

Bol Likert ordinalista alebo intervalista? Chyby pri tvorbe a vyhodnocovaní Likertových škál

Milan Kubiatico

Katedra pedagogických štúdií, Fakulta humanitných vied ŽU, Žilina

Anotácia: *Bol Likert ordinalista alebo intervalista? Chyby pri tvorbe a vyhodnocovaní Likertových škál.* Likertove škály patria medzi často používané výskumné techniky vo výskumoch pedagogického zamerania. Preto je štúdia zameraná na prezentovanie použitia a vyhodnocovania škál Likertovho typu. Prvá časť textu predstavuje základné informácie o tomto type škál, o ich tvorbe a používaní. Druhá a tretia časť textu sú kritické, najskôr sú rozobrané prístupy k ich chápaniu, podľa toho, za aký typ premennej ich jednotliví autori považujú, sa vytvorili dve skupiny, tzv. ordinalisti a intervalisti. Uvedené sú dôvody, ktorými jednotliví autori operujú, prečo práve ten ich prístup je správny. V samotnom texte nastáva prikláňanie sa k prístupu tzv. intervalistov, čo prechádza do tretej kapitoly, ktorá ponúka čitateľom desať najčastejšie sa vyskytujúcich chýb pri tvorbe, používaní Likertových škál a analýze dát, ktoré sú škálami produkované. Každá mylná predstava obsahuje krátke zdôvodnenie, ktoré sa jej snaží protirečiť. Záverečná kapitola prináša zhrnutie prezentovaného textu. PEDAGOGIKA.SK, 2016, ročník 7, č. 3: 177-190

Kľúčové slová: intervalisti, Likertova škála, ordinalisti, položka Likertovho typu

Was Likert Ordinalist or Intervalist? The Mistakes by the Creation and Evaluation of Likert Scales. Likert scales belong to often used research techniques in the researches focused on education. Because of this fact, the study is focused on the presentation of using and evaluation of Likert scales. The first part of the text provides the basic information about this type of scales, their creation and using. The second and third part are crucial, at first the approaches are characterized according to their understanding, then according to the type of variables that are considered by specific authors there were created two groups ordinalists and intervalists. There are described reasons which are presented by specific authors and these reasons explain why the selected approach is valid. The text opts for the view of so called intervalists that is described in the third chapter, in which there are listed ten most widespread misconceptions occurring in the creating and using Likert scales and also in the data analysis produced by the scales. Each misconception contains brief justification, whose role is to contradict to the misconception. The conclusion summarizes the presenting text. PEDAGOGIKA.SK, 2016, Vol. 7 (No. 3: 177-190)

Key words: intervalists, Likert scale, ordinalists, Likert type item

Úvod

Veľké množstvo kvantitatívne orientovaných výskumných šetrení v pedagogickom výskume používa výskumné nástroje obsahujúce položky Likertovho typu. Častokrát nastáva problém už pri samotnom používaní, keď niekedy výskumní pracovníci nahradzujú názvy bodov na škále vlastnými vytvorenými názvami, čo sa ukazuje ako jedna z mnohých chýb. Takisto vyhodnocovanie dát produkovaných položkami Likertovho typu sa častokrát odchyľuje od určitých zákonitostí, ktoré sú pre tento typ položiek zaväzujúce. Na základe toho môžu autori vyprodukovať informácie, ktoré sa môžu líšiť od pravdivých informácií, ktoré by dosiahli správnou analýzou dát. Na tento problém je upozorňované vo viacerých štúdiách novšieho, či staršieho dáta (napr. Strasser, 2007). O tom, ako treba chápať Likertove škály, ako pristupovať k ich vyhodnocovaniu, je venovaný tento príspevok.

1 Čo je vlastne Likertova škála?

V úvodnej kapitole budú stručne uvedené informácie o Likertovej škále. Táto škála nesie niekedy aj označenie metóda sumovaných intervalov, a ako už znie jej názov, bola navrhnutá Likertom v roku 1932. Uvedený typ škály používa položky vo forme tvrdení, či výrokov, pričom respondent vyjadruje svoj súhlas, resp. nesúhlas s nimi. Prvý krok pri jej tvorbe je príprava určitého počtu tvrdení, ktoré sa týkajú toho, aké informácie má škála zistiť. Pri zisťovaní postojov ide o výroky, ktoré sa týkajú objektu postoja tak, aby súhlas, resp. nesúhlas s výrokom identifikoval valenciu postojov (bol prejavom pozitívneho, resp. negatívneho postoja k danému objektu). Pripravený súbor položiek je posudzovaný na typickej (Likertovej) 5-bodovej odpovedajúcej škále.

Používajú sa 2 verzie:

- so silným súhlasom a nesúhlasom (rozhodne súhlasím, súhlasím, neviem, nesúhlasím, rozhodne nesúhlasím – namiesto rozhodne je možné použiť napr. silno);
- s oslabeným súhlasom a nesúhlasom (súhlasím, čiastočne súhlasím, neviem, čiastočne nesúhlasím, nesúhlasím – namiesto čiastočne je možné použiť napr. mierne).

Podľa zvoleného postupu potom vypočítame výsledné skóre každého jednotlivca ako súčet alebo priemer škálových hodnôt odpovedí na jednotlivé položky. Pozornosť je potrebné venovať polarizácii položiek.

