

## Možnosti využití Likertovy škály – základní principy aplikace v pedagogickém výzkumu a demonstrace na příkladu zjišťování vztahu člověka k přírodě

*Vlastimil Chytrý, Roman Kroufek*

### Abstrakt

Obsah článku je primárně věnován možnostem zpracování dat získaných na základě Likertova škálování. Představuje principy nakládání s takovým typem dat a následně ukazuje, jak s nimi lze pracovat. Užita jsou data získaná pomocí škály Nature relatedness scale, která je určena ke zjišťování vztahu jedince k přírodě kvantitativní cestou. Cílem článku je popsat způsob zpracování dat získaných pomocí Likertovy škály a reflektovat postup jejich zpracování na příkladu konkrétních, výzkumem reálně získaných dat, která informují o vztahu jedince k přírodě.

**Klíčová slova:** Likertova škála, vztah k přírodě, metodologie výzkumu, environmentální výchova.

## Possibilities of Using the Likert's Scale — Basic Principles of Application in Pedagogical Research and Demonstration on the Example of Human Relationship to Nature

### Abstract

This article is focused on the possibilities in data processing based on the Likert scale. The article outlines the principles guiding the use of data suitable for this particular scale analysis, and then demonstrates how to work with such data sets. The test uses data generated by Nature relatedness scale which is commonly used for analysing the relationship between test subjects and nature with the use of a quantitative method. The main goal of the article is to describe possible ways how to process and analyze such data through the Likert scale survey. In doing so, it also substantiates the procedure of the data processing with reference to empirical data which provide information about man-nature relationship.

**Key words:** Likert scale, nature relatedness, research methodology, environmental education.

# 1 ÚVOD

K sepsání článku nás vedla skutečnost, že se v současnosti stále zvyšuje počet výzkumných prací, v nichž jsou s menšími nebo většími nedostatky analyzována data pomocí Likertových škál<sup>1</sup> (Likert, 1932), které patří právě v dnešní době k často využívané metodě sběru dat. Správné zpracování dat, ať už v rámci diplomových nebo disertačních prací, případně článků publikovaných v recenzovaných časopisech, představuje širokou problematiku, na kterou upozorňuje například Ruskova práce (2015) věnující se analýze disertačních prací. Jedná se o velmi aktuální téma, neboť Likertovy škály jsou v pedagogickém výzkumu hojně používány (viz např. Šmejkal et al., 2016; Vácha & Ditrich, 2016).

Cílem příspěvku je analyzovat možnosti zpracování dat pomocí Likertovy škály. Závěry, ke kterým autoři dospěli, jsou prezentovány na ukázce reálně získaných dat, přičemž jejich záměrem není prezentovat jednotlivé škály jako výzkumné nástroje, ale nastínit možný postup zpracování dat za využití Likertových škál. Snahou je poukázat na základní nedostatky při tomto zpracování dat a podrobně specifikovat vhodný způsob práce s danou škálou. Ačkoliv se jedná o škály, které jsou velmi často používány v dotaznících zaměřených na ověřování postojů a názorů (Bendl, 2011), dochází často k formálnímu využívání statistických metod a mnohdy i k mylným závěrům. Takový způsob analýzy dat pak vede k tomu, že statistická významnost je používána i za předpokladu nesplnění podmínek, které představuje náhodný či randomizovaný výběr (Blahuš, 2000).

Významnou otázkou pedagogického výzkumu je kvantifikace aspektů afektivní dimenze osobnosti za účelem jejich následného statistického zpracování. Jde především o to, jak transformovat subjektivní charakteristiky jedince do objektivní reality (Říčan & Pešout, 2013; Joshi et al., 2015) a jakým způsobem získat z výzkumu afektivní složky osobnosti (postoje, hodnoty, názory, senzitivita apod.) data ve formě čísel. Nejčastější cestou může být právě škálování, tedy převod objektů na čísla podle určitého pravidla (Stevens, 1959).

K takovým účelům jsou v rámci empirických studií a kvantitativního zpracování dat velice často užívány Likertovy škály, které se společně s Thurstoneovým škálováním stejně se jevících intervalů a Gutmannovou kumulativní metodou řadí mezi tzv. jednodimenzionální škálování (Rod, 2012). Tyto škály, jež se víceméně plošně používají jako nástroj shromažďování dat (Weng & Cheng, 2000), jsou nejčastěji užívány pro měření postojů či názorů, kdy respondent na předem dané stupnici vyjadřuje míru svého souhlasu či nesouhlasu s výrokem.<sup>2</sup> Právě četnost užití těchto škál a nejednoznačnost ve způsobu jejich zpracování jsou důležitým jevem, na který je zapotřebí se zaměřit.

V mnoha prezentovaných studiích pak panuje názorová nejednotnost zejména v tom, jak by se mělo s danými škálami pracovat (viz oddíl 2.1). Diskutována je nejen problematika rozsahu škály, případně parity položek, ale také charakteristika a následné statistické vyhodnocení získaných dat. Clason a Dormody (1994) popisují různé typy škál, včetně těch, v nichž jsou vynechávány neutrální hodnoty. Za

---

<sup>1</sup> V českém prostředí se používá pojem škála jak pro samostatnou stupnici (například 1–5), tak také pro soubor položek. V angličtině se rozlišuje mezi Likert Type Scale (= samostatná stupnice) a Likert Scale Data (= soubor položek).

<sup>2</sup> Jinou z možností, jak analyzovat postoje či hodnotící názory, je využití sémantického diferenciálu tak, jak jej popisuje Kubiátko (2016). Problematika užívání sémantického diferenciálu však není předmětem našeho výkladu, a proto zde není podrobněji analyzována.

typické pak považují zejména pětibodové<sup>3</sup> stupnice. Existují však argumenty ve prospěch sedmistupňových škál nebo škál se sudým<sup>4</sup> počtem možností (Jamieson, 2004). K rozsahu škály se váže postřeh Boona a Boonové (2012), kteří se věnují využití Likertových škál ve svých příspěvcích publikovaných v časopise *Journal of Extension* v roce 2011. Právě v tomto roce vyšlo v tisku značné množství odborných studií, v nichž byla data zpracována prostřednictvím čtyř až sedmistupňové Likertovy škály.

