# Měření v psychologii

[TODO: je to spíš detailní osnova, tj. všechno přepracovat, rozvést, doplnit, dopsat a restrukturovat jako souvislý text]

* Psychologové rádi měří. A rádi říkají, že měří takové atributy jako intelektové schopnosti, obecné („inteligence“) i specifické (verbální, prostorové…), intenzity různých druhů počitků (hlasitost, jas…), postoje (k různým celospolečenským tématům), osobnostní rysy (extroverze, neuroticismus…), dovednostiv různých oblastech, od konkrétních (psaní, počítání…) po abstraktní (kreativita), a mnoho dalšího. Zakladatelé psychologie předpokládali, že měření je pro vědu nutné, a svoje měřící procedury za měření pokládali. Psychologie se ustanovila jako kvantitativní věda, a tato představa nebyl dlouho rozporována. [3]
* V posledních 25 letech se ovšem stále častěji ozývají hlasy, které se tuto stále přetrvávající iluzi snaží uvést na pravou míru: Michell ([3], [?]) sveřepě kritizuje, že psychologův netestují, zda jsou psychické atributy skutečně kvantitativní, a dokonce – poněkud přehnaně – označuje psychometriku za patologickou vědu ([50]); Cliff ([48]) poznamenal, že psychologové nebyli schopni využít poznatky RTM – další dodali, že ani nebudou ([42], [10]); Borsboom ([49]) si povzdechl, že psychologové nejsou schopni používat aktuální matematické nástroje psychometriky; ovšem Schönemann ([51]) na druhou stranu psychology naopak peskuje za to, že jsou příliš okouzlení moderními způsoby analýzy dat, jako je strukturní modelování či neuronové sítě, a že by se místo snahy využívat matematiku, podobně jako fyzikové, měli soustředit na empirické zkoumání „neurčitostí živých bytostí“ – oproti jiným, kteří naopak tvrdí, že teorie psychologů nestojí za nic, protože jsou příliš v zajetí empiricismu, či rovnou operacionalismu ([3], [?]).
* Jisté tedy je pouze to, že měření v psychologii rozhodně není tak jednoduché jako v přírodních vědách. [TODO: proč] Zatímco v přírodních vědách máme alespoň k některým měřeným atributům přímý přístup, v psychologii tomu tak není; nejenže s duševními atributy můžeme těžko manipulovat, ani nevíme, jakým způsobem a zdali vůbec jsou manifestovány ve fyzickém světě.
* Z popisu měření v předchozí kapitole vyplynula dvě zásadních tvrzení, které mají zásadní vliv na měření v psychologii: a) pro měření jsou nejvhodnější atributy, které mají kvantitativní strukturu; b) mezi měřeným atributem a výsledkem měření musí existovat kauzální vztah.
* Touha, aby byly psychické atributy kvantitativní, pramení z toho, aby je bylo možno mezi sebou vztahovat pomocí matematických rovnic – podobně jako ve fyzice jsou mezi sebou provázané veličiny. Kauzalita pak velmi těsně souvisí s validitou.

## [Jsou atributy kvatitativní?]

### Snaha o psychologii jako kvantitativní vědu

[TODO: Fechner a tak dál]

### Campbellův omyl

[TODO: rozšířit]

* Campbellovo pojetí fundamentálního měření kladlo velmi přísné podmínky na atributy, aby je bylo možno považovat za měřitelné: pokud atribut nemá operaci konkatenace, nelze ho měřit. Campbell definoval měření pomocí omezení epistemologického procesu zapojeném v přiřazování čísel objektům. [21], [3] [TODO: byla to chyba, proč převzít z předchozí kapitoly]
* Když Campbell popsal svoje pojetí měření, bylo jasné, že je pro psychology nepoužitelné, neboť psychologické atributy nemají žádnou operaci konkatenace, a zároveň je nelze definovat pomocí odvozeného měření. Byla tedy otázka, zdali jsou takové atributy kvantitativní a tímádem vůbec měřitelné. [21], [3]

