

## PROBLÉMY PSYCHOLOGICKÉHO TESTOVÁNÍ POMOCÍ POČÍTAČE A V PROSTŘEDÍ INTERNETU

JAN ŠIRŮČEK

Institut výzkumu dětí, mládeže a rodiny, Fakulta sociálních studií MU

Text problematizuje užívání psychologických diagnostických metod pomocí počítače a v prostředí internetu. Diskusí technických a metodologických problémů vyšetření asistovaného počítačem dokládá nutnost zvláštní standardizace diagnostických metod určených pro užití pomocí počítače nebo jiného nestandardního přístroje, nebo v prostředí internetu.

### Úvodem

V posledním desetiletí se testování s podporou počítače, resp. testování v prostředí internetu stalo doslova bestsellerem zejména v oblasti řízení lidských zdrojů a psychologie práce a marketingu. Další oblasti také nezůstávají pozadu – pomocí metod upravených pro online administraci jsou realizovány sběry dat v aplikačním i základním výzkumu, v oblasti školní psychologie i čistě populární „vědě“ – např. <http://www.axlife.com/luscher/>. Počítačem řízené testování je nejen předmětem výzkumu, ale stále více i běžnou technikou sběru dat.

Výhody testování pomocí počítače jsou více než nasnadě. Z ekonomického hlediska představují možnosti okamžitě evidovat odpovědi respondentů a v reálném čase vyhodnocovat výsledky šetření až do úrovně vážených skóre, uchovávat takřka neomezené množství dat a případně i provádět vyšetření na dálku nepopíratelnou výhodou. Z odborného hlediska přibývá další podstatná výhoda, totiž možnost počítačového adaptivního testování (viz např. Květon, Klimusová, 2002; Jelínek, Květon, Denglerová, 2006). Dále přibývá možnost využít v testu podnětů v „tradiční“ diagnostice nemyslitelných bez mimořádných nákladů – animovaných obrázků, interaktivních testů, zvukových stop atd. Tyto se mohou velmi dobře uplatnit ve výkonových testech, zaměřených např. na elementární kognitivní kapacity.

S užitím počítačových verzí psychologických testů se však pojí i řada rizik. Některá jsou ryze technického charakteru (kompatibilita software) a lze je i úspěšně technicky řešit, další se však vztahují k samotné podstatě testování s pomocí počítače a musí být brána v potaz. V případě testování pomocí internetu totiž není možné zajistit ekvivalenci hardware – jinými slovy, počítače či jiné rozhraní, které budou mít různí respondenti k dispozici pro řešení testu, se mohou i velmi značně lišit. Standardní podmínky vyšetření tím zcela berou za své, a to se všemi důsledky.

### Psychometrické charakteristiky počítačově administrovaných testů

Pro počítačem administrované testy platí totožné nároky na reliabilitu, validitu, objektivitu a standardizaci metody jako pro klasické metody papír-tužka. Určité odlišnosti prezentace podnětového materiálu na displeji počítače a interakce člověka s počítačem však mohou přímo měnit situaci vyšetření a tím vést k podstatným změnám v povaze reakcí na jednotlivé testové položky oproti klasickým verzím testů. V důsledku toho lze oprávněně pochybovat o obsahové a konstruktové validitě metod, které jsou převedeny z původní „papírové“ do počítačem administrované podoby.

Příkladem mohou být takové testy, které pracují s grafickými podněty nebo jsou založeny na nějakém časovém limitu (rychlostní testy). Použijme takové testy jako příklad, dokladující problémy šetření v prostředí internetu.

V obou případech se zde promítá hned několik faktorů, které mohou vést ke zkreslení výsledků vyšetření. Tyto působí na straně hardware, tedy „situace vyšetření“ a současně na straně respondenta.

- a) Neekvivalence grafického podnětu na papíře a displeji počítače. „Tentýž“ podnět na displeji působí jinak, než v podobě papírové předlohy (Klimusová, Květon, 2002) a může současně být citlivější vůči poruchám vizuální percepce na straně respondenta.
- b) Kompetence respondenta k práci s počítačem. U rychlostních testů může mít samotný nezvyk na pohyb kurzorem po displeji, resp. na tzv. „dvojklik“ zásadní vliv na odvedený výkon. Tyto dovednosti přitom nelze v žádném případě považovat za samozřejmé.
- c) Odlišnost grafického podnětu na různých displejích. Díky variabilitě displejů na trhu (úhlopříčka displeje, rozlišení displeje, barevné spektrum, obnovovací frekvence) je prakticky nemožné nastavit základní parametry podnětu (zejména velikost) tak, aby se zobrazoval všem vzdáleným uživatelům stejně či alespoň podobně. Nemožnost takto standardizovat podnětový materiál činí jakoukoliv další práci na metodě, která to vyžaduje, de facto zbytečnou.
- d) Odlišnost ovládacích prvků. Nejčastěji používaným zařízením pro ovládání webových aplikací je myš, popřípadě touch-pad či trackpoint. Rychlost pohybu kurzoru a jeho přesnost může značně variovat v rámci těchto zařízení i mezi nimi. Možnost standardizovat tuto charakteristiku vzdáleně je vyloučena.
- e) Specifika různých prohlížečů na různých platformách. Různá vykreslovací jádra, různý design ovládacích prvků, velikost a typ fontů, rozteč písma, řádkování, šířka řádku, rychlost zobrazování a zpracování dat, rychlost připojení a kontext dalších vizuálních podnětů v pozadí spuštěné aplikace, popřípadě vliv jiných spuštěných aplikací... těchto vlivů technického charakteru mohou být desítky a všechny společně vedou k nemožnosti zajistit standardní podmínky vyšetření náročnějšího účelu.

