveness of IBE is of ultimate importance for the problem of scientific realism. According to scientific realism (e.g., Kitcher 1993; Leplin 1997; Niiniluoto 1999a; Psillos 1999), the best explanation for the empirical success of mature scientific theories is their truth. Scientific theories are the products of IBE. Therefore, if scientific realism is true, then IBE seems to be truth-conducive. On

the other hand, IBE deprived of truth- conduciveness makes scientific realism less viable.

IBE is often considered as the development of C. S. Peirce's idea of abduction (Peirce 1932-1958). Nevertheless, the term inference to the best explanation was introduced only in 1965 by Harman (1965). The most vigorous defense of the truth-conduciveness of IBE is presented by Lipton (1993; 2001a; 2004) and Psillos (2002; 2007; 2009d). Different arguments for the truth-conduciveness of IBE are also put forward, for example, by Carruthers (1992), Goldman (1990), Josephson and Josephson (2003), Glass (2010), Niiniluoto (1999b) and Thagard (2007b). Van Fraassen (1980; 1989), on the other hand, presented the most important arguments against IBE.

**The aim of the thesis.** The thesis is going to analyze the problem of truth-conduciveness of IBE: does the nature of IBE warrant the truth-conduciveness of IBE? The problem has two aspects: given a pool of potential explanations, which one of the potential explanations is the best, and, given that we identified the best explanation, how confident can we be that it is true. More particularly, firstly, the thesis explicates the concept of IBE, i.e., it analyzes what the concept of explanatory power stands for: how to obtain a pool of potential explanation; given a pool of potential explanations, which one of these is the best; and is it always possible to discriminate the best explanation? Secondly, the thesis evaluates the tenability of the truth aspirations of IBE, i.e., it analyzes to what extent a hypothesis can be claimed to be true, given that it is the best explanation for some phenomenon of interest?

**Claims of the thesis.** The thesis claims that even though IBE could facilitate the determination of probability distributions and is a wide psychological practice, due to the bad lot argument, possible incommensurability of explanatory virtues, pessimistic induction and better-safe-than-sorry argument, all the four ways of justifying IBE in terms of truth-conduciveness cannot be taken for granted which leaves IBE only accidentally valid. More particularly, the thesis argues for the following claims:

1. IBE is a form of material inference that ascribes truth to the hypothesis that has the highest degree of explanatory virtues among its competitors: it is the most consistent with approved background knowledge, unifies the most the relevant



phenomena, is the deepest explanation, and is the simplest explanation. This explanationist account is more fundamental than probabilistic accounts of IBE, but coherence should not be treated as one of the explanatory virtues.

2. Currently, there are four basic ways of justifying the truth-conduciveness of IBE that can be discerned in literature: reliabilist-coherentist, evolutionary, probabilistic and empirical-historical.

3. None of the discerned ways of epistemic justification of IBE grants the truth-conduciveness of IBE. They are undermined by the bad lot argument, the argument of pessimistic induction, the better-safe-than-sorry argument, and the possibility of contradicting orders of explanatory power. Therefore, being the best explanation does not grant truth.

4. IBE is warranted pragmatically, i.e. IBE is a widespread psychological practice, there is no better method of ampliative inference and its use helps to successfully cope with the world.

**Methodology of the thesis.** The ongoing discussion on IBE is meaningful only if the correspondence theory of truth is held to be correct. Thus the thesis is going to assume the correspondence theory of truth. The thesis also assumes the naturalized epistemology perspective, in which findings from the sciences that study human reasoning are brought to bear on questions in philosophy. Even though, deprived of truth-conduciveness IBE makes scientific realism less viable, this would not refute scientific realism: scientific realism can be true even if IBE is not truth-conducive. Therefore, the thesis would not argue for nor against the truth of scientific realism. A clear distinction is drawn between IBE and abduction. Nevertheless, claims about abduction are used to illustrate points about IBE: this happens only if the author of the claim does not discriminate between IBE and abduction and means the same thing by the both concepts.