Možnosti práce s Likertovými škálami se pokusíme v následujícím textu prezentovat na konkrétní ukázce dat, která byla získána při výzkumu vztahu studentů oboru Učitelství pro 1. stupeň ZŠ k přírodě. Naším cílem nebude podrobně referovat o výsledcích a závěrech tohoto bádání (blíže např. Kroufek, Janovec & Chytrý, 2015; Kroufek, 2016), nýbrž poukázat na vhodný způsob, jak s takto získanými daty nakládat. Podrobný popis způsobu zpracování dat bude v další části textu nastíněn, a to včetně výzkumného problému, který se k dané problematice váže.

Jako vzorový výzkumný nástroj, který vhodně využívá Likertovo škálování, jsme zvolili škálu *Nature relatedness scale* – NRS (Nisbet, Zelenski & Murphy, 2009; Nisbet, 2011), jejíž český překlad poprvé publikoval Franěk (2012). Jedná se o škálu vycházející z hypotézy biofilie (Wilson, 1984; Kellert & Wilson, 1993) a složené z jednadvaceti položek, k nimž respondent vyjadřuje míru souhlasu na pětistupňové Likertově škále. Cílovým konstruktem, jenž NRS měří, je „nature relatedness“<sup>5</sup>, který vyjadřuje pochopení a oceňuje propojení s ostatními živými tvory na Zemi (Nisbet, Zelenski & Murphy, 2009). Výzkumný nástroj je trojfaktorový a dělí se na tři subškály, přičemž každá z nich je zaměřena na jinou kvalitu vztahu člověka k přírodě. Subškála NR-já (NR-self) se skládá z devíti položek a zaměřuje se na vnitřní propojení jedince s přírodou. NR-perspektiva (NR-perspective) má položek šest a věnuje se jedincovu pohledu na přírodní svět. Konečně faktor NR-prožitky (NR-experience) je také šestipoložkový a postihuje jedincovu touhu být s přírodou v interakci.

Stejní autoři dále prokázali, že konstrukt „nature relatedness“ je prediktorem osobní pohody a psychického zdraví jedince (Nisbet, Zelenski & Murphy, 2011). Sariçam, Şahin a Soyuçok (2015) pak ve své studii zjistili mimo jiné i negativní korelaci mezi „nature relatedness“ a stresem, úzkostí a depresí respondentů. K obdobným zjištěním dospěli také Martynová a Brymer (2014). Význam měření tohoto konstruktů je tedy evidentní a využití této škály je ve výzkumech zcela relevantní. Nevýhodou je však její délka, a proto se autoři pokusili tento handicap řešit vytvořením zkrácené verze škály, a to NR-6 (Nisbet & Zelenski, 2013). Jak už její název napovídá, skládá se z pouhých šesti položek, vybraných ze subškál NR-já a NR-prožitky. NR-6 velmi silně pozitivně koreluje s původním nástrojem ( $r = 0,91$ ) a její autoři jej doporučují jako jeho možnou a použitelnou alternativu (Nisbet & Zelenski, 2013). Jako kvalitní nástroj, jenž je svou subtilností vhodný do širokého spektra výzkumů, jej doporučují také Craig, Logan a Prescott (2016).

---

<sup>3</sup>Průcha a Veteška (2012) upozorňují na skutečnost, že nejčastěji jsou používány sedmibodové škály. Tito autoři, např. Vácha, Ditrich (2016), Rusek (2015), Hladě (2015) a další, považují za častěji užívané pětibodové škály. Vlčková et al. (2015) však poukazují na skutečnost, že z hlediska psychometrických vlastností není rozdíl mezi pěti a sedmibodovou škálou.

<sup>4</sup>Problematice sudého nebo lichého stupně škály se věnuje Rod (2012).

<sup>5</sup>Pojem „nature relatedness“ se v českých výzkumných pracích zpravidla nepřekládá (viz Franěk, 2012; Kroufek, 2016). Bývá vyjadřován jako konstrukt měřící vztah k přírodě či sepjetí s přírodou.

## 2 ZPRACOVÁNÍ LIKERTOVY ŠKÁLY

Struktura tohoto oddílu odpovídá standardnímu postupu, který je vhodné dodržovat při zpracování dat získaných pomocí Likertova škálování. Nejprve je nutné vymezit typ použitého měřítka, detekovat odlehlé hodnoty a následně ověřit psychometrické vlastnosti nástroje. Až pak může následovat vlastní analýza dat.

### 2.1 VOLBA TYPU MĚŘÍTKA

Vlastnímu vyhodnocení dat, která byla získána na základě Likertova škálování, musí předcházet rozbor jejich charakteru, respektive typu. Jinými slovy, je nutné si uvědomit, s jakými proměnnými pracujeme ve vztahu k použitému měřítku<sup>6</sup> podle Ary, Jacobse a Sorensena (2010), neboť právě od tohoto vztahu se odvíjí následná analýza. Podle Hendla (2012) rozlišujeme čtyři různá měřítka, viz tabulka 1.

Tab. 1: Popis základních měřitek

Typ měřítka	Popis	Příklad
<b>Nominální</b> měřítko	Měřítka určující rozlišnost mezi třídami, které není možné seřadit.	Pohlaví, barva, sport
<b>Ordinální</b> měřítko	Kromě odlišnosti tříd je možné je seřadit na základě intenzity nebo pořadí.	Prospěch
<b>Intervalové</b> měřítko	Podobné ordinálním, avšak s tím rozdílem, že mezi jednotlivými třídami jsou vzdálenosti dány určitou jednotkou.	Testy inteligence
<b>Poměrové</b> měřítko	Stupnice začínající nulou, přičemž všechny jednotky mají stejný rozměr a je možné s nimi provádět všechny matematické a statistické operace.	Výška, váha, věk, délka skoku

Ukazuje se, že ze statistického hlediska není parita ani počet stupňů určující. Problém nastává spíše ve chvíli, kdy řešíme povahu jednotlivých stupňů.

