

Posudek disertační práce Mgr. Lindy Nepivodové

In Their Own Words and By the Test Results: A mixed methods study of the comparison of two modes of test administration

Ve své disertační práci představuje Mgr. Linda Nepivodová výsledky výzkumu zabývajících se srovnáním dvou typů testů, a to vyhotoveného studenty tradiční, ručně psanou formou a testů vyplněných na počítači (dále je budu nazývat papírové a počítačové testy). Jde tak o studii testové reliability, což je vhodné téma pro disertační práci. Autorka se snaží zjistit, zda je na katedře anglistiky, kde působí, využití počítačových testů pro testování angličtiny oprávněné. Logicky tak pro svůj výzkum využívá data získaná od vlastních studentů v rozmezí tří studijních ročníků. Výsledky kvantitativního výzkumu přitom trianguluje výzkumem kvalitativním, v němž zpracovává postoje studentů k oběma typům testů, a snaží se vytvořit návod pro úspěšné sestavení a zdárnou administraci počítačových testů.

Po stručném úvodu nejprve popisuje vývoj jazykového testování, dále se zabývá charakteristikami testových metod jako je validita, reliabilita, proveditelnost a efekt dopadu testů. Stručně též popisuje základní typologii testů dle různých kritérií. V další kapitole se věnuje problematice využití počítačů při jazykovém testování a uvádí přehled výzkumu srovnávajícího papírové testy s počítačovými. V další kapitole popisuje výzkumnou metodu, data a výzkumné otázky, posléze prezentuje a analyzuje výsledky kvantitativního i kvalitativního výzkumu a na jejich základě odpovídá na výzkumné otázky a zamýšlí se i nad praktickým využitím získaných poznatků. Struktura práce je tak standardní a splňuje kritéria pro disertační práce. Autorka zjišťuje, že médium testu nemá na výsledky významný vliv a že si studenti na počítačové testy rychle zvykají.

V úvodu autorka zmiňuje, že v disertaci navazuje na vlastní diplomovou práci, ale překvapivě nezmiňuje, k jakým výsledkům v ní dospěla a proč se rozhodla v tomto výzkumu pokračovat, či jaké mezery ve vlastním výzkumu se rozhodla vyplnit. Zde je na závalu jistá vágnost, kdy tvrdí, že v diplomové práci "odkryla řadu zajímavých poznatků a zároveň četné otázky" (str. 11), aniž by však jakkoliv vlastní výzkum a vyplývající otázky shrnula. Není tak vůbec patrné, proč její diplomová a disertační práce mají stejné téma a jak se odlišují. Úvod není dostatečně zasazen do kontextu existujícího výzkumu, a tak z něj není zřejmé, v čem a jakým způsobem mají být existující poznatky autorčiným vlastním výzkumem doplněny. Omezuje se na konstatování, že její výzkum "osvětlí výhody a nevýhody počítačových testů", což však považují za problematiku již velmi dobře probádanou a popsanou. Jakou výzkumnou mezeru zde autorka objevila? Co nového má její výzkum přinést? V úvodním odstavci tvrdí, že cílem práce je "give testing the attention it undoubtedly deserves" (p. 11). Je problematika testování skutečně nějak marginalizována? Osobně ji považují za hluboce probádanou oblast, o níž jsou stovky knih, tisíce studií, řada specializovaných periodik, konferencí, univerzitních programů a dokonce i celých institucí, jež se výzkumem na tomto poli zabývají.

Srovnává-li autorka výsledky dosažené v různých formách administrace jazykových testů, ústředním tématem disertace je reliabilita. Té je však věnováno necelé jeden a půl strany. Místo toho se autorka věnuje v teoretické části například historii jazykového testování. Jak to souvisí s tématem práce? Domnívám se, že historický kontext je v disertaci cenný, ale pouze pokud se týká tématu. A tím je zde reliabilita, což je samo o sobě téma s dlouhou a bohatou historií (viz např. Brennan, 2001), o níž zde není zmínka. Historická část není z mého pohledu napsána příliš kvalitně. Autorka zde sice pracuje s bohatou literaturou, ale poněkud ztrácí ze zřetele, že text disertace má být primárně jejím vlastním erudovaným a osvěcujícím výkladem, který se náležitě opírá o sekundární literaturu. V předloženém textu se však se sekundární literaturou pracuje primárně, a to do té míry, že se vlastní autorčin hlas poněkud vytrácí. V úvodní větě historického exkursu se dozvídáme, že "In this chapter, Spolsky's (1977) categorization of the history of language testing [...] is introduced and discussed." Proč zrovna kategorizace Spolského? A neexistují novější dějiny testování? Podobným způsobem pracuje autorka s literaturou v celé disertaci. Literaturu sice pečlivě sesbírala, ale její výběr se díky chybějícímu kritickému zhodnocení může jevit jako značně nahodilý. Takovou práci s literaturou mj. trpí i textová koherence. Ta je sice na první pohled skryta za kohezními prvky, ale jsou-li