Základný problém Likertových škál spočíva v tom, že nediferencujú spoľahlivo v strednom pásme výsledné skóre. Skóre nachádzajúce sa v teoretickom strede môže byť výsledkom konštantnej voľby odpovede „neviem“, ale takisto aj môže byť výsledkom približne rovnakého výskytu odpovedí v prospech a proti objektu postoja (hodnoty „1“ a „5“). Preto sa tento typ škál viac hodí na zisťovanie rozdielov medzi skupinami a na výskumné

ciele vôbec, než na individuálnu diagnostiku jednotlivca (Krosnick, Judd & Wittenbrink, 2005).

Pri zostavovaní škálovaných položiek musí zostavovateľ určiť počet bodov na škále tak, aby bol rovnaký pre všetky položky. Počet bodov na škále sa mení v závislosti od typu výskumu, ale aj samotného výskumníka. Medzi najpoužívanejšie patria 5-bodové škály Likertovho typu, menej sú používané 7-bodové a 11-bodové škály.

Keď je respondent konfrontovaný so škálovanou položkou (napr. Mám rád prírodopis), tak najskôr si musí uvedomiť vlastný postoj k danému výroku a následne nájsť na škále hodnotu, ktorá najviac odpovedá jeho postojovi k tomuto výroku. Na to, aby respondent takto odpovedal, je nutné, aby boli splnené určité podmienky. Body na škále by mali byť ordinálne, ich počet by mal byť taký, aby pokrýval celé merateľné kontinuum. Každý respondent musí rozumieť položke a vedieť jej priradiť určitý bod na škále. Počet bodov na škále závisí na stupni rozvinutia postoja respondentov k danému javu. Ak má napríklad jedinec extrémne negatívny alebo pozitívny postoj k určitému javu, tak je pre neho postačujúca dichotomická škála s dvoma možnosťami (súhlasím – nesúhlasím).

Pre respondenta s neutrálnym postojom je dichotomická škála nevyhovujúca, neponúka mu možnosť pre neutrálnu stanovisko. Taktiež škála s tromi možnosťami (súhlasím – neviem – nesúhlasím) nemusí byť vyhovujúca pre respondentov s jemnejším vnímaním, ktorí nemajú neutrálnu stanovisko, ale nezastávajú ani extrémny postoj. Ako najvhodnejšie sa javia 5-bodové škály (úplne súhlasím – súhlasím – neviem – nesúhlasím – úplne nesúhlasím). Možné je použiť aj 7-bodovú škálu u respondentov, ktorí majú veľmi citlivé vnímanie postojov. Táto škála sa uplatňuje skôr pri meraní postojov u marginálnych skupín, ako sú napríklad nadpriemerne inteligentní jedinci. Samozrejme, cennejšie výsledky sú získané z viacbodovej škály. Ale na jej uplatnenie musia byť splnené dve podmienky. Prvou podmienkou je, že respondent dokáže označiť na škále ktorýkoľvek bod. Druhou podmienkou je, že si je plne vedomý jeho významu. Pri viac ako 5-bodových škálach dochádza k ignorovaniu krajných možností, a takto získané dáta sa stávajú skreslenými a pre výskumníka je ich význam malý.

Snahou výskumníkov zaoberajúcich sa postojmi je dôležité, aby ich výskumný nástroj bol validný a reliabilný. Aby boli tieto podmienky splnené, tak pre respondenta musí byť škálovaná položka zrozumiteľná a musí jej vedieť priradiť bodovú hodnotu. Preto musí byť každý bod na škále špecifikovaný. Z uvedených informácií vyplýva, že najjasnejšie sú pre respondentov položky s dvoma možnosťami. Preto zrozumiteľnosť položky klesá s narastajúcim počtom bodov na škále. Niekedy sú používané ešte aj viac ako 7-bodové škály, kde dochádza k problémom pomenovať jednotlivé body na škále a často

nastáva situácia, keď výskumník pomenuje len krajné body (úplne súhlasím – úplne nesúhlasím) a medzi tým nechá len číselnú stupnicu. V takom prípade býva pre respondenta veľmi náročné určiť postoj ku skúmanému javu. Pri určovaní počtu bodov na škále by si mal autor výskumného nástroja uvedomiť i rozsah kognitívnych schopností respondentov. Nárast kognitívnych schopností sa predpokladá so stúpajúcim vekom respondentov. Pre žiakov základných škôl je zbytočné dávať viac ako 5-bodové položky (Krosnick, Judd & Wittenbrink, 2005). Existuje aj niekoľko odborných prác, ktoré poukazujú na nevhodnosť 5-bodových škál a za vhodné prezentujú používanie 11-bodových škál (napríklad Leung, 2011). Ako argument uvedení autori uvádzajú, že len pri viacbodových škálach je možné dosiahnuť normálne rozdelenie získaných dát. Autori však výskum realizovali na študentoch vysokých škôl, či na dospeljej populácii. Problémom je tiež, že autori neuvádzajú priemerné skóre, aké bolo dosiahnuté pri jednotlivých položkách, a ani príklad, ktorým by poukazovali na označovanie jednotlivých bodov na škále. Či naozaj respondenti označovali všetky body alebo niektoré boli ignorované, nebolo uvedené. Avšak už Hirner (1978, s. 123) sa prikláňal k 5-bodovým škálam, pretože respondenti nemusia dostatočne citlivo diferencovať pri nižšom počte bodov a pri vyššom počte bodov nevyužívajú dostatočne škály. V tomto zmysle navrhovali 5-bodové škály ako najvhodnejšie aj vyššie zmienení autori (Krosnick, Judd & Wittenbrink, 2005).