### Fergusonova komise

[TODO: rozšířit]

* British Association for the Advancement of Science se pokusila problém měřitelnosti psychologických atributů vyřešit ustanovením setkání expertů, fyziků (včetně Campbell) a psychologůů, které je známé jako Fergusonova komise. Protožoe psychologové nebyli schopni prokázat empirickou smysluplnost matematické operace scčítání psychologickýh atributů, komise pod vlivem Campbellová tábora dospěla k závěru, že není možné nazývat procedury, kterými psychologové určují hodnoty atributů, měřením. [21], [3]

### Stevensův omyl

[TODO: rozšířit]

* Stevens, jehož škála Sone byla rovněž předmětem jednání Fergusonovy komise, se – stejně jako ostatní psychologové – nechtěl měření vzdát. Vzal tedy Campbellovu definici měření jako „přiřazování číslic objektům nebo událostem podle pravidla“, odstranil upřesnění „pravidla“ v podobě konkrétní specifikace empirických podmínek, které musí atribut splňovat, a ještě navíc Campbella drze ocitoval. Stevensův operacionalistický přístup k měření převládá, ke vzteku Michella a jiných ([?]), dodnes. [21], [3]

### [Revolution that never happened]

[TODO: rozšířit]

* Problém s neexistencí operace konkatenace měla vyřešit reprezentační teorie měření, obzvláště spojené měření, ale to se nestalo. [TODO: proč? Chyby RTM v předchozí kapitole, [48], [?]]

### [Současná situace + latent variables models]

[TODO: rozšířit]

* Přístup latentních proměnných řeší problém chyby měření tím, že předpokládá, že data obsahují šum, a modeluje očekávané hodnoty pozorovaných proměnných. Tím překonává obtíže s vytvořením deterministického relačního systému. Model s latentními problémy schovává relační systém pod „nepozorovatelnost“ proměnných, tudíž fakt, že relační systém nelze vytvořit není dále potřeba řešit. Na druhou stanu, struktura atributu nemůže být založena na pozorovaném empirickém relačním systému. Data, obsahující chybu, mohou být k latentní struktuře vztažena celou řadou často rozdílných modelů. Další problémy se objeví s intepretací měření: a) je těžké vytvořit intepretaci očekávaných hodnot; b) konceptualizace kauzálních vztahů je obtížnější. [NOTE: ne že by u jiných teorií měření byla snazší…] [21]
* Modely s latentními proměnnými fungují následujícím způsobem. Předpokládá se, že psychologický atribut má určitou strukturu; tento předpoklad by měl být podpořený teorií. Struktura může být kategorická, pořadová či spojitá. Na těchto nepozorovatelných, latentních proměnných statisticky závisí množina proměnných pozorovaných, které mohou být rovněž kategorické, pořadové či spojité. To znamená, že očekávané hodnoty dat jsou pomocí rovnic vztaženy k latentní proměnné (či více latentním proměnným). Rovnice takového typu nicméně obsahují více neznámých než známých proměnných, takže nemají jedinečné řešení. Pokud ovšem více pozorovaných proměnných závisí na stejné latentní proměnné a předpokládádme určitý funkční vztah s latentní strukturou, parametry modelu mohou být odhadnuty, model má testovatelné důsledky a je falzifikovatelný. Jedním z důsledků je lokální nezávsilost: protože kormě náhodné chyby závisí pozorované proměnné zcela na latentní, jsou, podmíněně k této latentní proměnné, vzájemně statisticky nezávislé. [21]
* Použití regresních modelů vyžaduje, aby data byla kvantitativní, nicméně v psychologii je to netestovaný předpoklad – možná proto je tolik výsledků výzkumů jen „pomíjivým popisem situaci, které již nikdy nenastanou a které lze popřít téměř jakoukoliv replikací výzkumu“. [52]

#### Bayesovské frameworky pro spojené měření

[TODO: snahy o pravděpodobnostní ověřování axiomů RTM; možná to hodit jako „rámeček“ mimo hlavní proud textu?]