tyto vyjmuty, často se v textu stává, že logická souvislost mezi jednotlivými pasážemi se hledá obtížně (viz např. str. 19 "[...] what Mastery Learning does is bring about the idea of streaming and thus greater variety" – Co to je "streaming"? A jaká logická souvislost implikovaná adverbium *thus* existuje mezi *streaming* a *greater variety*?). Obecně historická pasáž obsahuje mnoho detailů, ale málo vysvětlení a logických souvislostí (viz např. str. 21 – "One can clearly see the parallels to Assessment for learning." Osobně je tam na první pohled nevidím. V čem spočívají? Proč to není explicitně zmíněno?). Historickou pasáž neuzavírá shrnutí, a čtenář tak neví, jaké vývojové tendence autorka považuje za klíčové a vidí-li ve vývoji nějakou logiku či směřování. Pasáž tak trochu postrádá smysl, protože není vidět jasná vazba na vlastní výzkumné otázky. Rovněž vazba na další oddíl (znovu bez úvodní pasáže) není zřejmá. Autorka bez komentáře přechází k "Test Qualities", ale jaká je návaznost a souvislost?

V další pasáži se zabývá charakteristikami testových metod. V sekci o validitě píše: "The present author adopts here, as well as in her master thesis, the unitary concept of the notion of validity [...]." Nevysvětluje však, čím je její volba dána. V této sekci dále zmiňuje různé typy validit, ale ne vždy je definuje (např. 2.2.1.2 Face validity). Dalším drobným nešvarem a příkladem vágnosti je v sekci 2.2.1.4 práce se zdroji. "There are, of course, many limitations pointed out by, for example, Hughes (1989), Alderson et al. (1995), or Bachman (1990) and Weir (2005), as discussed in the author's master thesis (pp 32, 33)" To je mnoho textu, ale žádná informace. Doktorand by měl poskytnout syntézu, pokud něco zmiňuje, a ne jen odkázat na zdroje. Čtenář se zde nedozvídá nic. Podobně prázdné jsou i dvojí citace, mj. třeba na str. 23: "Weir (2005) acknowledges Anastasi's (1988) guidelines [...]"

V teoretické části, jak jsem již předeslal, mám největší výhrady k části věnované reliabilitě. Toto téma považuji za hlavní téma práce, ale zde je pouze velmi povšechně pojednáno na necelých 1 a půl stranách. Není zde základní definice pojmu, jakkoliv některé pasáže jsou dobře napsané. Reliabilita je předmětem výzkumné části, a to mj. i v podobě, která v disertaci není explicitně zkoumána, tedy mezianotátorské shodě. Zcela chybí zmínka o "generalizability theory", která s tématem práce úzce souvisí. Rovněž zmíněna není "inter-method reliability", která je přitom do hloubky zkoumána ve výzkumné části. Reliabilita měla být mj. rozebrána a vysvětlena i s ohledem na vývoj metod jejího výpočtu, protože jedině tak může být vytvořena souvislost s praktickou částí a objasněním autorčiny volby příslušných testů a metodických postupů. Více prostoru zabírají sekce o testové proveditelnosti a dopadu testů, které však ve výzkumu pojednávány nejsou. Namísto historie testování v disertaci mohl být přehled studií reliability, třeba i v historickém pohledu, jak jej prezentuje např. Brennan, R. L. (2001) (An essay on the history and future of reliability from the perspective of replications. *Journal of Educational Measurement*, 38(4), 295 – 317.). Tato studie však v bibliografii není.