Medzi výskumníkmi, ktorí používajú škálované položky, sa často vedú diskusie o tom, či je vhodné používať, resp. nepoužívať prostredný neutrálny bod. Neutrálny bod slúži v prípadoch, keď respondent nemá vytvorený postoj ku skúmanému javu (Krosnick, Judd & Wittenbrink, 2005). Rod (2012) vo svojej prehľadovej štúdii uvádza, že mnohí výskumníci sa prikláňajú k párnemu počtu bodov na škále, čiže vynechávajú neutrálny bod. Podľa nich to vedie k menšiemu skresleniu výsledkov. Rod (2012) ďalej uvádza, že väčšina autorov sa však prikláňa k nepárnemu počtu bodov na škále, pretože absencia neutrálneho bodu znižuje vierohodnosť nameraných výsledkov. Niektorí autori sa vyjadrujú o vzájomnom vzťahu reliability a počtu bodov na škále, pričom uvádzajú, že nepárny počet bodov na škále spôsobuje vyššiu spoľahlivosť výskumného nástroja v porovnaní s párnym počtom bodov na škále (napr. Cicchetti, Showalter & Tyrer, 1985; Fabiola et al., 2012).

2 Ako vyhodnocovať Likertove škály?

Ak nahliadneme do rôznych metodologických odborných knižných publikácií, z ktorých množstvo výskumníkov, ale aj študentov čerpá, je tam možné sa stretnúť poväčšine prípadov len so strohou informáciou, že meranie na škálach je brané ako ordinálna premenná (napríklad Ferjenčík, 2000, s. 222; Walker, 2013, s. 136). Prípadne druhá séria publikácií, kde sú zas škály brané ako

intervalová premenná (napríklad Pelikán, 1998, s. 109; Kerlinger, 1972, s. 472). Podobný prístup majú aj štúdie zo slovenského prostredia, či už staršieho dáta (Hirner, 1978, s. 121) alebo novšieho (Gavora, 2008, s. 114). V ďalšej časti kapitoly bude pokus o bližšie predstavenie oboch prístupov k nazeraniu na Likertove škály.

Pri posudzovaní položiek Likertovho typu dochádza vo výskumnom svete k rozdielnemu nazeraniu na ne. Napríklad Jamieson (2004) sa vo svojej štúdií prikláňa k názoru, že položky Likertovho typu majú výsostne charakter ordinálnych premenných. Autor vo svojej štúdií pojednáva o aplikácii uvedeného typu položiek do oblasti medicíny. Vo svojej publikovanej štúdií to však rieši len na teoretickej úrovni. Jamieson pravdepodobne vychádza z práce Stevensa (1946), kde autor uvažuje o Likertových škálach ako o ordinálnych premenných. Podľa neho jedna položka Likertovho typu predstavuje atóm a spojenie viacerých položiek do jedného celku (škály) je pri analógii s biológiou molekulou. Ale podľa Stevensa nemôže byť škála braná ako intervalová premenná, pretože atómy v molekulách sú navzájom na sebe závislé, preto škála musí byť braná ako ordinálna premenná. Podobne na teoretickej úrovni rozoberajú nahliadanie na položky Likertovho typu ako na ordinálne Göb, McCollin & Ramalhoto (2007). Podobne ako Jamieson aj tu sa je možné stretnúť s argumentáciou, že medzi jednotlivými bodmi na škále nie je rovnaká vzdialenosť, a preto nie je možné považovať ich za intervalové. Na základe toho autori z oboch štúdií odporúčajú používať len percentuálne zhodnotenie označenia toho-ktorého bodu na škále. A následne odporúčajú používať len neparametrické metódy spracovania dát, ako sú Mann-Whitneyho test, či Kruskal-Wallisov test. Na praktickej úrovni, kde autori vychádzajú práve z práce Jamiesona, je možné sa stretnúť v oblasti klinickej a zdravotnej psychológie, kde autori používajú sadu dotazníkov obsahujúcich položky Likertovho typu, ale v celej metodologickej časti práce je možné sa stretnúť s odvolávaním sa na to, že sa ne prihliada ako na ordinálne premenné (Fonseca-Pedrero, et al., 2013). Autori však používajú štatistické postupy, ako sú reliabilita prostredníctvom Cronbachovej alfa, či konfirmačná faktorová analýza. Podobný názor zdieľajú aj Zumbo, Gadermann & Zeisser (2007), ktorí odporúčajú považovať položky Likertovho typu za ordinálne. Okrem iného vo svojej štúdií, že aj keď sú položky Likertovho typu považované za ordinálne premenné, je možné používať na zisťovanie reliability koeficient Cronbachova alfa. Podobne naznačujú vo svojej práci Forero, Olivares & Pujol (2009), že by sa malo na tento typ položiek pozeráť ako na poradové, autori sa o tomto fakte vo svojej štúdií len zmieňujú. Väčšina obsahu prezentovanej štúdie sa však týka aplikácie faktorovej analýzy na položky Likertovho typu. Obdobne sa môžeme stretnúť s pozeraním sa na položky Likertovho typu ako na poradové v oblasti finančnictva (Gardner & Martin, 2007). Vo svojej štúdií rozoberajú