## Kauzalita

[TODO: rozšířit]

* Předpokladem psychometrických modelů je, že mezi atributem a testovými skóry existuje nějaká asociace; někteří autoři požadují, aby tato asociace měla podobu kauzální relace. Předpokládá se, že nějaký rys testovaných osob způsobuje odchylky v jejich skórech. Pokud měřící nástroj fnguje tak, jak má, tento rys je to, co chceme měřit. Kauzalita je tedy hlavní podmínkou intepretace testových skórů jako výsledků měření a validity testu. [21]
* Vztah mezi atributem a jedím skórem lze konceptualizovat jako kauzální relaci ve dvou směrech: atribut způsobuje skór, nebo skór způsobuje atribut. Toto rozlišení vede ke dvou základním konceptům modelů: reflektivní model (atribut -> skór) a formativní model (skór -> atribut). Každému z nich lze přiřadit určité filozofické pohledy; kauzalita od atributů ke skórům vyžaduje určitoi formu realismu vzhledem k atributům, zatímco opačná kauzalita vede spíše k pojetí atributů jako konstruovaných z pozorování. [21]
* [TODO: rozvést: protože pro měření jsme určili jako podmínku kauzalitu atribut->výsledek měření, použitý model musí být koncepčně reflektivní]
* Z jediného vztahu atribut–skór nemůže usoudit na směr kazuality; teorie je vzhledem k datům podurčená. Reflektivní model se proto rozšiřuje tak, aby zahrnoval víc pozrovoaných proměnných, tj. test skórů, závisejíccíh na stejné latentní proměnné, tj. atributu; výsledkem je model společné příčiny (common cause model). Ten implikuje, že proměnné následku jsou vzájemně nezávislé podmíněně na společné příčině. [21]
* [TODO: formativní modely stručně popsat]
* Základní předpoklady o kazualitě: daný směr (příčina->následek, nikoliv následek->příčina); časová posloupnost (příčina, potom následek); tranzitivita (a->b, b->c => a->c); korelace neimplikuje kauzalitu, kauzalita implikuje korelaci; manipulace s příčinami je doprovázena pozorovanými změnami v následcích, naopak nikoliv; příčiny jsou odlišné od následků, nic nemůže způsobovat sebe, nebo část sebe. [21]
* Co přináší kauzální pojetí měření oproti statistickému: kauzalita přináší důvod, proč intervence s konstruktem ovlivní skóry; kauzalita přináší důvod věřit, že pravděpodobnostní vztah přetrvá i do budoucna; kauzalita přináší důvod proč věřit, že tento pravděpodobnostní vztah bude odolný proti vnějším vlivům. [21]
* Regularity theories – základní problém indukce do budoucnosti: předpokládáme, že výsledky (v podobě pravděpodobnostních vztahů atributy<->skóry) budou platit pro populaci a pro budoucí měření. Pokud člověk tuto indukci přijme, výhoda kauzální intepretace spočívá v posílení podpory takového závěru. Co taková intepretace přináší: kauzální model je robustnější proti změnám v populaci, které by mohly změnit pravděpodobnostní vztahy. [21]
* Kontrafaktuální teorie – vysvětlují kauzalitu na individuální úrovni. Teorie sice rovněž zobecňují minulé události, ale dovolují víc než predikci – vysvětlní: pokud by nenastala změna v atributu, nezískal by respondnet určité skóre. [21]
* Procesní teorie – popisují proces, který dává vzniknout kauzálním vztahům definovaným rovnicemi/grafem modelu; kauzální proces generuje kauzální strukturu, která generuje spojenou pravděpodobnost. [21]
* Je sotva možné konstruovat uvažování tvůrce testu pomocí ne-kauzálního přístupu. Položky mají mnoho společných vlastností, a změna některé – např. písma, konkrétní formulace otázky, změna číslic na slova vyjadřující čísla… – by neměla ovlivnit výsledek, protože tyto vlastnosti nehrají kazuální roli ve vyvolání konkrétního procesu odpovídání. [21]