Hlubšímu rozboru zařazení dat získaných na základě Likertova škálování do konkrétního měřítka se věnuje celá řada autorů, např. Jakobsson (2004), Kuzon, Urbanek a McCabe (1996), Albaum (1997), případně i Hodge a Gillespie (2003). V rámci práce s jednotlivými položkami například Brown (2011) rozporuje, že by získaná data byla ordinální. Tento názor je v kontradikci se závěry Carifia a Perly (2007), podle nichž je chybným mýtem tvrzení, že data pocházející z Likertovy škály jsou ordinálního charakteru a měly by na ně být aplikovány neparametrické statistické metody. Přikláníme se k názoru Subediho (2016), že s danými daty není možné pracovat jako s nominálními, nýbrž ordinálními, protože mají specifické uspořádání, které McLeod (2008) popisuje jako lineární (ve stupních od „absolutně souhlasím“ po „absolutně nesouhlasím“). Carifia a Perly (2007) navíc dodávají, že není vhodné analyzovat jednotlivé položky, ale celé škály.

Pokud budeme hovořit o vyhodnocení celé Likertovy škály, máme na mysli jednu proměnnou, která vznikla sloučením (v našem případě součtem) minimálně čtyř různých položek do jedné proměnné (Boone & Boone, 2012; Joshi et al., 2015). V tomto případě budeme vycházet ze závěrů a způsobu užití dané škály u řady

<sup>6</sup>Zpravidla jsou konzultovány pouze tři stupnice (nominální, ordinální a metrické – zahrnují intervalové a poměrové stupnice). Více viz Brown (1988).

autorů (např. Baggaley & Hull, 1983; Allen & Seaman, 1997; Maurer & Pierce, 1998; Vickers, 1999) a danou proměnnou považovat za intervalovou.<sup>7</sup>

Nutnost vymezit typ škály vzhledem k použitému měřítku vyplývá z možnosti následné deskriptivní analýzy, která je přehledně zanesena do tabulky 2. V této tabulce se omezíme pouze na tři typy měřitek, neboť ze statistického hlediska není rozdíl mezi intervalovým a poměrovým typem měřítka, a to vzhledem ke zmíněným parametrům / možnostem operace.

Tab. 2: Možnosti deskriptivní analýzy vzhledem k použitému měřítku

Statistický parametr / možnost operace	Typ měřítka		
	Nominální	Ordinální	Intervalové
Četnost	✓	✓	✓
Modus	✓	✓	✓
Medián		✓	✓
Aritmetický průměr			✓
Směrodatná odchylka			✓
Možnost kvantifikace rozdílu mezi hodnotami			✓
Možnost přidání nebo ubrání hodnoty			✓

Způsoby analýzy Likertových škál jsou často odlišné, což vede k závěrům, které jsou do jisté míry sporné. Budeme-li uvažovat získaná data jako ordinální, není možné počítat průměr, směrodatnou odchylku, faktorovou analýzu, analýzu rozptylu a další.<sup>8</sup> Do tabulky 3 je zaneseno, jakým způsobem je vhodné zpracovat data, která byla získána na základě Likertova škálování za předpokladu, že jednotlivé položky jsou hodnoceny na ordinálním měřítku a škála jako celek na intervalovém měřítku.

Tab. 3: Schéma možné analýzy pro hodnocení jednotlivých položek a hodnocení škály jako celku (podle Subedi, 2016)

Statistická metoda	Hodnocení položek prostřednictvím parametru / testu	Hodnocení škály jako celku prostřednictvím parametru / testu
Míra ústřední tendence	Medián nebo modus	Průměr
Míra variability	Frekvence (četnost)	Směrodatná odchylka
Míra asociace	Kendalovo tau b nebo c, Spearmanovo rhó	Pearsonovo R
Ostatní statistiky	$\chi^2$ test, Mann-Whitney U-test	ANOVA, t-test

<sup>7</sup>Boone a Boone (2012) popisují, že při zpracování dat na intervalové stupnici je zapotřebí využít parametrické statistické metody. My se však s tímto tvrzením neztotožňujeme a vycházíme z Hendlova (2012: s. 197) názoru, že neparametrické metody jsou vhodné pro data z intervalového měřítka, jež nemají normální rozdělení četností. Normalita dat je tedy považována za nutný předpoklad pro využití parametrických statistických metod.

<sup>8</sup>Kubiatko (2016) popisuje možnou analýzu dat vzešlých z Likertova škálování a věnuje se základním miskonceptům při použití tohoto škálování.

## 2.2 DETEKCE ODLEHLÝCH HODNOT

V úvodu tohoto oddílu považujeme za důležité zmínit terminologické nejasnosti související s podezřelou, odlehlou nebo vybočující hodnotou. Jak uvádí Hebák (2013: s. 199), „česká statistická literatura se zcela obecně o pojmově jednotnou klasifikaci statistických pojmů či symbolů nepokouší a zahraniční ekvivalenty mají příliš široký význam“. Vzhledem k zaměření článku a pro významné zjednodušení budeme každé pozorování, které není konzistentní s většinou ostatních nebo neodpovídá předpokladu o populačním pravděpodobnostním rozdělení sledované veličiny, označovat za odlehlou hodnotu. Využijeme-li jiného termínu, zmíníme, z jakého autora při komentování výsledků vycházíme.