okrem iného dostatočný počet položiek vo výskumnom nástroji. Wu (2007) sa pozerá na problematiku použitia a vyhodnocovania položiek Likertovho typu z čisto matematického hľadiska. Tu je možné vidieť veľmi zaujímavý náhľad na vyhodnocovanie tohto typu položiek. Autor sa prikláňa k vyhodnocovaniu len prostredníctvom metód typických pre poradové premenné. Zdôvodňuje, aj matematicky znázorňuje, že pri dátach získaných z tohto typu položiek je ťažké získať normálne rozloženie, preto je použitie parametrických metód nemožné. Podobne ako v predošlých štúdiách, ani tu však autor nemá zábrany používať viacrozmerné štatistické metódy, ako je napríklad faktorová analýza.

Na druhej strane stojí, dá sa povedať, väčšia skupina autorov, ktorí sa prikláňajú k realite, že na Likertove škály je nutné sa pozerieť ako na intervalové premenné. Zaujímavovo vyznieva fakt, že takmer všetci autori z tejto druhej skupiny vo svojich prácach poukazujú na to, kde podľa nich urobili „chybu“ autori zo skupiny ordinálnych premenných. Podľa nich prvá skupina hodnotí len jednotlivé položky Likertovho typu, vtedy o nich možno hovoriť ako o ordinálnych dátach. Ale podľa autorov z druhej skupiny by sa položky Likertovho typu nemali hodnotiť jednotlivo, ale ako skupina položiek, ktorá predstavuje Likertovu škálu (Carifio, 1976; Carifio, 1978; Carifio & Perla, 2007; Vickers, 1999). Podporu „intervalovosti“ Likertových škál je možné nájsť vo viacerých odborných prácach, ako príklad je možné uviesť konštatovanie autorov Allen & Seaman (1997), že to je základná vlastnosť Likertových škál pôsobiť ako intervalové premenné, to už bola podstata pri ich vytváraní prvotným autorom (Likert). Nepredpokladá sa, že by jednotlivé body na škále mali byť ako nálepky. Brown (2011) pridáva porovnanie s vyhodnocovaním testov, keď sa vyhodnocuje jedna úloha, pozerá sa na to ako na nominálnu premennú (správna odpoveď – nesprávna odpoveď), ale pri vyhodnocovaní celého testu sa už s odpoveďami za jednotlivé otázky pracuje ako s intervalovými premennými, vypočítava sa priemer, prípadne smerodajná odchýlka, pri porovnávaní úspešnosti medzi demografickými premennými prichádza na rad induktívna štatistika. Podľa autorov by sa podobne malo hľadiť aj na Likertove škály. Ak sa vyhodnocuje len jedna položka, malo by sa k nej pristupovať ako k ordinálnej premennej, ale keď sa berie do úvahy Likertova škála, malo by sa na ňu hľadiť ako na intervalovú premennú. Pell (2005) potvrdzuje vo svojej práci vyššie zmienené fakty, že položky Likertovho typu by sa nemali vyhodnocovať jednotlivo, ale ako Likertova škála, v ktorej bude zastúpených viac položiek. Rovnaké interpretácie je možné nájsť aj v práci zo slovenského prostredia (Hirner, 1978), kde autor hovoril o určení postoja, či názoru respondentov, ak sa do úvahy berie celkové skóre zo všetkých položiek, nielen z jednej. Autor vo svojej práci taktiež uprednostňoval použitie parametrických štatistických metód pri vyhodnocovaní dát získaných prostredníctvom Likertových škál. Clason & Dormody (1994) dodávajú

k tomuto faktu okrem vyššie zmienených skutočností, že vo výskumných prácach by sa malo používať buď len vyhodnotenie položiek Likertovho typu, alebo len Likertových škál. Gu, Wen & Wu (1995) poukazovali vo svojej práci na použitie položiek Likertovho typu a taktiež na ich vyhodnotenie. Pri vyhodnotení sa jednoznačne opierali o indukčnú štatistiku typickú pre dáta intervalovej povahy. Čiže používali testy, ako sú analýza rozptylu, či multivariátna analýza rozptylu. Autori však upozorňovali na jeden problém, ktorý môže pri použití tohto typu položiek nastať, a to je zámena klasického pomenovania bodov na škále (úplne nesúhlasím – ... – úplne súhlasím) za názvy bodov (nikdy – občas – niekedy – často – stále), prípadne ich podobné znenie. Podľa nich práve v tejto situácii môže nastať problém, pretože respondenti v tejto situácii nevedia presne identifikovať, aký je rozdiel medzi „občas“ a „niekedy“. Preto by bolo vhodnejšie používať len body, ktoré navrhol Likert (úplne nesúhlasím – ... – úplne súhlasím). Za malú zmienku stojí výskumná štúdia od Murrayho (2013), ktorý zistil rovnaký výsledok pri analýze Likertových škál tak parametrickou, ako aj neparametrickou metódou. Do úvahy treba brať i fakt, že výskum bol robený na dátach ekonomického typu. Veľmi interesantne vyznievajú informácie v štúdiu od Normana (2010), ktorý okrem iného analyzuje aj vyhodnocovanie dát z položiek Likertovho typu. Za zmienku stojí, že táto štúdia leží kdesi na priesečníku dvoch vyššie zmieňovaných pohľadov na vyhodnocovanie dát z uvedeného typu položiek. Autor uvádza, že analýzou mnohých výskumných prác zistil, že položky Likertovho typu sa môžu vyhodnocovať tak parametrickými, ako aj neparametrickými metódami, uvádza aj ďalšie zvláštnosti, ako je napríklad nerovnaký rozptyl porovnávaných skupín, malá výskumná vzorka, či nenormálne rozloženie dát. Autor upozorňuje ešte na to, že tento prístup bol bežný až do začiatku 90. rokov, ale pretrváva až do súčasnosti.