## Validita a validace

* Validita není komplexní, fasetovaná, nebo závislá na nomologických sítích – je to jednoduchý koncept: test je validní pro měření atributy právě tehdy, když atribut existuje a variace v tomto atributu kauzálně vytváří variace výstupu měřící procedury. Základní tvrzení je tedy ontologické, a logicy předchází procesu měření, protože je to nutná podmínka, aby bylo měření možné; toto tvrzení je nicméně odlišné od samotného procesu měření. Sjednocení všech procedur týkajících se ověřování této vlastnosti testů, oněch různých „validit“, testování vhodnosti, atd., nejsou validitou, ale *validizací*. [38]
* Měření je kauzální koncept, s jasně daným směrem svět → nástroj; nemůže být tedy o korelaci. Kriteriální validita vede k tomu, že test je „validní“ pro spoustu věcí – šance potkat nulovou korelaci je téměř nulová. Dále – pokud by měly dva testy perfektní korelaci, znamenalo by to, že měří stejnou věc pod dvěma názvy. Korelace je rovněž statistika závislá na populaci – a proto by byla i validita na ní založená. [38]
* Skór testu může a nemusí kauzálně záviset na změně měřeného atributu, což je důležitý faktor vzhledem k validitě testu. Tvrzení „atribut X způsobuje skór Y“ musí mít podmínky, za kterých je pravdivé. Tyto podmínky závisí na analýze termínů atributu, skóru a kauzality, a empiricky testovatelných implikacích jejich intepretací. [21]
* Regularity theories vyžadují nejen závislost skórů na atributu, ale i vyloučení jiných vysvětlení (tj. alternativních modelů). Kontrafaktuální teorie vyžadují další důkazy, že závislost platí napříč možnými případy, nejen mezi hodnotami pozorovanými v populaci. Proesní teorie vyžadují vytvoření vysvětlení závislosti, tj. popis, jak kauzální proces generuje odpovědi na položky. [21]
* Průřezová data napříč jednotlivci se mohou lišit od procesů probíhajících v jednotlivci. [TODO: Simpson’s paradox].
* Validace inter-individuální kauzality vyžaduje méně důkazů než kauzalita napříč i uvnitř jedinců. Intra-individuální kauzalita znamená, že obdržíme stejné výsledky napříč jedinci i napříč různými okolnostmi u jednoho jedince. Empirické důkazy pro toto se hledají těžko: intra-individuální efekt nemusí existovat, či funguje jinak než inter-individuální. Musíme tedy používat metodologie s opakovaným testováním jedince, či manipulováním formátu položek ([TODO: návaznost na automatické generování položek]). Nejlepší důkazy o validitě se sestávají z různých druhů studií, které potvrzují jednu intepretaci. [21]

### [Jak zajistit kvantitativnost atributů a kauzalitu v testech?]

* [TODO: pořádné teorie a pravé experimenty; Problém kvantitativnosti psychických atributů by se vyřešil tím, pokud by mohly fungovat v teoretických vztazích mezi sebou.]
* Šum nastává díky kauzální komplexitě světa, a můžeme rozlišit dva druhy chyby: náhodné a systematické. S náhodnou chybou se můžeme vypořádat pomocí statistických metod (TODO: gauss a jeho teorie chyby). Se systematickými chybami se lze vypořádat pouze tím, že je kontrolujeme (držíme je konstantní, izolujeme, zahrneme do modelu, znáhodníme…) v experimentu. V přírodě nicméně není rozdíl mezi zkoumaným fenoménem a šumem – jsou jen příčiny a následky. [45]
* Selhání při kontrole systematických chyb vede k tomu, že odhalování kvantitativní struktury je téměř nemožné. [45]
* Kvantitativní strukturu atributu nelze odhalit bez pomocí přístrojů a aparatur. [45]
* V experimentu manipulujeme příčinami a pozorujeme, jaký je vliv na následky, v podobě fenoménu ovlivňují fenoménu. Zároveň kontrolujeme šum. [45]
* Michell [3] tvrdí, že vědecký úkol předchází instrumentálnímu; [NOTE: což nicméně není pravda, vzájemná iterace, atd. [23]].
* Vědecký pokrok je možný i bez experimentů, astronomie i teorie evoluce se bez nich obešly. [47]

### [Proč bayesovská statistika?]

[TODO: Proto.]