Pokud dojde k detekci odlehlé hodnoty například pomocí grafických metod, jako je kvartilový graf (angl. Box-and-Whisker Plot), případně jeho dvojrozměrná verze Bagplot, je vhodnější zabývat se samotnou otázkou vzniku dané hodnoty. Jestliže je odlehlá hodnota identifikována<sup>9</sup>, ukazuje se jako nejvhodnější ověřit, zda se nejedná o chybu v měření. Pokud pracujeme s velkými soubory, lze očekávat, že se vždy najde takováto hodnota (Hendl, 2012). Jejich mechanické odstraňování není vždy vhodné, jelikož pro některá měření jsou odlehlé hodnoty typické. V této části se výzkumník musí opírat o vlastní názor, jelikož se často objevují sporné hodnoty, které vyžadují rozhodnutí o ponechání / vyřazení z matice dat (Hebák, 2013). Podrobně se problematice průzkumové analýzy dat a řešením situace, kdy jsou detekovány odlehlé hodnoty, věnuje Madansky (1988).

Samotné vyhledávání odlehlých hodnot je možné například pomocí metody vnitřních hradeb (ukázka v oddíle 3.2). Pokud bychom chtěli použít některý z testů pro detekci těchto hodnot, pak se zpravidla využívají dva: Grubbsův test (využívá se za předpokladu normálního rozdělení dat) a Dean-Dixonův test (testování malých<sup>10</sup> souborů nebo souborů s neznámým rozdělením). V praxi jsou odlehlé hodnoty často ze souboru dat vyřazovány, jelikož zpravidla vznikají chybou v měření a uměle zvyšují hodnotu směrodatné odchylky jako míry rozptýlu dat.

## 2.3 RELIABILITA

U nástrojů využívajících Likertových škál je nejčastěji používaným koeficientem, jímž se zjišťuje jejich vnitřní konzistence, Cronbachova alfa (Cronbach, 1951; McGartland Rubio, 2005). Tento koeficient může nabývat hodnot v intervalu  $\langle 0, 1 \rangle$ , přičemž obecně akceptovatelné hodnoty<sup>11</sup> jsou mezi 0,7 a 0,95 (Tavakol & Dennick, 2011). Hopkins (1998) je podstatně přísnější a udává, že při standardizaci výzkumného nástroje je možné považovat za dostatečné pouze hodnoty reliability vyšší než 0,90.

---

<sup>9</sup>Vzhledem k zaměření článku je zmiňována pouze identifikace odlehlých pozorování v jednorozměrných statistických souborech. Grafická identifikace vícerozměrných extrémů například pomocí Mahalanobisovy vzdálenosti, případně popisu efektů, jež zmiňuje Ben-Gal (2005), zde není záměrně brána v úvahu.

<sup>10</sup>Za malé soubory jsou považovány soubory o četnostech 3–10 pozorování (viz např. Meloun & Militký, 2012).

<sup>11</sup>Podle autorů Nun-Nally (1978) je reliability dostačující také v mezích 0,50–0,60 za předpokladu, že byl nástroj publikovaný poprvé. Oproti tomu Sekaran (1992) stanovil minimální přijatelnou hodnotu 0,60. Na základě tvrzení Shoukriho a Edge (1996) je možné říci, že obecný koeficient reliability (označme jej  $r$ ) je možné považovat za excelentní, pokud je  $r > 0,75$ , dobrý, pokud je  $r$  v intervalu 0,40–0,75, a nízký, pokud  $r < 0,40$ .

Využití Cronbachovy alfy však bývá často špatně interpretováno, a to zejména z důvodu, že tento koeficient je vhodný pro nástroje pracující s jednou latentní proměnnou (Garrido, Abad & Ponsoda, 2013). Jeho rutinní využívání nemusí být vhodné (Martínková & Vlčková, 2014) a odporuje tvrzení, že není možné použít pouze jeden odhad reliability a ten považovat za optimální. Nabízejí se porovnání s více odhady, jak například popisují Revelle a Zinbarg (2009). Základními předpoklady (Martínková & Vlčková, 2014) pro použití Cronbachovy alfy jsou:

- náhodná chyba je jediným zdrojem chyby měření,
- položky popisují jedinou latentní proměnnou,
- položky jsou nekorelované,
- lze předpokládat, že položkové skóry se řídí přibližně normálním rozdělením.

Pokud je nutné uvažovat více než jednu latentní proměnnou, bude Cronbachova alfa podhodnocovat skutečnou reliabilitu (Novick & Lewis, 1967). Za předpokladu více latentních proměnných je vhodné využít odhad  $\omega_t$ , popsany McDonaldem a Roderickem (1978).

Závěry Gadermanna, Guhna a Zumby (2012) prokazují, že Cronbachova alfa se využívá v 76 % prezentovaných výzkumů jako koeficient spolehlivosti. Domníváme se, že využití Cronbachovy alfy je natolik rozšířené zejména z důvodu dostupnosti software jako Statistica nebo SPSS. Pro další koeficienty spolehlivosti (viz Revelle & Zinbarg, 2009) sice existují speciální aplety, které je možné najít na internetu, avšak jejich dostupnost je omezená.

Ve vybraných případech zjišťování reliability je možné testovat split-half reliabilitu (Johnson & Penny, 2005). Při ní se test rozděluje do dvou polovin a zjišťuje se jejich korelace, která bývá zpravidla doplněna o korekci pomocí Spearman-Brownova vzorce (Johnson & Penny, 2005), jenž koriguje problém s dělením testu v průběhu zjišťování split-half reliability, jak blíže charakterizuje Chráška (1999). Pokud má výzkumník možnost vidět se s respondenty opakovaně, je vhodné využít test-retest reliability, založený na korelaci výkonů testovaných osob při dvojnásobném po sobě jdoucím měření stejným testem (Koníček et al., 2007). Mezi prvním a opakovaným testováním je vhodné dodržet delší časový úsek, a to proto, aby byla větší pravděpodobnost, že si respondenti nebudou pamatovat své předchozí odpovědi (Martínková & Vlčková, 2014). Možnostmi odhadu spolehlivosti se věnují například Richardson (1939), Zvára (2002), Pulpán (2004) a další.