Výhody a výzvy, které přináší počítačové testování, se tak v prostředí internetu stávají Achillovými patami: podmínky, které lze dobře kontrolovat v případě laboratorního vyšetření u standardního a prověřeného přístroje, popř. zaznamenat v inquiry a zohlednit při vyhodnocení výsledků, jsou při práci se vzdáleným, neznámým PC či notebookem zcela neovlivnitelné a téměř nezjistitelné.

Uvedené problémy pravděpodobně nenarušují platnost odpovědí např. v běžném osobnostním dotazníku, který není časově limitován, a respondent pouze volí mezi odpověďmi na všeobecně oblíbené x-bodové škále popisující jeho míru souhlasu s předloženým výrokem. Potom je respondentova práce v zásadě totožná s tou, kterou odvádí při vyplňování dotazníku metodou „papír-tužka“ či při třídění karet MMPI, je-li dotazník důmyslněji, tak, že administruje položky postupně, bez možnosti odpovědi zpětně upravovat. Zbývá pouze nutnost zajistit, aby se k počítači s otevřeným dotazníkem posadil skutečně ten, koho je cílem vyšetřit, pokud možno sám, bez podpory široké rodiny, přátel, kolegů či literatury a vhodného internetového vyhledávače. V oblasti HR, kde jsou respondenti skutečně motivováni k poskytnutí „správných“ odpovědí, může být naplnění této podmínky velmi obtížné.

V případě metod, které budou z hlediska využití možností počítače progresivnější, jako jsou námi uvažované příklady, však přináší uvedené nedostatky řadu komplikací, které vyúsťují v nutnost zohlednit všechny jednotlivé aspekty v procesu konstrukce a standardizace metody. Tedy, vyloučit intervenující faktory, které by mohly hrát roli při vzdálené administraci testu. Můžeme si položit otázku, zda je něco takového vůbec možné, pokud jde o úpravu prostředí pro testování. Lze se totiž pouze spolehnout na vůli respondenta k poctivé spolupráci.

Pokud jde o vyrovnání podmínek daných variabilitou hardware, nabízí se možností více. Z hlediska psychometrického nejde o nic jiného, než kontrolu a omezení zdrojů chybového rozptylu na různých úrovních a zajištění konstruktové, resp. obsahové validity metody. Pouze po zvládnutí těchto úkolů je možné vytvořit pro prakticky nový test adekvátní normy – a tedy věcně správně interpretovat výsledky.

### ***Konstruktová a obsahová validita***

K nutnosti zajistit konstruktovou resp. v užším smyslu obsahovou validitu metody se vztahují problémy spojené s odlišnou úrovní uživatelských kompetencí předpokládaných respondentů – internetem šířená metoda se nesmí stát testem těchto kompetencí, nebo respektive respondentova hardware. Podmínka se může zdát přehnaná, ale je třeba na ní trvat. Jedná se v zásadě o podobnou výchozí situaci, jako v případě kulturně podmíněných testů rozumových schopností. Jestliže má být metoda „počítačově férová“, je třeba zcela eliminovat takové úkony, které vyžadují větší zvyk na práci s počítačem (zminěný dvojklik, rychlý pohyb kurzorem, vyhledávání kláves, rychlé psaní atd.), doplnit sadu zácvičných úkolů, nebo zásadním způsobem omezit indikovanou populaci. Oba přístupy použitelnost internetem šířených testů výrazně omezují.

### ***Reliabilita – kontrola nespecifikovatelných zdrojů chybového rozptylu***

K problematizaci přesnosti měření přispívá variabilita software a hardware na straně respondenta – viz body c) a d) výše. Tyto faktory lze kompenzovat v zásadě jen systémem doporučení a pravidel pro vyplňování testu, jejich dodržení dost dobře nelze kontrolovat, a současně ani nemusí být v technických možnostech respondentů. V praxi se tento zdroj chybového rozptylu pravděpodobně projeví výrazným snížením reliability metody se všemi důsledky - omezením rozlišovací účinnosti, rozšířením intervalů spolehlivosti, zkrátka snížením jistoty interpretace. Nejednotnost zobrazení může zcela znemožnit smysluplné vyšetření třeba testem využívajícím doplňování matic, problémy s pomalou myší nebo trackpointem zase znemožní individuálního respondenta v libovolném rychlostním testu.

### **Shrnutí**

Možnosti využití počítačového testování (v prostředí internetu) přináší do psychologické diagnostiky řadu výzev. Na jedné straně umožňuje experimentovat s konstrukcí testového materiálu dříve nemyslitelným způsobem, současně ale přináší nové a velmi důležité výzvy pro zajištění technické a psychometrické kvality metod. Pokud není současně s adaptací testových metod do počítačového prostředí věnována zásadní pozornost jejich standardizaci, není možné tyto metody adekvátně používat a jejich výsledky věcně správně interpretovat – snad jen s výjimkou těch nejprimitivnějších.

## **Literatura**

P. Květon, M. Jelinek, D. Denglerová, D. Vobořil (2008). Software pro adaptivní testování: CAT v praxi. *Československá psychologie*, 52, 2, 145-154.

Květon, Petr - Klimusová, Helena (2002). Metodologické aspekty počítačové administrace psychodiagnostických metod. *Československá psychologie*, 46, 3, 251-264.

**Mgr. Jan Širůček**  
**IVDMR FSS MU**  
**Joštova 10**  
**602 00 Brno**  
**e-mail: sirucek@fss.muni.cz**