[TODO: bayesovská statistika a filozofie vědy – možná spíš do přílohy?]

## [Ontologický status atributů]

[TODO: Mind-body problem, reduction problém]

## Modely

### CTT

[TODO: rozšířit]

[38]

#### Formální pohled

##### Syntax

*True* skór je očekávaná hodnota naměřeného skóru.

Chybový skór je rozdíl mezi skutečným a pozorovaným skórem.

##### Sémantika

Očekávaná hodnota měřeného skóru vzniká na úrovni jedince, ale toho ve skutečnosti nemůžeme měřit opakovaně; taková měření by totiž musela být paralelní, a pravý skór by musel být konstantní. Tedy opakovaná měření by musela pocházet ze stacionárního náhodného procesu, což je ale nerealistický předpoklad – pokud by měření pocházela z náhodného procesu, jistě by nebyl stacionární.

Lord and Novick proto přišli s myšlenkovým experimentem s Mr. Brownem, kterému je mezi pokusy „vymýván mozek“. Nulová očekávaná hodnota chyb tak nemůže být brána jako empirický předpoklad – zatímco teorie chyb počítá s tím, že chyby se při opakovaných měření vyruší, a očekávaná hodnota měření tak může sloužit jako odhad skutečné hodnoty, u CTT toto neplatí, protože místo toho, aby byla nulová očekávaná hodnota chyb brána jako předpoklad, na kterém lze založit hypotézu o tom, že očekávaný skór se rovná pravému, Lord a Novick nejprve definují pravý skór jako , a nulová očekávaná hodnota chyb je odvozena jako důsledek (protože ).

#### Empirický pohled

Protože CTT neříká nic o datech, je trvale imunní proti falzifikaci, a je třeba ji rozšířit o další předpoklady o vztahu mezi pravými skóry různých testů:

* Paralelismus – Testy a jsou v populaci paralelní, pokud přináší stejné očekávané hodnoty a rozptyly měřených skórů pro všechny subpopulace včetně těch, které se skládají z jednoho člověka)
* Tau-ekvivalence – Testy a jsou tau-ekvivalentní, pokud přináší stejné očekávané hodnoty, ale jiné rozptyly.
* Tau-ekvivalence v podstatě – Testy a jsou v podstatě tau-ekvivalentní, pokud nepřináší stejné očekávané hodnoty ani rozptyly, ale očekávané hodnoty jsou lineárně spojené rovnicí kde *c* je konstantní přes osoby.

##### Reliabilita

Reliabilita je na populaci závislý index přesnosti měření:

Protože toto stále obsahuje nepozorovatelný true skór; pokud bychom však použili respondenty s brain-washingem mezi měřeními, mohli bychom tyto replikace ( a ) považovat za paralelní testy a korelace mezi naměřenými skóry by se rovnala:

Nyní je třeba interpretovat kovarianci mezi dvěma proměnnými jako rozptyl jedné proměnné a součin dvou směrodatných odchylek odlišných proměnných jako rozptyl jedné proměnné – to ovšem vyžaduje, aby tyto proměnné byly jedna a ta samá: . Korelaci mezi paralelními testy lze považovat za reliabilitu jednoho testu pouze pokud tyto dva testy měří stejný *true* skór.