## 2.4 VALIDITA

Validita je jedním z nejdůležitějších kritérií samotného bádání (Mazalová & Marečková, 2012). Jedná se o základní pojem psychometrie (Gavora, 2013) odpovídající na otázku: „Do jaké míry zkoumáme to, co zkoumat chceme?“ (Kerlinger, 1972: s. 435; Chráška, 2007: s. 37). Mezi nejčastější požadavky zaručující validitu patří primárně relevantnost nároků na respondenta (počet položek a časové zatížení), srozumitelnost položek výzkumného nástroje a adekvátnost měřící stupnice. Za neopomenutelné je nutné považovat reprezentativnost výběrového souboru a uspořádání položek v rámci dotazníku.

Parker a Lunney (1998) udávají, že v odborných textech se zpravidla hovoří o třech typech validity:

1. *Obsahová* – založená na posudku, úsudku kompetentních osob a popisujících, nakolik daný test odpovídá zkoumanému fenoménu (Kerlinger, 1972).

2. *Kriteriální* – pod pojmem kriteriální validizace se skrývá jak prediktivní, tak souběžná validita. V obou případech vychází proces validizace z hledání kritéria, s nímž je daný test dáván do vztahu (Pelikán, 2011).
3. *Konstruktová* – využívá se ve chvíli, kdy nástroj měří konkrétní rys. Ve své podstatě hovoří o míře platnosti závěru oproti danému předpokladu (Polit & Beck, 2008).

### 3 DEMONSTRACE MOŽNÉHO ZPRACOVÁNÍ ŠKÁLY NRS

Jak bylo již zmíněno v úvodu článku, je naší snahou co možná nejzřetelněji demonstrovat, jakým způsobem je možné zpracovat data získaná na základě Likertova škálování. Následující text má právě z tohoto důvodu strukturu analogickou k textu předchozí části, v níž seznamujeme s principy práce s Likertovou škálou. Demonstrace aplikace zpracování získaných údajů je tak založena na reálných výzkumných datech. Ta zde však neslouží ke zjištění vztahu k přírodě, neboť výsledky výzkumu byly již na základě shromážděných dat publikovány Kroufkem, Janovcem a Chytrým (2015) a Kroufkem (2016).

#### 3.1 VÝZKUMNÝ PROBLÉM A CÍL VÝZKUMU

Podle Gavory (2010) i Chrásky (2016) může mít výzkumný problém podobu tvrzení či otázky. „Správně formulovaný výzkumný problém je otázka, která by měla vyjadřovat vztah mezi proměnnými“ (Chráska, 2016: s. 13). Výzkumný problém může být trojího typu, a to deskriptivní, relační či kauzální (Gavora, 2010). Ze správně definovaného výzkumného problému jsou potom tvořeny hypotézy a případné subhypotézy.

*Výzkumným problémem je tedy otázka:* Jaká je využitelnost vybraného výzkumného nástroje (NRS) pro zjišťování vztahu k přírodě studentů Učitelství pro 1. stupeň ZŠ?

K vlastnímu výzkumnému problému se pak váže následující cíl: Zjistit využitelnost vybraného výzkumného nástroje (NRS) pro zjišťování vztahu k přírodě studentů Učitelství pro 1. stupeň ZŠ.

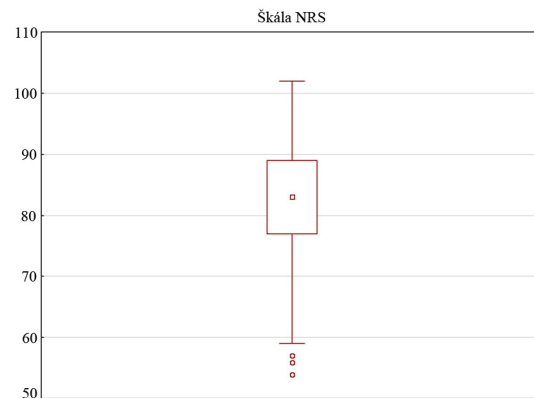
#### 3.2 VZOREK RESPONDENTŮ, VÝCHOZÍ DATA UŽITÁ PRO DEMONSTRACI

Respondenty výzkumu ( $N = 503$ ) byli náhodně vybraní studenti oboru Učitelství pro 1. stupeň ZŠ Univerzity J. E. Purkyně ( $N = 326$ ), Jihočeské univerzity ( $N = 104$ ) a Univerzity Karlovy ( $N = 73$ ). Relativně rovnoměrně byli pak zastoupeni studenti prezenční ( $N = 227$ ) a kombinované ( $N = 277$ ) formy studia. Zastoupení žen a mužů ve vzorku respondentů bylo výrazně odlišné. V celém vzorku bylo pouhých 30 mužů, což je dáno specifikou daného studijního oboru. Průměrný věk respondentů byl 28,08 let (UJEP = 28,1; UK = 28,35; JU = 28,09).

#### 3.3 ODSTRANĚNÍ ODLEHLÝCH HODNOT

Vlastnímu šetření, zaměřenému na odlehlé hodnoty, předcházelo ověřování normality pomocí Shapiro-Wilkova testu normality (Shapiro & Wilk, 1965), kdy testujeme proti nulové hypotéze, že posuzovaná data mají normální rozdělení. U škály





Obr. 1: Kvartilový graf škály NRS

NRS nulová hypotéza o normálním rozložení dat odmítnuta nebyla ( $W = 0,98$ ;  $p = 0,255$ ). Bylo tedy možné pro odstranění odlehlých hodnot použít metody zmíněné v oddíle 2.1. Při detekci odlehlých hodnot jsme vycházeli z kvartilového grafu na obrázku 1.

