

## **Oponentský posudek disertační práce**

**Mgr. Zuzany Bažantové: Využívání počítačů k testování.**

**Praha, PedF UK 2007.**

Autorka zvolila pro svoji disertační práci velmi aktuální téma. V devadesátých letech 20. století se i v českém prostředí začala dostávat do popředí pozornosti tematika didaktických testů. Bylo to důsledkem podnikatelských aktivit některých subjektů (např. Scio) a diskusemi a postupnou přípravou nové maturitní zkoušky v Cermatu. Rozvoj informačních a komunikačních technologií spolu s jejich rozšiřováním do škol nastoluje aktuálnost tématu předložené disertační práce.

Cílem práce je seznámit se současným stavem využívání počítačů při testování. Práce má tedy deskriptivní charakter. Skládá se z teoretické části a z části výzkumné. K dosažení cíle autorka zdařile směřuje v obou částech práce. Vyčtený cíl se podařilo autorce splnit.

Práce je rozsahu 160 stran s 20 stranami příloh. Text je psán velikostí písma 12 s jednoduchým řádkováním, takže rozsah 1 strany práce přesahuje 1 normostranu více než 2krát.

K práci mám následující připomínky diskusní povahy.

Rozsah práce je mimořádný a pro účely disertační práce možná i nadbytečný. Uvítal bych efektivnější jazyk vědecké práce. V některých pasážích se autorka opakuje (např. rozdíly mezi KTT a IRT „v IRT jsou charakteristiky položek měřeny nezávisle na schopnostech testovaných ...“ – s. 1, 80, 85, 118, 147), někde se detailně věnuje tématům, která jsou dostatečně popsána i v česky psané literatuře (některé části kapitoly 4. Tvorba didaktického testu). Někde se autorka věnuje problému, který s tématem práce souvisí spíše okrajově (kapitola 3 Využívání počítačů ve výuce). Na druhou stranu je téma zpracováno uceleně a výkladově srozumitelně. Práce může sloužit i jako studijní materiál pro každého, kdo se problematikou testování chce zabývat.

Pro výzkumnou část práce si autorka zvolila dva problémy: A. Jaký je současný stav využívání počítačů při testování. B. Čím se liší využití metod klasické teorie testů a item-response theory při tvorbě testu a analýze jeho výsledků (teoreticky a při konkrétní aplikaci).

Ve výzkumné části se objevují pasáže, které by spíše náležely do části teoretické (např. kapitola 6.2 Administrace počítačového testu.)

Ačkoliv autorka explicitně nepojmenovává použité metody vědecké práce, jsou z textu zřejmé (pro zmapování situace v zahraničí je to především na studiu založená analýza odborných publikací, internetových stránek společností zabývajících se testováním, analýza dostupných SW, v českém prostředí k tomu přispívá i vlastní shromažďování údajů od společností, které se testování věnují). Takto se autorce podařilo na 4 stranách (s. 100 – 103) zmapovat aktuální situaci ohledně testování u nás a ve světě a vyřešit tak problém A. Autorka sice chce vyslovit závěry k problému A až v kapitole 8 (soudím podle názvu kapitoly), ale ty zde jsou už příliš obecné a tato kapitola obsahuje formulace, které přísluší spíše do předcházejících kapitol (opět vracení se k výhodám počítačového testování, doporučení pro tvorbu testovacího programu apod.)

Druhý výzkumný problém autorka řeší za pomoci statistické analýzy dat z testu obecných studijních předpokladů, který byl použit v rámci přijímacího řízení na Pedagogické fakultě

UK v roce 2006. Autorka řešení problému zvládla a vhodně prezentovala výsledky pomocí statistických parametrů, tabulek a grafů.

K práci mám rovněž připomínky dílčího charakteru.

Na s. 29 nejsou zcela přesné formulace „Ačkoliv nikdy neznáme pravé skóry, můžeme zjistit neshody mezi neznámým pravým a známým naměřeným skórem. Rozdílem těchto skórů je standardní chyba měření“. Nemůžeme zjistit neshodu mezi konkrétním pravým skórem a naměřeným skórem. Známe pouze rozdělení pravděpodobnosti chyby. Kde standardní chyba měření je charakteristikou tohoto rozdělení.

Na s. 29 ve druhém řádku odspoda došlo patrně „k překlepu“ ve vzorci.

Logicky nekonzistentně působí tvrzení na s. 33 „Reliabilita zde (myšleno v IRT) nezávisí na délce testu. Reliabilitě testových skórů v KTT odpovídá v IRT množství informace... Množství informace lze určit informačními funkcemi.“ a níže „Ze vzorce plyne, že hodnota informační funkce testu závisí na počtu úloh v testu ... Jelikož s počtem položek roste množství informace, měří test jako celek danou schopnost mnohem přesněji než jedna položka.“ Předpokládám, že autorka disertační práce u obhajoby vysvětlí, za jakých okolností závisí a na jakých předpokladech stojí závislost reliability na délce testu v KTT a v IRT a jak si tedy vysvětlit zdánlivou nekonzistenci v jejím textu.

Na s. 125 nesouhlasím s hodnocením Cronbachova alfa jako „nejpřesnějším“ koeficientem reliability. Cronbachovo alfa souvisí se specifickou chybou měření ověřovanou pro vnitřní konsistenci testu. Nepokrývá jiné chyby měření, které autorka dobře mapuje na s. 30 a které se zjišťují jinými postupy. Další teorií zabývající se chybami měření v celku je teorie zobecnitelnosti (Generalizability Theory), které se autorka ve své práci nevěnuje.

Přínos práce je především v její aktuálnosti a komplexním uchopení problému. Problematika testování a využívání počítačů v něm nastoluje řadu výzkumných otázek pro pedagogy, psychology, statistiky, programátory. Vzhledem k současnému vývoji je nanejvýš potřebné tyto otázky pojmenovat a začít se jimi zabývat. Předložená práce je toho nutným předpokladem.

Autorka prokázala vysokou úroveň porozumění tématu (především u nás méně známým a přitom aktuálním problémům IRT a adaptivního testování) a schopnost samostatné vědecké práce na poli testování.

Práce má velmi pěknou grafickou úpravu, vhodně zpracované přílohy s cenným přehledem nejvýznamnějších světových organizací a firem činných v oblasti testování. Cenná a kvalitu práce dokumentující je i četná odborná literatura převážně zahraniční, ze které autorka čerpala.

**Vznesené připomínky diskusního charakteru a připomínky dílčí spíše formulačního rázu nijak nesnižují kvalitu předložené disertační práce. Práce splňuje požadavky kladené na tento typ prací a doporučuji její kladné přijetí.**

V Praze, 1. května 2007

PhDr. Martin Chvátal, Ph.D.