Abychom mohli někdy mluvit o změně pravého skóru a počítat *true gain score*, musí mít smysl odlišit *true* skóry mezi různými časovými okamžiky; ono vymývání mozku mezi replikacemi stejného testu tedy musí mít podobou časového „skoku zpět“ do původního okamžiku první administrace – pravý skór je tedy definovaný pro jediný okamžik, a můžeme psát:

My ovšem nechceme počítat reliabilitu na základě nerealizovatelných myšlenkových experimentů; přesný odhad reliability nicméně nelze získat empiricky, lze získat pouze dolní hranici odhadu, pokud předpokládáme:

* dvě *skutečné* replikace testu jsou paralelní (test–retest metoda), nebo
* dva *různé* testy jsou paralelní (metoda paralelních testů)

Test–retest metoda – Je nutné považovat měřený atribut za stabilní v čase, spolu s předpokladem paralelnosti replikací testu; vlastnosti myšlenkového *brain-washing* experimentu musí být rovněž zachovány: paralelismus a nezávislost opakovaných měření, což je značně nerealistické.

Metoda paralelních testů – Nejenže bylo hledání paralelních testů doposud – kvůli přísným empirickým kritériím – neúspěšné, představa dvou různých testů, které jsou paralelní, je sémanticky nekonzistentní.

Aby byla kovariance mezi dvěma *true* skóry interpretována jako rozptyl jednoho, musíme opět předpokládat . Ovšem pravé skóry rozdílných testů jsou odlišitelné právě proto, že jsou definovány vzhledem k různým testům; i kdyby byly empiricky rovnocenné, nejsou logicky stejné. Syntax CTT nedokáže popsat, co pro znamená pro dva různé testy, že měří jeden atribut. Dva testy mohou měřit to stejné pouze pokud se jedná o opakovanou administraci jednoho testu.

#### Ontologický pohled

Mohlo by se zdát, že true skór odpovídá konstruktu, který se snažíme měřit. Pokud by to však platilo, CTT by mohla přepsat vztah měřený skór → konstrukt jako měřený skór → pravý skór – to ale nedokáže.

Lord a Novick: „Pozorovaná proměnná je mírou teoretického konstruktu, pokud jeho očekávaná hodnota monotonicky roste s konstruktem…“ Ovšem true skór je sám definován jako očekávaná hodnota, takže nahrazení do předchozí definice („Pozorovaná proměnná je mírou [true skóru], pokud jeho [true skór] monotonicky roste s [true skórem]“) vede k nesmyslné tautologii. Vztah měření tak musí být vztahem mezi pravým skórem a *něčím jiným*.

Navíc, z definice *true* skóru plyne, že CTT nepředpokládá, že základem měření nějaký konstrukt vůbec je. Nejenže má každý test vlastní pravý skór, ale je rovněž možné vymyslet situace, kdy validní konstruktový skór neodpovídá *true* skóru: máme-li test, který diagnostikuje nějakou chorobu s určitou senzitivitou (např. 0,8), tak pacient mající chorobu (tj. konstrutkový skór 1) má jiný pravý skór (tj. očekávanou hodnotu 0,8). Existuje-li možnost špatné klasifikace, očekávaná hodnota měřeného skóru se nikdy nemůže rovnat konstruktovému skóru.

Existuje vůbec true skór? Pravý skór nelze konceptualizovat jako skutečné opakování testové situace, stejně to platí i pro chybové skóry. Co tedy způsobuje náhodnou chybu, tj. fluktuace, které definují pravděpodobnostní rozložení pravého skóru?

Buď nic (zdroj chyby je tedy vskutku záhadný), nebo vše (ale pak se Mr. Brown může změnit tak, že bude prakticky odpovídat jiné osobě), nebo jen něco – my pak definujeme, které faktory jsou při replikacích konstantní, a které ne. Definice náhodné chyby je pak ovšem cirkulární → co se má měnit, abychom našli vhodnou intepretaci náhodné chyby? – nesystematické věci – které to jsou? – Mr. Brown má bolest hlavy, hluk za oknem, atd. – proč tyhle věci? – mají vliv a mění se napříč replikacemi – proč se mění? – protože jsou zahrnuty jako variující v našem myšlenkovém experimentu.

Realistická intepretace *true* skóru je tedy k ničemu – nevede k testovatelným hypotézám, a protože každý test má svůj pravý skór, vede to ke zbytečnému množení jsoucen. Díky těmto důvodům neposkytuje CTT sama o sobě žádné spojení s pozorovanými hodnotami.

