

*Počítačové adaptivní testování a možnosti jeho využití v psychodiagnostice*

Adaptivní testování je považováno za efektivnější variantu klasických testů, které mají dopředu daný soubor položek. Abychom vyhodnotili, do jaké míry se nám vyplatí investice do adaptivních testů, je vhodné vědět, o kolik přesně je toto testování časově efektivnější při dosahování srovnatelné informační hodnoty. Tato otázka je hlavním výzkumným tématem hodnocené práce.

V teoretické části autorka srovnává klasickou testovou teorii a s ní spojený přístup s teorií odpovědi na položku. Osobně vítám tuto práci jako pokročilou nadstavbu toho, co studenty učíme na psychometrii a teorii psychodiagnostiky. Rovněž vítám, že se autorka pustila do oblasti, která nebývá studenty psychologie příliš navštěvována, neboť vyžaduje větší znalosti psychometrie, ale i programování. To je asi zároveň i odpověď na autorčinu otázku, proč je v České republice teorie odpovědi na položku oproti klasickému testování tak málo využívána. Přestože je tato mladší sestra klasické testové teorie ve světě známa již více než padesát let, byla první česká monografie o této metodě napsaná až po půlstoletí. Podle autorčina šetření existuje v české verzi zatím jen jeden adaptivní test. Porovnání obou přístupů je vyvážené, dobře postihuje výhody i nevýhody každého z nich. Sama diplomová práce ctí ideu adaptivního testování, která se snaží i při menším rozsahu a s ním spojené menší časové investici uživatele přinášet velké množství informací. Autorka přehledně popisuje základní ideje a předpoklady. V první kapitole stručně sumarizuje typy modelů a odhadů jejich parametrů. V druhé kapitole podává přehled hlavních rozdílů (a tedy výhod i nevýhod) teorie odpovědi na položku a klasické teorie testů, a pak přidává informaci o srovnávacích studiích. V kontextu těchto výzkumů je zajímavé, že složitost IRT postupů není vždy vyvážená jednoznačným ziskem. Většina studií ukazuje srovnatelnost výsledků, což by nás mělo vést k jasnější identifikaci podmínek, za kterých máme volit jeden nebo druhý postup. Jelikož je využívání adaptivního testování vázané na využívání počítačů, připojuje autorka kapitolu o specifičnosti počítačového a on line testování. Na některých místech bych měl pouze výhrady k přílišné, někdy až programátorské stručnosti, i zde doplňuje informace o srovnatelnosti výsledků testů administrovaných systémem papír-tužka a testů

administrovaných počítačově. Zde především, v závislosti na provedených studiích, upozorňuje autorka na nutnost ekvivalenci obou forem dotazování prověřovat a automaticky se na ní nespolehat. Ve čtvrté kapitole na základě výzkumů resumuje hlavní výhody a nevýhody adaptivního testování. Informuje zde o procentu ušetřených položek při stejné či obdobné informaci. Tím se vytváří prostor pro empirickou část práce. Ještě předtím ovšem autorka uzavírá teoretickou část přehledem kroků nutných k vytvoření a aplikaci adaptivního testu.

Teoretická tvrzení v empirické části prověřuje na Eysenckově osobnostním dotazníku. V empirické části provádí autorka nejprve simulaci výsledků adaptivního tetování oproti výsledkům klasického testování. Simulace nebývá v psychologii často užívanou metodou, tím spíše ji oceňuji, o to více, že jsou její výsledky následně porovnány se skutečnými daty. Autorka nesrovnává pouze adaptivní a neadaptivní testování, ale zajímá se rovněž o rozdíl počítačové administrace a administrace tužka-papír. Tyto výzkumy však nutně předchází samotné vytvoření databanky položek a naprogramování adaptivního testování. Autorka ověřuje základní předpoklady – unidimenzionalitu, lokální nezávislost a shodu modelu a dat. Po popisu metod i vzorku autorka náležitě testuje hypotézy. Jisté omezení spočívá právě v nesplnění základních předpokladů, které jsou však robustní, a tudíž nebrání tomu, aby byly dané metody použité.

Ukázalo se, že i při poloviční redukci délky testu zůstává korelace s testem celé délky 0,882. Korelace mezi administrací tužka a papír a počítačovou administrací byla 0,876. Při porovnání výsledků simulace a reálné administrace bylo rovněž dosaženo signifikantních korelací. Matematické vyhodnocení výsledků je velmi nadstandardně provedené, i s následnou diskusí. Samotný závěr již je příliš obecný a o konkrétních výsledcích nehovoří, ocenil bych stručné shrnutí hlavních výsledků výzkumu. V literatuře uvádí autorka více než 100 zdrojů. Přílohou práce je přiložený program. Na některých místech bych měl pouze výhrady k přílišné, někdy až programátorské stručnosti. Diplomová práce je ovšem jinak přehledná a užitečná, tematicky neobvyklá a v dané oblasti vysoce nadstandardní. Navrhuji hodnocení „výborně“.

Radvan Bahbouh

Katedra psychologie