One difficulty in discussing inference to the best explanation lies in the ambiguity of the term denoting it. IBE can denote a particular kind of derivation. IBE can denote the conclusion of a derivation. Finally, IBE can also denote a theory or a set of claims about the properties of inference to the best explanation. In order to avoid this ambiguity

the following convention is going to be used in this thesis. “IBE” will stand for IBE as a derivation or process of inference. “A conclusion of IBE” will stand for the conclusion of IBE as a process of inference. “A theory of IBE” will stand for a theory or a set of claims about the properties of IBE. The term “truth-conduciveness” will stand for the reliability of the method of inference: hence, IBE is truth-conducive if and only if IBE tends to generate true or approximately true conclusions. The term “truthlikeness” will stand for closeness to truth or the degree of verisimilitude: hence, if IBE is truth-conducive, it will tend to generate truthlike conclusions.

**Novelty of the thesis.** The thesis reconstructs the concept of IBE and examines the results of the reconstruction for internal consistency, conceptual coherence and fit with empirical and historical data. The novelty of this thesis lies in two main aspects. Firstly, the thesis presents an original account of IBE and argues how to evade possible inconsistencies that are present in theories of IBE. It adopts and argues for the explanationist (e.g., Lipton 2004; Psillos 2002) rather than probabilistic (e.g., Glass 2007; Schupbach and Sprenger 2011) approach to IBE. Secondly, the thesis discerns four ways in which philosophers argue for the truth-conduciveness of IBE and shows that these ways are insufficient to grant the truth-conduciveness of IBE. Van Fraassen’s (1989) bad lot argument and Stich’s (1990) argument about the better-safe-than-sorry beliefs are defended and elaborated to argue against the truth-conduciveness of IBE. Laudan’s (1981) pessimistic induction argument is applied to the issue of IBE. An argument about the possible incommensurability of explanatory virtues is produced for the problem at hand.

**Structure of the thesis.** The thesis has two main chapters that correspond to the two aspects of the main problem delineated above. The first chapter reconstructs IBE and argues for the explanationist account as an adequate account of IBE. IBE is commonly described as consisting of two steps or two filters: the first gathers a pool of potential explanations and the second selects the best explanation out of the pool of potential ones. Therefore section 1.2 provides an account of abduction as the first step of IBE and the next two sections present two alternative accounts of the second step: the explanationist account (1.3.1) and its defense (1.4.2), according to which, the

explanatory power of a hypothesis is the degree of explanatory virtues present in the hypothesis and the probabilistic (Bayesian) account (1.4.1) and its critique (1.4.2), according to which, the explanatory power of a hypothesis is a particular product of the prior probability and likelihood of the hypothesis. The second chapter evaluates whether IBE can be granted truth-conduciveness. The chapter enumerates four basic ways in which the truth-conduciveness of IBE is justified, argues that these means of justification do not grant the truth-conduciveness of IBE and finishes with the section (2.6) which summarizes arguments that refute the truth-conduciveness of IBE. Subsections on psychological adequacy (2.3.1) and ontological commitments of IBE (2.5.1) are apposite to some of the justifications of IBE and are provided in the relevant parts of the second chapter.

**Conclusions.** 1. IBE is a form of material inference that ascribes truth to the hypothesis that has the highest degree of explanatory virtues among its competitors: it is the most consistent with approved background knowledge, the most unifying, the deepest and the simplest. If a hypothesis is not consistent with or does not restore consistency in the relevant background knowledge, then the hypothesis cannot be an abductive conclusion and, because of that, cannot be included in the set of potential explanations. In addition to consistency unification as an explanation of several different kinds of phenomena, explanatory depth as an explanation of why the proposed explanation should be true, and simplicity as fewness of posited explanatory entities or shortness of expression of an explanation are features attributed to explanations having a sufficient rationale to be accepted. Therefore, if explanations exhibiting these features are good explanations, then explanations exhibiting the highest degree of consistency, unification, depth and simplicity are the best explanations. Nevertheless, there are two inconsistencies in the theories of IBE:

(a) The theories that treat coherence as the main explanatory virtue are circular, because coherence itself stands for consistency plus explanatory relations. Therefore, explanatory power and coherence should be considered to stand for the very same phenomenon and explicated as derivable from the rest of explanatory virtues.