Použitelná je pouze operacionalistická interpretace, a to přestože jsou tyto operace jen hypotetické (myšlenkový *brain-washing* experiment): CTT vede k multiplikaci teoretických termínů, true skór je definován jako očekávaná hodnota určitého testu. Pravé skóry je nicméně možné definovat jako dispozice, ale *brain-washing* je stále třeba zachovat pro smysluplnou interpretaci.

### Latentní proměnné

[TODO: rozšířit]

[38]

#### Formální pohled

##### Syntax

Model vztahuje pozorovanou odpověď na položku U k latentní proměnné θ pomocí funkce:

kde g je spojovací funkce; E(Uij) je interpretováno jako očekávaná odpověď respondenta i na položku j, nebo jako očekávaní odpovědi na položku v populaci respondentů s pozicí θi na latentní proměnné; αj a βj jsou specifické parametry položky.

Model nelze testovat kvůli přítomnosti latentní proměnné přímo, pouze nepřímo díky implikacím pro sdružené rozložení pravděpodobnosti odpovědí na položky; odpovědi položky jsou nezávislé, podmíněné pouze latentní proměnnou.

##### Sémantika

To, že je odpověď na položku realizací náhodné proměnná, lze interpretovat dvěma způsoby.

Stochastický subjekt – Funkce odpovědi na položku určuje pravděpodobnost, s jakou respondent s určitou pozicí na latentní proměnné odpoví na danou položku. Vyžaduje podobný myšlenkový experiment jako CTT.

Mezisubjektový pohled – Soustředí se na charakteristiky populace, a nevyžaduje tak onen „brain-washing“. Pravděpodobnostní rozložení odpovědí na položky, podmíněné latentní proměnnou, je koncipováno jako opakované vzorkování z populace se stejnou hodnotou latentní proměnné; funkce odpovědi na položku určuje očekávání odpovědí na položku v populaci s určitou pozicí na latentní proměnné. Očekávaná hodnota tak může být brána jako subpopulační průměr.

Srovnání – Pokud bychom použili první interpretaci, např.. že člověk s hodnotou rysu R má pravděpodobnost P, že správně odpoví na položku, naznačuje to spojení s interpretací druhou – protože i v tom případě by poměr lidí z populace s hodnotou rysu R byl roven P. Tato podmínka lokální homogenity (tj. že model má stejnou formu mezi i uvnitř subjektů) nicméně nevyplývá z modelu, a navíc je netestovatelná, neboť vyžaduje „*brain-washing*“.

#### Empirický pohled

Testy vhodnosti modelu trpí problémem podurčenosti – je mnoho způsobů, jak mohou vzniknout data, která odpovídají modelu.

Z empirického pohledu je latentní proměnná viděna jako operační latentní proměnná – která nicméně tak latentní není, neboť je funkcí pozorovaných proměnných.

#### Ontologický pohled

Předpoklad, že to byl právě daný model, který generoval data, předchází procesu odhadu – pokud odhadujeme hodnotu latentní proměnné pro nějaké respondenty, činíme tak na základě daného modelu. Tento odhad ovšem není odhadem formální latentní proměnné; protože ovšem používáme formální část modelu k získání informací o světě, je tedy nutné ontologicky propojit operační a formální latentní proměnnou.

##### Operacionalistická intepretace

Model nelze interpretovat z operacionalistického pohledu, neboť různé testy (sestavené z různých položek) by nemohly vycházet ze stejné latentní proměnné.

##### Realismus a konstruktivimus

Konstruktivisté odmítají tvrzení realistů; ti zastávají realismus teorií (teorie jsou pravdivé či nepravdivé) a realismus entit (aspoň některé teoretické entity existují).

Model latentních proměnných se drží realismu entit, protože taková forma realismu stojí za volbou obvyklého modelu měření v psychologii: kovarianci mezi indikátory lze vysvětlit regresí těchto indikátorů v závislosti na latentní proměnné (reflektivní model). Alternativou je formativní model: latentní proměnná je naopak závislou na indikátorech. Realistický přístup ovšem vyžaduje první model, neboť latentní proměnná předchází variaci v indikátorech, a determinuje je.