Ako je možné vidieť z predošlých štúdií, existujú autori, ktorí sa prikláňajú k ponímaniu položiek Likertovho typu ako výsostne ordinálnych dát. Za pripomenutie stojí, že takmer žiadna z týchto štúdií neleží v oblasti pedagogického výskumu, väčšina sa pohybuje v oblasti medicíny, či ekonomiky. Obdobne si je možné všimnúť aj práce, ktorých autori už uvažujú o použití Likertových škál, ktoré sa dajú používať ako intervalové premenné. Do tohto spektra štúdií patria aj tie, ktoré sa už týkajú pedagogických výskumov.

3 Miskoncepcie a mýty používania Likertových škál

Do vyššie zmienenej kategórie patria aj štúdie od J. Carifia, ktorý vo svojich prácach, najmä v tej publikovanej v roku 2007, vo dvojici Carifio & Perla, popisujú dôvody nazerania na Likertove škály ako na intervalové premenné. Vo svojich prácach obšírnym spôsobom zdôvodňuje, prečo tento typ položiek

by sa mal vyhodnocovať vo väčšine prípadov parametrickými štatistickými metódami, prečo by sa nemali vyhodnocovať jednotlivé položky, ale samotné škály a ďalšie informácie dôležité pre korektné používanie Likertových škál (Carifio & Perla, 2007; Carifio & Perla, 2008). Zdôvodnenia vychádzajú z používania škál vo výskumných šetreniach týkajúcich sa prevažne psychologických a pedagogických výskumov, ich analýzy a interpretácie. Na základe toho stanovil 10 tzv. miskoncepcií a mýtov o používaní škál Likertovho typu, ktoré sú bežne zaužívané a používané vo výskumných šetreniach. Nižšie sú prezentované miskoncepce o Likertových škálach, ktoré sú voľne prevzaté a preložené so súhlasom autorov Carifio & Perla (2007). Mýty boli preložené za asistencie angličtinára preto, aby nedochádzalo k odchýlkam od pôvodných informácií. Okrem samotných mýtov je uvedené aj ich stručné vyvrátenie miskoncepce, prípadne aj navrhnuté správne používanie škál, či položiek.

1. Nie je nutné rozlišovať medzi škálou a formátom odpovedí. V podstate ide o jednu a tú istú vec, a čo je pravda o jednom, je pravda aj o druhom.

V podstate sa jedná o to, že „ordinalisti“ v podstate nerozlišujú medzi škálou a formátom odpovedí, na rozdiel od „intervalistov“. Dôležitý rozdiel je v tom, že škála tvorí súhrn položiek, napríklad všetkých 23 položiek dotazníku zameraného na postoje k prírodovedným predmetom (Pardo & Calvo, 2002) vytvára škálu. Odpoveďový formát je ale 5-bodový (konkrétne: úplne nesúhlasím – takmer nesúhlasím – neviem sa prikloniť k žiadnej možnosti – takmer súhlasím – úplne súhlasím). V tomto prípade škála je intervalová premenná a keď vezmeme do úvahy len jednu položku, tak sa jedná o ordinálnu premennú. Na doplnenie len treba povedať, ak spojíme výrok s odpoveďovým formátom, dostaneme položku Likertovho typu, či škálovanú položku.

2. Škálované položky pôsobia nezávisle a autonómne, bez žiadnej významnej konceptuálnej, logickej, či empirickej štruktúry, ktorá ich spája a syntetizuje.

Poprieť danú miskoncepciu sa dá na príklade atómov a molekúl, kým atóm by bolo možné porovnať k jednej položke, tak škála by mala svoj obraz v molekule, teda škálu vytvára viac položiek tak, ako je atóm tvorený viacerými molekulami. V odbornej literatúre je možné sa stretnúť s viacerými prístupmi k počtu položiek, ktoré už vytvárajú škálu, pričom jeden z najviac akceptovaných prístupov je ten, kde škálu vytvárajú minimálne tri položky (napríklad Raubenheimer, 2004), ale je možné sa stretnúť aj s ojedinelými štúdiami, kde autori uvádzajú možnosť tvorenia škály dvomi položkami (napríklad Eisinga, Grotenhuis & Pelzer, 2013). Škály na rozdiel od uvedenej miskoncepce pôsobia štruktúrované a ako jednotný celok. Tým aj splňajú

empirické kritériá pre stanovenie reliability, či oboch typov validity, ako sú logická a empirická. Ako Carifio & Perla (2007) upozorňujú pri hodnotení jednotlivých položiek sa stanovovanie validity a reliability dostáva do úzadia alebo sa úplne ignoruje.