Oddíl 2.2 dále bez snahy o logické propojení či shrnutí postupuje k oddílu 2.3 pokrývající typy testů, aniž by přitom byla zvažována otázka reliability u jednotlivých z nich. V oddíle 2.3.4 je zmíněna Chomského dichotomie kompetence vs. performance, která by snad měla opodstatnění, kdyby byla zmíněna v kontextu testové validity, ale oddíl je zde zařazen bez této reference a bez explicitního vyjádření, jak souvisí s předcházejícím textem či vlastním výzkumem.

Obecnými výhradami k celé 2. kapitole je přílišná závislost na použité literatuře – kde je autorčin vlastní výklad, vlastní pohled na věc, erudice či kritické zhodnocení? Kde je logický vývoj celku i jednotlivých částí? Text místy připomíná spíše koláž náhodně vybraných útržků z literatury, ke kterým se autorka nestaví vlastním názorem, kategorizací či vytknutím logických souvislostí. Vazba mezi jednotlivými částmi (kterých je celá řada) druhé kapitoly a zároveň mezi nimi a důvody, proč byly zařazeny do disertace, je tak kvůli značné absenci autorčina hlasu velmi nejasná.

Třetí kapitola se zabývá využitím počítačů v jazykovém testování. Zde vůbec nerozumím zařazení oddílu 3.1 Digital literacy (digitální gramotnost), kterou na sklonku 2. desetiletí 21. století považuji v kontextu vysokoškolských studentů za zcela překonanou. K čemu je Tabulka č. 1, jež zabírá takřka půl strany, ale s jediným z konceptů, které zahrnuje, se v práci dále nepracuje? Mnoha termínům v ní uvedeným naprosto nerozumím (např. "remix literacy"), termíny nejsou vysvětleny, ani dále využívány. Autorka nás nicméně

(str. 39) vyzývá, abychom se na ně dobře podívali, obzvláště pak na ty, jež jsou v kontextu práce relevantní, ale vzápětí říká, že na to zde není prostor. Čtenář tak neví, na co se má vlastně dívat a proč, o relevanci v kontextu práce ani nemluvě. Tato část je obecně velice nedbale napsána. Za vše cituji ze str. 40: "With reference to a number of studies dealing with ICT access, usage, skills, knowledge and interest, Bennet and Maton (2011) dismantle Prensky's claims, for example, Otto et al., 2005, Livingstone and Helsper, 2007, Selwyn, 2008, Jenkins, 2009, Kennedy et al., 2009, Maton and Bennet, 2010, and Jones et al., 2010." Celkem čtyři řádky textu, které naprosto nedávají smysl a nevhodným způsobem "pracují" s literaturou (cf. též nesystematická vrocení). Další pasáže oddílu 3.1 vůbec nejsou relevantní obsahu disertace. Paradoxně stať o digitální gramotnosti zabírá v disertaci téměř trojnásobný prostor, než vlastní zkoumané téma reliability!

Dalších 18 stran textu je věnováno popisu vlastností a typů testových úloh. To na první pohled dává smysl, ovšem do té doby, než zjistíme, že autorka v disertaci nepředkládá testy (ani ukázky z nich), které pro výzkum použila. Můžeme se tak jen dohadovat, nakolik je těchto 18 stran skutečně relevantních. K oddílu 3.3 diskutujícímu výhody a nevýhody testování počítačem nemám vážnější výhrady, pozastavuji se pouze nad tím, že většina literatury, s níž se v části pracuje, je deset a více let starší. Tento obor se však dynamicky vyvíjí a v posledním desetiletí prošel výraznými změnami. V tomto ohledu mi chybí novější literatura a příklady důležitých počítačových testů (alespoň těch "high-stake"). Myslím si, že ve světě ICT není dnes možné pracovat se zdroji staršími než 20 či možná dokonce 10 let. Rozhodně existují novější a rozhodně se neustále a rapidně proměňují jak technické tak softwarové možnosti.

Relativně nejlépe zpracovanou sekcí teoretické části disertace je oddíl 3.5 zabývající se porovnáváním papírových a počítačových testů. To je přitom předmětem vlastního výzkumu a mělo by mít vedle reliability v disertaci čelní místo. Autorka podává zprávu o výsledcích existujícího výzkumu, což je výborné, ale nezamýšlí se nad metodologickými aspekty prezentovaného výzkumu. Jde o reliabilní studie? Jsou mezi sebou navzájem srovnatelné? V čem je metodologický posun autorčina vlastního výzkumu, když se rozhoduje provádět stejný výzkum? Jsou její metody odlišné a proč? Je to právě v této oblasti, kde měla autorka hledat mezery ve výzkumu a pokusit se zlepšit metodologii. Jinak se zkoumá něco, co již bylo vyzkoumáno.