Výsledky výpočtu odpovídající kvartilovému grafu jsou uvedeny v tabulce 4. Tyto hodnoty jsou důležité pro výpočet horních a dolních hradeb.

Tab. 4: Výběrové charakteristiky

$q_1$	medián	$q_3$	průměr	$Q = q_1 - q_3$	Min	Max
77	83	89	82	12	54	102

První „náповědou“ týkající se existence odlehlé hodnoty je zpravidla velký rozdíl mezi mediánem a průměrem, což v našem případě neplatí. Budeme-li se držet obvyklého úzu a využijeme-li popisu, který uvádí Hebák (2013: s. 192), pak dojdeme k následujícímu:

$$\text{Dolní hranice vnitřních hradeb: } q_1 - 1,5Q = 77 - 1,5 \cdot 12 = 59$$

$$\text{Horní hranice vnitřních hradeb: } q_3 + 1,5Q = 89 + 1,5 \cdot 12 = 107$$

$$\text{Dolní hranice vnějších hradeb: } q_1 - 3Q = 77 - 3 \cdot 12 = 41$$

$$\text{Horní hranice vnějších hradeb: } q_3 + 3Q = 89 + 3 \cdot 12 = 125$$

Máme-li k dispozici tyto čtyři údaje, je pak již následná detekce hledaných hodnot jednoduchá. Hebák (2013: s. 192) uvádí: Hodnoty uvnitř vnitřních hradeb jsou považovány za běžné, hodnoty mezi vnitřními a vnějšími hradbami jsou považovány za podezřelé, hodnoty za vnějšími hradbami jsou označovány jako odlehlé. Jak z výpočtu, tak také z grafu na obr. 1 je zřejmé, že hodnoty  $x < 59$  jsou považovány za podezřelé. Odlehlé hodnoty se v matici dat nevyskytují. V další části výpočtu budeme pokračovat tak, že vyřadíme hodnoty, které byly na základě této detekce identifikovány jako podezřelé. Ukázalo se, že u těchto hodnot došlo k pochybení při přepisování dat do datové matice.

### 3.4 RELIABILITA

Ke každé položce škály NRS vyjadřuje respondent míru souhlasu na standardní pětistupňové Likertově škále (1 – souhlasím, 2 – spíše souhlasím, 3 – neutrální postoj, 4 – spíše nesouhlasím, 5 – nesouhlasím). Reliabilita představené škály byla testována pomocí Cronbachovy alfy a split-half reliability. Zjištěné hodnoty srovnáváme s reliabilitou identifikovanou při prvním publikování nástroje jeho autory (Nisbet, Zelenski & Murphy, 2009) a prvního českého překladu (Franěk, 2012). Zjištěné hodnoty jsou uvedeny v tabulce 5.

Tab. 5: Reliabilita škály NRS

Způsob vyhodnocení reliability	Získaná hodnota pro škálu NRS
Cronbach $\alpha$	0,82
split-half	0,83
Cronbach $\alpha$ (Nisbet et al., 2009)	0,87
Cronbach $\alpha$ (Franěk, 2012)	0,84

Vzhledem ke zjištěným hodnotám a tvrzení z oddílu 2.2 je možné konstatovat, že škála vykazuje vysoké hodnoty spolehlivosti a jejich přesnost je pro využití v pedagogickém výzkumu dostatečná.

Škála NRS je autory dále členěna, a to na tři segmenty. Hodnoty reliability je možné nalézt v tabulce 6, kde je zároveň srovnáváme s originálními hodnotami tvůrců škály a hodnotami prvního použití českého překladu (Franěk, 2012). V případě těchto subškál je však již každá zaměřena na jednu latentní proměnnou a koeficient Cronbachovy alfy se tak stává vhodnějším než v předcházejícím případě.

Tab. 6: Reliabilita škál NR – já, NR – perspektiva a NR – prožitky

Způsob vyhodnocení reliability	NRS – já	NR – perspektiva	NR – prožitky
Cronbach $\alpha$	0,79	0,54	0,65
Split-half	–	–	–
Cronbach $\alpha$ (Nisbet et al., 2009)	0,84	0,66	0,80
Cronbach $\alpha$ (Franěk, 2012)	0,76	0,43	0,72

Z tabulky 6 je patrné, že nedostatečných hodnot dosahuje zejména škála NR-perspektiva. Zbylé subškály je možné při správné interpretaci využít separátně.

### 3.5 VALIDITA

Validita použitého nástroje byla zjišťována několika způsoby. V případě, že to podmínky umožňovaly, byla zjišťována souběžná kriteriální validita (Hendl, 2012). V relevantních případech je diskutována konstruktová validita (Cronbach & Meehl, 1955). Predikční validita, o které byla zmínka v oddíle 2.3, nebyla vzhledem k cílům a zaměření předkládaného textu testována.<sup>12</sup>

#### KRITERIÁLNÍ VALIDIZACE

Souběžná kriteriální validita použitého nástroje byla ověřována jeho vzájemnou korelací s dalšími nástroji, které byly zaměřeny na příbuznou doménu. Konkrétně jde o škály New ecological paradigm – NEP (Dunlap, Van Liere & Mertig, 2000) a 2-MEV (Johnson & Manoli, 2008). U obou zmíněných škál došlo k zamítnutí nulové hypotézy o normalitě dat, která byla ověřena na základě Shapiro-Wilkova testu normality: NEP ( $W = 0,96$ ;  $p = 0,007$ ) a 2-MEV ( $W = 0,95$ ;  $p = 0,002$ ). Z tohoto

<sup>12</sup>Zmíňme jen, že u škály NRS ji zkoumal Franěk (2012), který odhalil signifikantní rozdíly ve vztahu k přírodě mezi studenty přírodovědně a technicky zaměřených fakult Univerzity J. E. Purkyně. K obdobnému závěru došli také Kroufek, Janovec a Chytrý (2015).