3. Likertova škála naznačuje odpoveďový formát a naopak.

K tejto nepravdivej legende len toľko, že počet bodov na škále k položkám nenaznačuje dĺžku ani obsah Likertovej škály. Neplatí to ani naopak, ak je škála nazvaná určitým názvom, tak pomenovanie škály, ani jej rozsah (počet položiek, ktoré ju vytvárajú) neurčuje, koľko bodov na škále bude každá položka obsahovať.

4. Likertove škály nemôžu byť rozlíšené do makro a mikro konceptuálnych štruktúr.

Práveže škály by sa mali rozlišovať do makro a mikro konceptuálnych štruktúr. Pri Likertových škálach sa ako makroštruktúra myslí napríklad súbor položiek týkajúci sa jednej ústrednej témy alebo, zjednodušene povedané, dotazník obsahujúci určitý počet položiek Likertovho typu. Naproti tomu stojí mikroštruktúra, kde sa už zapája aj respondent výskumu, sú to vlastne jeho odpovede v dotazníku, ktoré sa určitým spôsobom vyhodnocujú a interpretujú.

5. Položky Likertovho typu by mali byť analyzované oddelene.

Tento typ analýzy pri Likertových škálach je častou chybou. Pravdepodobne ani samotný Likert, keď vytváral svoje škály, nepredpokladal a neuvažoval o separátnom vyhodnocovaní jednotlivých položiek. Treba si uvedomiť jeden dôležitý fakt, a to, že jedna položka nevytvára škálu v zmysle škály, ktorá sa dá merať. Môžeme to prirovnať k prísloviu „Jedna lastovička leto nerobí“. Na lepšie priblíženie, je nemožné stanoviť napríklad IQ na základe odpovede na jednu otázku, prípadne určiť postoj k fyzike z položky „Fyzika je zbytočná“ (voľne preložené z Duda & Garrett, 2008). Na presné stanovenie postojov, či IQ je potrebné použiť väčšie množstvo položiek, ktoré určia presný postoj k sledovanému javu, či IQ jedinca.

6. Pretože Likertove škály majú charakter ordinálnej premennej, tak na ich vyhodnocovanie môžu byť použité len neparametrické štatistické metódy.

V podstate pri tejto miskonceptii by boli zopakované informácie z predošlého textu, kde píšeme o charaktere Likertových škál, prípadne informácie o vyhodnocovaní intervalových premenných sa dajú nájsť v každej metodologickej knihe (napr. Chráska, 2007). Túto miskonceptiu pravdepodobne pomohla utvoriť a podporiť aj štúdia od autorov Glass, Peckham & Sanders (1972). Uvedení autori uviedli použitie neparametrickej F-metódy

na Likertove škály s tým, že metóda je robustná a vhodným spôsobom ukazuje skúmané rozdiely. Okrem toho autori ešte uviedli použitie danej metódy aj na jednotlivé položky, ak majú dostatočný počet bodov na škále (5 až 7). Štúdia sa dostala do širokého vedeckého povedomia, a stála za tým, že pretrváva aj v súčasnosti pohľad na dáta produkované z Likertových škál, ako na tie, ktoré je možné analyzovať len neparametrickými štatistickými metódami.

7. Likertove škály sú empirické a matematické nástroje, ktoré nemajú podstatný a hlboký základ a význam.

Intuitívne, tejto mylnej predstave sa dá oponovať všetkými matematickými procedúrami, ktoré je nutné vykonať v rámci analýz Likertových škál. Len ilustratívne, je nutné vypočítať reliabilitu použitého výskumného nástroja, stanoviť normalitu rozloženia dát, ktorá určuje, či môžu byť použité parametrické alebo neparametrické štatistické metódy, potom použitie štatistických metód, ktoré určujú rozdiely, či vzťahy. Použitie multidimenzionálnej štatistiky, ktorá sa riadi určitými pravidlami, preto tvrdiť, že matematické metódy sú vágne a neriadia sa pravidlami a zákonitosťami, je naozaj chybné.

8. Odpoveďový formát (rozumej body na škále) môže byť beztretné vyňatý z Likertovej škály.

To, o čo sa v tomto prípade jedná, je detailne popísané napríklad v prácach od autorov Sisson & Strocker (1989), či Clason & Dormody (1994). Likertova škála je vo svojej podstate pevne spojená s odpoveďovým formátom (názvami bodov na škále). Ale ako opisujú spomínaní autori, niekedy je možné sa stretnúť s doslova vyňatím odpoveďového formátu a jeho následným nahradením. Tento postup je však nesprávny, lebo Likertova škála a jej body tvoria jeden kľúčový celok, ktorý bol overený rôznymi štatistickými procedúrami a jeho bezdôvodné nahradenie môže viesť k zníženiu a porušeniu validity a reliability celého výskumného nástroja.

9. Likertova škála nie je systém na zachytenie a kódovanie informácií, ani neslúži na získanie skrytých názorov respondentov na určité objekty výskumu.