Poznámky k praktické části

Výzkumná část disertace má dvě složky, a to kvantitativní a kvalitativní. Snaží se tak o zkoumání problematiky metodou smíšeného výzkumu, což je přínosné. Kvantitativní výzkum je proveden na srovnání výsledků papírových a počítačových testů ve třech po sobě následujících ročnicích studentů prvního ročníku anglistiky. Autorka výzkum označuje za longitudinální, což neodpovídá realitě. V longitudinálním výzkumu se zkoumá vývoj u **stejných** subjektů, kdy hlavní proměnnou je **čas**. V představeném výzkumu jde pouze o opakování stejných (resp. postupně modifikovaných) testů u **jiných** studentů ve třech letech po sobě, a čas jako proměnná zde tedy nehraje žádnou roli.

V úvodní pasáži popisu metody formuluje výzkumnou otázku jako zjištění, zda je využití počítačových testů v daném kontextu obhajitelné ("justified"), aniž by přitom sdělila, v jakém ohledu by obhajitelné mělo být. Problém je formulován jako výzkumná otázka nikoliv hypotéza. Jaká byla nulová hypotéza? Zda výsledky nejsou srovnatelné, či že naopak jsou? V další výzkumné podotázce se táže, zda má na výsledky vliv pohlaví testovaných. Z čeho by měla pramenit hypotéza, že by zde měl být významný rozdíl? Ze stereotypické představy, že muži jsou v užívání počítačů zběhlejší? Taková otázka se mi jeví jako přinejmenším podružná.

Datový soubor je dobře popsán, chybí mi jen informace o úrovni studentů. Bohužel jsem nemohl provést zhodnocení testů samotných, protože testy nejsou součástí přílohy a jsou jen stručně popsány. Posuzovatel tak nemůže hodnotit parametry testů a udělat si představu o tom, zda je tento výzkumný nástroj dostatečně reliabilní. Z uvedeného vyplývá jen to, že testy jsou primárně lexikálně-gramatické, což mi přijde zvláštní v kontextu teoretické části, kde autorka popisuje i problematiku komunikativní kompetence. Rovněž by bylo bezesporu zajímavé, jakým způsobem ovlivní výsledky zařazení úloh na čtení s porozuměním, kdy

bych si dokázal představit, že prezentace textu ke čtení v tištěné podobě může vést v závislosti na grafickém ztvárnění a délce textů k lepším výsledkům než v počítačové verzi. Rovněž není zvaženo testování poslechu a psaní, takže výsledek výzkumu hovoří jen o velmi úzce zaměřeném didaktickém testu klasického typu. Bez konkrétních ukázek není možné hodnotit ani testovou validitu – co měly testy vlastně testovat? Znalost gramatiky a slovní zásoby? Nebo komunikativní dovednosti či stupeň osvojení různých řečových dovedností? Zkoumané testy jsou dle mého názoru málo komplexní, a výsledky výzkumu tak jen v malé míře zobecnitelné.

K vlastnímu srovnání papírových a počítačových testů mám několik výhrad:

- 1) Testy nebyly ve všech ohledech identické. V počítačové verzi se objevuje časovač, který dle všeho nebyl přítomen při papírových testech.
- 2) Použité počítačové testy ukazovaly studentům jejich průběžnou úspěšnost. To se bezesporu stává proměnnou, která však není zvažována. Průběžné výsledky mohou ovlivňovat stres účastníků. Navíc jde o další odlišnost od papírové verze. Při sestavování testů mělo být toto zvaženo a průběžné výsledky neměly být zobrazovány, aby si byly testy podobné v co největší míře. Vyloučením veškerých jiných možných proměnných by bylo dosaženo skutečné srovnatelnosti testů, jež by se bývaly lišily pouze v použitém médiu.
- 3) Administrace papírových testů byla dle mého názoru nevhodná a vedla k nižší reliabilitě. Dle vlastních vyjádření účastníci neměli informace o čase, navíc byli rušeni povídáním examinátorů (viz pozn. na str. 130). U experimentu je nutné za účelem maximalizace reliability zajistit co nejideálnější testové prostředí, což mělo být snadné a automatické. Uvedené podmínky se tak staly dalšími proměnnými, jejichž přítomnost byla zbytečná, ale přitom mohla výrazně ovlivnit výsledky celého výzkumu.
- 4) Je chvályhodné, že autorka zvažuje statistické parametry distribuce dat (str. 89). Ty by však měly vést k volbě odpovídajících statistických testů, nikoliv ke konstatování, že statistická srovnávání nejsou možná. Proč nebyly použity neparametrické testy (viz str. 90, srovnání výsledků dle pohlaví)?
- 5) Proč nebyly vyloučeny mezní hodnoty (outliers), když byly dosaženy zahraničními studenty s nižší úrovní AJ, kteří nesplňovali kritéria zkoumané populace?
- 6) Proč nebyly papírové a počítačové testy a jejich verze obsahově absolutně shodné? (Týká se prvního a druhého roku experimentu.) Jakékoliv srovnávání je tak téměř bezvýznamné.
- 7) Chybí informace o mezihodnotitelské shodě. Byly testy vyhodnoceny více examinátory? Jaká byla shoda? Vyhodnocoval-li test pouze jeden examinátor, je významně snížena reliabilita, a do výzkumu tak vstupuje další proměnná. To se týká papírových testů, i otevřených otázek v testech počítačových.
- 8) Autorka zveřejnila část dat na diskovém úložišti Google. Proč nevyužila standardní postup vložení CD či DVD s daty do výtisku disertace?

Oceňuji, že autorka mezi jednotlivými roky zlepšovala testy a poučila se z nedostatků pilotní studie. Je škoda, že se to netýkalo i některých z výše uvedených parametrů. Nicméně metodická vylepšení administrace testů jsou nanejvýš chvályhodná. Rovněž oceňuji statistické zpracování výsledků a jejich prezentaci, jakkoliv bych přivítal vedle prezentace výsledků i zřetelnější závěry (viz např. výsledky na str. 105). Na str. 114 autorka zmiňuje, že omezením studie č. 2 bylo, že "the research design was general", čemuž nerozumím.

V kvalitativní části výzkumu autorka vyhodnocuje reakce studentů na jednotlivé varianty testů. Tyto reakce získala na základě jednoduchého dotazníku (viz příloha č. 4C). K němu mám tyto výhrady: Proč byly dotazníky k vyplnění v angličtině? Mohlo to nějak snížit reliabilitu? Proč jsou studenti vyzýváni k tomu, aby uvedli pouze jeden aspekt, který se jim na daných testech "líbil" či "nelíbil"? Co když studenti chtěli uvést aspektů více? Rovněž nevím, proč autorka nezačlenila otázky týkající se např. zhodnocení vlastní úrovně počítačové gramotnosti, zkušenosti s jinými počítačovými testy, postoj k IT obecně atp. Takové údaje by jistě bylo možné korelovat s odpověďmi v dotazníku i s výsledky samotných testů. Autorka sice uvádí, že pro tento účel sestavila dotazník, ale výsledky nejsou v disertaci prezentovány a autorka se omezuje pouze na sdělení, že korelují s výsledky z kvalitativního výzkumu. Vzhledem k tomu, že dotazník nebyl zařazen do přílohy, není vůbec možné si o této části výzkumu udělat obrázek.

Zpracování údajů z kvalitativních dotazníků proběhlo kategorizací na základě otevřeného kódování. Výsledky jsou podány přehledně. Nemám k nim výhrad, jakkoliv většina sdělení získaných od studentů je spíše banálních (např. vadí mi hluk psaní na klávesnici atp.).

Na závěr autorka shrnuje odpovědi na jednotlivé výzkumné otázky. Zde chybí odpověď na otázku č. 3 na str. 76 týkající se vztahu mezi skóre a typem úlohy. Právě srovnání reliability jednotlivých typů úloh realizovaných na papíře a na počítači by podle mého názoru mělo mnohem větší smysl než srovnávání podle pohlaví.