Tab. 7: Korelace škál měřících afektivní složku EG – Spearmanovo  $\rho$

	NEP	2-MEV	NRS
NEP		0,53	0,56
2-MEV	0,53		0,78
NRS	0,56	0,78	

důvodu byla nadále použita neparametrická Spearmanova korelace. U škály NRS nulová hypotéza o normálním rozložení dat odmítnuta nebyla ( $W = 0,98$ ;  $p = 0,255$ ). Zmíněné korelace jsou uvedeny v tabulce 7.

Především vysoká pozitivní korelace mezi NRS a dvoudimenzionální 2-MEV garantuje jejich vysokou kritériální validitu.

## OBSAHOVÁ A KONSTRUKTOVÁ VALIDIZACE

Obsahovou a konstruktovou validitu NRS potvrzuje její široké nasazení v řadě výzkumů (srov. např. Nisbet, Zelenski & Murphy, 2009; Nisbet, Zelenski & Murphy, 2011; Craig, Logan & Prescott, 2016), v nichž opakovaně prokázala schopnost měřit požadovaný konstrukt.

## 4 ZÁVĚR

Cílem článku bylo ukázat, jakým způsobem je možné zpracovat Likertovy škály v pedagogickém výzkumu na základě kvantitativní analýzy dat.

Význam adekvátního zpracování dat získaných pomocí Likertových škál podtrhují Ruskovy (2015) závěry, v nichž se konstatuje, že například v disertačních pracích zaměřených na chemii převládá kvantitativní analýza dat oproti kvalitativní a že nejčastěji využívaným výzkumným nástrojem je dotazník. Ke stejnému závěru došli také Žák (2015) a Pavlasová (2015). Žák (2016: s. 27) udává, že se jedná „zejména o různé konceptuální testy a dotazníky určené k měření postojů, příp. zjišťující zpětnou vazbu z realizované výuky“. Tento závěr není však nijak překvapivý, jelikož se jedná o metodu, která umožňuje snadný a rychlý sběr dat.

Konkrétní zpracování dat bylo představeno na škále Nature relatedness scale (Nisbet, Zelenski & Murphy, 2009), která se využívá k měření vztahu jedince k přírodnímu prostředí. Bylo podrobně popsáno, jakým způsobem lze zjišťovat reliabilitu a validitu této škály, která svým pojetím představuje standardní typ výzkumného nástroje, využívajícího ke sběru dat právě Likertovo škálování. Potvrdili jsme využitelnost tohoto nástroje v rámci zkoumané demografické skupiny (studenti Učitelství pro 1. st. ZŠ).

Ačkoli není možné prezentované informace jednoduše zobecnit, lze z nich však vyvodit následující sérii doporučení, jichž by se měli výzkumníci pracující s daty získanými pomocí Likertova škálování držet:

1. Nehodnotit jednotlivé položky, ale škálu jako celek. Jednotlivé položky zpravidla nemají dostatečnou vypovídající hodnotu o zkoumaném fenoménu.
2. Data získaná na základě bodu 1 analyzovat jako intervalové proměnné a zaměřit se na statistický parametr, popřípadě operace zmíněné v tab. 2 oddílu 2.1.
3. Ověřit rozdělení dat, což je nutné z důvodu volby vhodné statistické metody a také kvůli testování odlehlých hodnot, jelikož testy k tomuto určené (například Grubbsův test) předpokládají, že je známé rozdělení dat. Pokud jsou tyto

hodnoty detekovány, bývají zpravidla vyřazeny z matice dat a analyzovány samostatně. Zde je nutné zmínit, že autoři se neshodují v otázce automatického vyřazování těchto hodnot. Je vždy nutné zkontrolovat, zda se nejedná o hodnoty pro daná měření obvyklá, která by mohla nést důležité informace. Pokud jsou odlehlé hodnoty detekovány, není vhodné je automaticky odstranit, ale ověřit příčinu jejich vzniku. Jsou-li touto příčinou překlepy, selhání měřicí techniky atd., je možné data vyřadit z dalšího pozorování. Ve všech ostatních případech se výzkumník odstraněním těchto hodnot připravuje o důležité informace o jevech vyskytujících se s nízkou četností. Doporučujeme tedy tyto hodnoty izolovat<sup>13</sup> a vyhodnocovat samostatně.

4. Pokud nejsou známy informace o distribuci dat, je vhodné využít neparametrických statistických metod.
5. Ověření psychometrických vlastností nástroje jako je validita a reliabilita tak, jak je uvedeno v oddílech 3.3 a 3.4. Pokud nástroj nebude vykazovat dostatečné psychometrické vlastnosti, není možné jej použít.

Věříme, že respektování těchto pravidel před vlastní analýzou a interpretací získaných dat povede ke zvýšení kvality kvalifikačních prací i výzkumných statí zaměřených na kvantitativní analýzu dat získaných pomocí Likertova škálování.

## PODĚKOVÁNÍ

Autoři děkují oběma recenzentům za podnětné připomínky a náměty, jež pomohly významně zvýšit kvalitu celého textu. Příspěvek vznikl díky financování v rámci specifického výzkumu, projekt Studentské grantové soutěže UJEP-SGS-172-03-01.

## LITERATURA

- Albaum, G. (1997). The Likert scale revisited: An alternate version. *Journal of the Market Research Society*, 39, 331–349.
- Allen, E. & Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress*, 40, 64–65.
- Ary, D., Jacobs, L. C. & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). California: Thomson Wadsworth.
- Baggaley, A. & Hull, A. (1983). The effect of nonlinear transformations on a Likert scale. *Evaluation of the Health Professions*, 6, 483–491.
- Barnett, V. & Lewis, T. (1978). *Outliers in statistical data*. New York: Wiley.
- Ben-Gal, I. (2005). Outlier detection. In O. Maimon & L. Rokach (Eds.), *Data mining and knowledge discovery handbook: A complete guide for practitioners and researchers* (131–146). Kluwer Academic Publishers.
- Bendl, S. (2011). *Školní kázeň v teorii a praxi: učebnice pro studenty učitelství*. Praha: Triton.
- Blahuš, P. (2000). Statistická významnost proti vědecké průkaznosti výsledků výzkumu. *Česká kinantropologie*, 4(2), 53–72.