Ako konštatujú Carifio & Perla (2007) „Otázky sú jadrom“ (voľne preložené) a následne danú informáciu rozvíjajú, že jednotlivé položky v Likertovej škále majú zachytiť požadované informácie, ktoré boli charakterom položky doslova vylákané od respondenta a následne musia byť spracované do zmysluplného významu nielen pre samotného výskumníka, ale pre čo najširšie publikum. Je asi zbytočné uvádzať, že získané informácie slúžia na získanie odpovedí na výskumné otázky, či na testovanie hypotéz.

10. Malá pozornost', nedostatočné množstvo vedomostí, slabý vhl'ad, nedokonalé porozumenie nie sú prekážkou pri tvorbe a použití Likertových škál.

Tento bod je vo svojej podstate zhrňujúci všetkých vyššie prezentovaných bodov. Výskumníci používajúci typ výskumných nástrojov, ktorý obsahuje položky Likertovho typu by mali neustále sledovať trendy, nové informácie týkajúce sa ich používania, či vyhodnocovania. Pri analýzach autori musia dodržiavať postupy, ktoré sú správne a vedú k pravdivým informáciám tak, aby nespôsobovali chybu prvého, či druhého typu. Autori, ktorí sa venujú používaniu Likertových škál v humanitných vedách upozorňujú aj na chybné vyhodnocovanie (štatistické) jednotlivých položiek.

4 Záver

Prezentovaný text sa snažil aspoň v malej miere osvetliť použitie a analýzu škál Likertovho typu. Dôležité je uvedomiť si smerovanie nášho výskumu, lebo ako je možné vidieť, výskumné šetrenia, ktoré nie sú primárne položené na pedagogickom, psychologickom, či sociologickom základe, tak uvažujú o používaní Likertových škál, ako ordinálnych premenných, tak ako aj o ich vyhodnocovaní, ktoré je primárne založené na jednotlivých položkách, nie na škálach. Ale, ako je možné sa dočítať v prácach od samotného Likerta, tak podstata škál v sumovaných intervaloch, to znamená, že položky by sa nemali vyhodnocovať jednotlivo, ale sumárne, keď vytvárajú určitú skupinu, či dimenziu (Likert, 1932; Likert & Hayes, 1957). Na tomto základe je možné sa pozeráť na Likertove škály, ako na intervalové premenné a možnosť ich vyhodnocovať (parametrickými) metódami indukčnej štatistiky, ale aj metódami multidimenzionálnej štatistiky. Ako je možné vidieť vo výskumnom svete, najmä v tom, kde sú používané Likertove škály, pretrváva značné množstvo mýtov a mylných predstáv o ich správnom používaní a vyhodnocovaní, preto by malo byť snahou metodológov a aj ostatných akademických pracovníkov tieto nesprávne možnosti používania eliminovať.

Pod'akovanie

Autor by rád poďakoval Kristíne Maškovej za pomoc pri práci s anglickými materiálmi.

LITERATÚRA

- ALLEN, E. – SEAMAN, C. A. 2007. Likert scales and data analyses. In *Quality Progress*, 40, p. 64 – 65.
- BROWN, J. D. 2011. Likert items and scales of measurement? In *SHIKEN: JALT Testing & Evaluation SIG Newsletter*, 15, p. 10 – 14.

- CARIFIO, J. 1976. Assigning students to career education programs by preference: scaling preference data for program assignments. In *Career Education Quarterly*, 1, p. 7 – 26.
- CARIFIO, J. 1978. Measuring vocational preferences: ranking versus categorical rating procedures. In *Career Education Quarterly*, 3, p. 34 – 66.
- CARIFIO, J. – PERLA, R. J. 2007. Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. In *Journal of Social Sciences*, 3, p. 106 – 116.
- CARIFIO, J. – PERLA, R. J. 2008. Resolving the 50-year debate around using and misusing Likert scales. In *Medical Education*, 42, p. 1150 – 1152.
- CHRÁSKA, M. 2007. *Metody pedagogického výzkumu*. Praha : Grada. ISBN 978-80-247-1369-4.
- CICCHETTI, D. V. – SHOWALTER, D. – TYRER, P. J. 1985. The effect of number of rating scale categories on levels of interrater reliability: A Monte Carlo investigation. In *Applied Psychological Measurement*, 9, p. 31 – 36.
- CLASON, D. L. – DORMODY, T. J. 1994. Analyzing data measured by individual Likert-type items. In *Journal of Agricultural Education*, 35, p. 31 – 36.
- DUDA, G. – GARRETT, K. (2008). Blogging in the physics classroom: A research-based approach to shaping students' attitudes toward physics. In *American Journal of Physics*, 76, p. 1054 – 1065.
- EISINGA, R. – GROTENHUIS, M. – PELZER, B. 2013. The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? In *International Journal of Public Health*, 58, p. 637 – 642.
- FABIOLA, G. – IWIN, L. – JENNIFER, L. – ZAIRA, V. 2012. The effect of the number of answer choices on the psychometric properties of stress measurement in an instrument applied to children. In *Evaluar*, 12, p. 43 – 59.
- FERJENČÍK, J. 2000. *Úvod do metodologie psychologického výzkumu*, Praha : Portál. ISBN 978-80-7367-815-9.
- FONSECA-PEDRERO, E. – PAINO, M. – LEMOS-GIRALDEZ, S. – MUNIZ, J. 2013. Maladaptive personality traits in adolescence: Psychometric properties of the Personality Diagnostic Questionnaire-4+. In *International Journal of Clinical and Health Psychology*, 13, p. 207 – 215.
- FORERO, C. G. – OLIVARES, A. M. – PUJOL, D. G. 2009. Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS Estimation. In *Structural Equation Modeling*, 16, p. 625 – 641.
- GARDNER, H. J. – MARTIN, M. A. 2007. Analyzing ordinal scales in studies of virtual environments: Likert or lump it! In *Presence*, 16, p. 439 – 446.
- GAVORA, P. 2008. *Úvod do pedagogického výzkumu*. Bratislava : Univerzita Komenského. ISBN 978-80-223-2391-8
- GLASS, G. V. – PECKHAM, P. D. – SANDERS, J. R. 1972. Consequences of failure to meet assumptions underlying the analyses of variance and covariance. In *Review of Educational Research*, 42, p. 237 – 288.
- GÖB, R. – MCCOLLIN, C. – RAMALHOTO, M. F. 2007. Ordinal methodology in the analysis of Likert scales. In *Quality & Quantity*, 41, p. 601 – 626.