Shrnutí připomínek k teoretické části:

- chybí mi vlastní autorčin výklad;
- chybí mi identifikace mezer ve výzkumu;
- chybí mi shrnutí východisek z autorčiny vlastní diplomové práce a shrnutí, jaké mezery v ní objevila a proč pokračuje ve stejném tématu;
- nízká míra logického propojení jednotlivých částí;
- nedostatečná práce s literaturou a absence kritického přístupu k ní;
- nevhodně zvolená témata (možná dáno tím, že autorka nevytváří logická propojení, aby zdůvodnila, proč některá témata zařazuje);
- marginalizace hlavního tématu disertace (reliability) (zdá se mi, že si autorka sama neuvědomila, co je skutečným hlavním tématem práce);
- častá vágnost, nedostatečná vysvětlování, chybějící definice.

Shrnutí připomínek k praktické části:

- mylné označení designu studie za longitudinální;
- nezveřejnění porovnávaných testů;
- malá komplexnost testů a z toho vyplývající nízká zobecnitelnost výsledků;
- neidentičnost srovnávaných testů a rozhraní;
- špatná administrace testů;
- možné matoucí proměnné pramenící z nestejnosti podmínek a administrace testů;
- nepopsaný postup vyhodnocování testů (např. bylo hodnotitelů více?).

Závěr

Jakkoliv jsem si vědom toho, že začínající posuzovatelé disertací, mezi něž patřím, mají tendenci být příliš přísní, domnívám se, že mnou zmíněné nedostatky bohužel převažují pozitiva práce a že by podstatné části práce měly být přepracovány. V důsledku toho navrhuji klasifikaci **neprospěla**.

PhDr. Tomáš Gráf, Ph.D.
V Praze 2.6.2017

Dílčí připomínky

Mnoho sekcí je označeno číslem ve decimální systému, ale samotná pasáž tam chybí. Viz např. oddíl 2 a bezprostředně navazující oddíl 2.1. Zde před oddílem 2.1 mohla a dle mého mělo být vysvětleno, jaké jsou ony "fundamental concepts" a proč je autorka za důležité považuje. Mohla tak ku příkladu vysvětlit, jak je historie relevantní tématu disertace. Osobně tam spojitost těžko hledám, ale autorka měla šanci vlastní vidění souvislostí nastínit. Stejně tak oddíly 2.2, 2.3. a 3. aj.

- vágní tvrzení, např. str. 18, 19, 26 ("second administration of the test proves to be rather problematic (Hughes, 1989)" – nedozvídáme se ale v jakém ohledu).
- str. 18 – neexistující sloveso "contradistincts"
- str. 18 – "language test is a test not a tea party" – co přesně tím autor mínil?
- "This is a thought-provoking statement" – co je thought-provoking na tvrzení, že "communicative language testing [...] belongs to history"?
- str. 21–23 – změna typu zarovnání textu
- "there are, of course, many limitations [...]" – proč nějak neshrnout, jaká to jsou?!
- str. 33 – "it needs to be emphasized that they serve different functions as well as purposes (Bachman, 1990)" - má-li to být zdůrazněno, proč to sama autorka bez diskuse takto povšechně přechází?
- v celé disertaci chaotické užívání velkým písmen v názvech oddílů;
- str. 49 – je nutná Tab. 3? Nestačilo by dát do jedné věty?
- str. 52 – autorčina "comprehensive table of some commonly used task formats" je málo "comprehensive" – srov. např. Purpura (2004), který nabízí mnohem detailnější a úplnější seznam;
- občasné užívání hovorových výrazů (quite a few, a bit more freedom, tiniest detail, pretty straightforward, slightly better, nicely complement, quite a lot, get hold of, incredibly difficult)
- str. 56–58 – čím je Effective Item Checklist (který je téměř celý přejatý z Brown (2005) relevantní pro autorčin výzkum?
- str. 58, 64 – nevhodné dělení tabulky
- str. 66 – v citaci "relatively a few performance thresholds, it is somewhat more prone to overexposure (Davey, 2011, p.11)" – autorka doslova cituje, aniž by vysvětlovala – co je v tomto kontextu overexposure? Jak to má čtenář vědět?
- str. 80 – "Dörnyei (2008)" – autorem je Dörnyei
- str. 89 – amount se pojí s nepočítatelnými substantivy;
- str. 90 – předposlední odstavec na straně téměř doslova duplikuje první;
- str. 113 – "the effect is roughly equal to 1 and a half additional item achieved correct" – negramatické;
- str. 155 – Alderson et al. – chybný formát položky v bibliografii