Realistická intepretace teorií souvisí s korespondenčním pohledem na pravdu (shoda mezi poznáním či teorií a skutečností). Teorie latentních proměnných nic o pravdě neříká, ale předpoklad korespondenční pravdy je formulován při aplikaci:

* při odhadu pozice subjektu na latentní proměnné (protože se při tomto mohu mýlit, musí existovat skutečná pozice)
* při odhadu parametrů (podobná situace jako předchozí – odhad je možný, jen pokud je co odhadovat)
* při podmíněném uvažování založeném na předpokladu, že model je pravdivý (protože model nemusí být pravdivý)

##### Kauzalita

Spojení mezi formální a operační latentní proměnnou vyžaduje realistickou ontologii; tímto spojením ke kauzalita (X může být příčinnou Y, pokud: a) X a Y kovariují, b) X předchází Y a c) *ceteris paribus*, Y nenastane pokud X nenastane).

Mezisubjektová kauzalita – Odpovídá definici kauzality, ovšem latentní proměnnou nelze odlišit od jejich indikátorů – tj. nelze odlišit následek a příčinu, což je ovšem problém kauzálních měření obecně.

Vnitrosubjektová kauzalita – Protože pozice respondenta na latentní proměnné je konstantní, nemůže mezi jí a odpovědí na položku být kovariance. Lze ovšem použít kontrafaktuální přístup (X způsobuje Y, protože pokud by nenastalo X, *ceteris paribus*, Y by nenastalo) – pokud by měl subjekt A jinou úroveň latentní proměnné, splnil by méně položek; to je ovšem pouze zamaskovaná mezisubjektová kauzalita, neboť stejně tak lze říci, kdyby měl subjekt A úroveň subjektu B, atd.

Mezisubjektový model latentní proměnné nemá jak popsat vnitrosubjektové procesy, které vedou k odpovědím. Vnitrosubjektový model kauzality by šel vytvořit pro takové proměnné, které jsou schopné empirické konkatenace, nebo pokud bychom dokázali, že mezisubjeketová a vnitrosubjektová struktura modelů je totožná. To je ovšem další předpoklad – nelze tvrdit, že model měření platí pro každého respondenta; rozhodně nelze tvrdit na základě mezisubjektových dat, že stejná struktura platí pro všechny jedince.

Poslední možností je použít eliptický přístup ke kauzalitě – vnitrosubjektové vysvětlení by bylo zkratkou pro individuální procesy. To by ovšem znamenalo, že pro různé lidi s různou pozicí na latentní proměnné (i se stejnou pozicí) by musely být procesy vedoucí k odpovědi homogenní – ve skutečnosti jsou pravděpodobně tyto procesy u různých osob kvalitativně odlišné.

#### Implikace pro psychologii

Lokálně homogenní konstrukty jsou takové, které lze vysvětlit latentní proměnnou pro vnitro i mezisubjektové rozdíly.

U lokálně heterogenních toto neplatí.

Lokálně irelevantní konstrukty jsou pak relativně stabilní rysy, jejichž struktura je odlišná – kovariance opakovaných měření nebude odpovídat latentnímu modelu.

Ty konstrukty, které nelze konceptualizovat na úrovni subjektu, lze tedy brát jako emergentní vlastnosti.

### Rasch jako případ spojeného měření?

[TODO: je Raschův model pravděpodobnostní variantou spojeného měření? (Ne.)]

### Bayesovské frameworky pro IRT

[TODO: úvod do bayesovské IRT]

### Hierarchické IRT modely

[TODO: hierarchické IRT modely]

[TODO: návaznost na AIG]

## [Technické záležitosti]

[TODO: paper vs. computer]

[TODO: CAT]

[TODO: návaznost na praktickou část: integrovaný (softwarový) testovací framework?]

[TODO: novinky v IT pro psychometriku: virtuální realita?]