- GU, Y. – WEN, Q. – WU, D. 1995. How often is often? Reference ambiguities of the Likert-scale in language learning strategy research. In *Occasional Papers in English Language Teaching*, 5, p. 19 – 35.
- HIRNER, A. 1978. *Primárne dáta v sociológii*. Bratislava : Pravda.
- JAMIESON, S. 2004. Likert scales: how to (ab)use them. In *Medical Education*, 38, p. 1217 – 1218.
- KERLINGER, F. 1972. *Základy výzkumu chování*. Praha : Academia.
- KROSNICK, J. A. – JUDD, C. M. – WITTENBRINK, B. 2005. Attitude measurement. In D. ALBARRACIN, B. T. JOHNSON, M. P. ZANNA (Eds.) *Handbook of attitudes and attitude change*. p. 21 – 76, Mahwah, NJ : Erlbaum. ISBN 978-0-9849328-1-8.
- LEUNG, S. O. 2011. A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales. In *Journal of Social Service Research*, 37, p. 412 – 421.
- LIKERT, R. 1932. A technique for the measurement of attitudes. In *Archives of Psychology*, 22, p. 1 – 55.
- LIKERT, R. – HAYES, S. 1957. *Some applications of behavioural research*. Paris : Unesco.
- MURRAY, J. 2013. Likert data: What to use, parametric or non-parametric? In *International Journal of Business and Social Science*, 4, p. 258 – 264.
- NORMAN, G. 2010. Likert scales, levels of measurement and the “laws” of statistics. In *Advances in Health Science Education*, 15, p. 625 – 632.
- PARDO, R. – CALVO, F. 2002. Attitudes toward science among the European public: a methodological analysis. In *Public Understanding of Science*, 11, p. 155 – 195.
- PELIKÁN, J. 1998. *Základy empirického výzkumu pedagogických jevů*. Praha : Karolinum. ISBN 80-7184-569-8.
- PELL, G. 2005. Uses and misuses of Likert scales. In *Medical Education*, 39, p. 970 – 970.
- RAUBENHEIMER, R. 2004. An item selection procedure to maximise scale reliability and validity. In *Journal of Industrial Psychology*, 30, p. 59 – 64.
- ROD, A. 2012. Likertovo škálování. In *E-Logos Electronic Journal for Philosophy*, 13, 13 s.
- SISSON, D. A. – STOCKER, H. R. 1989. Analyzing and interpreting Likert-type survey data. In *The Delta Pi Epsilon Journal*, 31, p. 81 – 85.
- STEVENS, S. 1946. On the theory of scales of measurement. In *Science*, 103, p. 668 – 690.
- STRASSER, N. 2007. Avoiding statistical mistakes. In *Journal of College Teaching & Learning*, 4, p. 51 – 58.
- VICKERS, A. 1999. Comparison of an ordinal and a continuous outcome measure of muscle soreness. *International Journal of Technology Assessment in Health Care*, 15, p. 709 – 716.
- WALKER, I. 2013. *Výzkumné metody a statistika*. Praha : Grada. ISBN 978-80-247-3920-5.
- WU, C. H. 2007. An empirical study on the transformation of Likert-scale data to numerical scores. In *Applied Mathematical Sciences*, 1, p. 2851 – 2862.

ZUMBO, B. D. – GADERMANN, A. M. – ZEISSER, C. 2007. Ordinal versions of coefficients alpha and theta for Likert rating scales. In *Journal of Modern Applied Statistical Methods*, 6, p. 21 – 29.

Milan Kubiato pracuje na Fakulte humanitných vied Žilinskej univerzity. Profiluje sa v metodológii pedagogického výskumu a v štatistickej analýze dát. Výskumne sa zaoberá problematikou mylných predstáv, postojov k vyučovacím predmetom a hľadania faktorov ovplyvňujúcich úroveň mylných predstáv a postojov. Ďalej sa zaoberá témami súvisiacimi s motiváciou k učeniu, typológiou osobnosti a jej vplyvom na pedagogické javy, problematikou IKT a ďalšími pedagogicko-psychologickými otázkami.

PaedDr. Milan Kubiato, PhD.
Katedra pedagogických štúdií
Fakulta humanitných vied ŽU
Univerzitná 1
01026 Žilina
mkubiato@gmail.com