(b) Probabilistic measures of explanatory power are superfluous: if probability distributions are known, then the posterior probability formula is the correct way to identify the most likely hypothesis, but this ceases to be IBE; if probability distributions are not known, then one cannot apply probabilistic measures of explanatory power and has to rely on the explanationist account of explanatory power. Either way, there is no actual need for the probabilistic measures of the explanatory power. The explanationist account of explanatory power is sufficient for the task at hand.

2. Present theories of IBE can be classified into four basic ways of justifying the truth-conduciveness of IBE:

(a) The reliabilist-coherentist way claims that IBE is truth-conducive, because it is reliable—prone to produce true beliefs rather than false ones—and it is reliable, because it enhances the total coherence of knowledge.

(b) The evolutionary way claims that IBE is truth-conducive, because reasoning following IBE is a survival-enhancing human cognitive capacity and the best explanation for why IBE is survival-enhancing is because IBE provides true beliefs.

(c) The probabilistic way claims that IBE can be defined probabilistically and IBE is truth-conducive, because hypotheses with the highest degree of explanatory virtues appear to have the highest posterior probability.

(d) The empirical-historical way claims that IBE is truth-conducive, because hypotheses with the highest degree of explanatory virtues are empirically confirmed and truth or at least approximate truth is the best explanation for the empirical confirmation.

3. None of the four ways of justification discerned above grants the truth-conduciveness of IBE:

(a) Better-safe-than-sorry argument refutes the evolutionary way of justification: there are survival-enhancing beliefs that do not have to be true.

(b) Pessimistic induction refutes the empirical-historical way of justification: there were instances of best explanations that were accepted as true, but later appeared to be false.

(c) The possibility of contradicting orders of explanatory power undermines the reliabilist-coherentist way of justification: one can never be sure that there is no alternative, and more adequate, ranking of explanatory power.

(d) The bad lot argument undermines the reliabilist-coherentist way of justifying IBE: the set of potential explanations can never be exhaustive and the use of catch-all hypotheses cannot correct that. Pessimistic induction strengthens the bad lot argument: every example of a superseded theory from the pessimistic induction argument is an example of the bad lot situation in action; a better theory, which superseded the old one, was not considered when the old one was accepted. The possibility of contradicting orders of explanatory power strengthens the bad lot argument: even if we picked a particular hypothesis as the best one, another hypothesis that is better according to some other explanatory virtues can be the actual explanation.

(e) The material nature of IBE contradicts an application of the probabilistic way to argue for the truth-conduciveness of IBE. Probabilistic justification would make IBE deductively, i.e., formally valid. However, IBE is a non-deductive kind of inference; its alleged validity is not formal, but material, dependent on the substantive considerations.

(f) Further research on the concept of material inference has to be conducted in order to assess whether there is any way to show that IBE is truth-conducive.

4. Even though, for the time being, IBE cannot be showed to be truth-conducive, IBE is a widespread psychological practice to provide hypotheses that can later be empirically tested. After successful empirical tests, conclusions of IBE can be given credence, but not before. Therefore, IBE is warranted at most pragmatically: there is currently no better kind of ampliative inference and its application helps us to successfully cope with the world. However, pragmatic warranty does not imply the truth-conduciveness of IBE.

**Publications on the topic of dissertation:**

Psychological Adequacy and Ontological Commitments of Inference to the Best Explanation, *Problemos* 79, pp. 41-54, 2011.

Geriausio paaiškinimo išvedimas. Tarp dedukcijos, indukcijos ir abdukcijos, *Problemos* 76, pp. 150-161, 2009.

**Presentations at conferences:**

What Does the Mind Tell About the Inference to the Best Explanation?, Fourth Conference of the Dutch-Flemish Association for Analytic Philosophy (VAF IV), Catholic University of Leuven, January 2010.

How to Measure and an Epistemic Justification of Inference to the Best Explanation?, Central European University Summer School, course "Probabilistic Causality", July 2008.

**Education:**

International scholar at Institute of Philosophy, Katholieke Universiteit Leuven, Belgium (October 2009 - July 2010).

Doctoral studies in Philosophy at Department of Logic and History of Philosophy, Vilnius University, Lithuania (2007-2011).

Master's degree in Philosophy at Vilnius University, Lithuania (2007).

Bachelor's degree in Philosophy at Vilnius University, Lithuania (2005).

Erasmus exchange studies in Philosophy, University of Skövde, Sweden (September - December 2003).