---

<sup>13</sup>Testování odlehlých hodnot, je možné se vyhnout využitím robustních statistických metod (těmto metodám se autoři v rámci článku nevěnují).

- Boone, H. N. & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension*, 50(2), 1–5.
- Brown, J. D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.
- Brown, J. D. (2011). Likert items and scales of measurement? *SHIKEN: JALT Testing & Evaluation SIG Newsletter*, 15(1), 10–14.
- 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. *Journal of Social Sciences*, 3(3), 106–116.
- Clason, D. L. & Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education*, 35(4), 31–35.
- Craig, J. M., Logan, A. C. & Prescott, S. L. (2016). Natural environments, nature relatedness and the ecological theater: connecting satellites and sequencing to shinrin-yoku. *Journal of Physiological Anthropology*, 35(1).
- Cronbach, L. J. (1951). Coefficient Alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.
- De Winter, J. C. F. & Dodou, D. (2010). Five-point Likert items: t test versus Mann-Whitney Wilcoxon. *Practical Assessment, Research & Evaluation*, 15(11). Dostupné z <http://pareonline.net/getvn.asp?v=15&n=11>
- Dunlap, R. E., Van Liere, K. D., Mertig, A. G. & Jones, R. E. (2000). Measuring endorsement of the New Ecological Paradigm: a revised NEP scale. *The Journal of Social Issues*, 56(3), 425–442.
- Franěk, M. (2012). Nature Relatedness Scale. Český překlad škály měřící spojení s přírodou. *Envigogika*, 7(1). Dostupné z <https://www.envigogika.cuni.cz/index.php/Envigogika/article/view/69/73>
- Gadermann, A. M., Guhn, M. & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, 17(3). Dostupné z <http://pareonline.net/getvn.asp?v=17&n=3>
- Garrido, L. E., Abad, J. J. & Ponsoda, V. (2013). A new look at Horn's parallel analysis with ordinal variables. *Psychological Methods*, 18(4), 454–474.
- Gavora, P. (2010). *Úvod do pedagogického výzkumu*. Brno: Paido.
- Gavora, P. (2013). Validita a reliabilita výskumných nástrojov: princípy a reálna prax. *Pedagogická orientace*, 23(4), 511–534.
- Hebák, P. (2013). *Statistické myšlení a nástroje analýzy dat*. Praha: Informatorium.
- Hendl, J. (2012). *Přehled statistických metod*. Praha: Portál.
- Hodge, D. R. & Gillespie, D. (2003). Phrase completions: An alternative to likert scales. *Social Work Research*, 27, 45–55.
- Chráška, M. (2016). *Metody pedagogického výzkumu: základy kvantitativního výzkumu*. 2. aktualizované vydání. Praha: Grada.
- Jakobsson, U. (2004). Statistical presentation and analysis of ordinal data in nursing research. *Scandinavian Journal of Caring Sciences*, 18, 437–440.

- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38, 1212–1218.
- Johnson, B. & Manoli, C. C. (2008). Using Bogner and Wiseman's Model of Ecological Values to measure the impact of an earth education programme on children's environmental perceptions. *Environmental Education Research*, 14(2), 115–127.
- Johnson, R. & Penny, J. (2005). Split-Half Reliability. In K. Kempf-Leonard (Ed.), *Encyclopedia of Social Measurement* (649–654). Elsevier.
- Joshi, A., Kale, S., Chandel, S. & Pal, D. K. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 7(4), 396–403.
- Kellert, S. & Wilson, E. O. (Eds.). (1993). *The Biophilia Hypothesis*. Washington: Shearwater Books.
- Kerlinger, F. N. (1972). *Základy výzkumu chování: pedagogický a psychologický výzkum*. Praha: Academia.
- Koniček, L., Malčík, M., Maťašeje, H. & Mazurová, V. (2007). *Hodnocení výsledků vzdělávání – praktická část*. Ostrava: Ostravská univerzita v Ostravě.
- Kroufek, R., Janovec, J. & Chytrý, V. (2015). Pre-service primary teachers and their attitudes towards nature. In O. Fleischmann, R. Seebauer, H. Zoglowek & M. Aleksandrovich (Eds.), *The Teaching profession: New Challenges — New Identities*. Wien: Lit Verlag GmbH & Co. KG.
- Kroufek, R. (2016). *Environmentální gramotnost studentů Učitelství pro 1. stupeň základní školy a možnosti jejího zjišťování* [Disertační práce]. Dostupné z [http://theses.cz/id/i5rowz/Kroufek\\_Roman\\_disertace.pdf](http://theses.cz/id/i5rowz/Kroufek_Roman_disertace.pdf)
- Kubiátko, M. (2016). Bol Likert ordinalista alebo intervalista? Chyby při tvorbe a vyhodnocovaní Likertových škál. *Pedagogika.sk*, 7(3), 177–190.
- Kuzon, W. M., Urbanchek, M. G. & McCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery*, 37, 265–272.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*, 22, 5–55.
- Madansky, A. (1988). *Prescriptions for Working Statisticians*. New York, Springer.
- Martinková, P. & Vlčková, K. (2014). Hodnocení reliability znalostních a psychologických testů. *Informační bulletin České statistické společnosti*, 4, 1–15.
- Martyn, P. & Brymer, E. (2014). The relationship between nature relatedness and anxiety. *Journal of Health Psychology*, 21(7), 1436–1445.
- Maurer, J. & Pierce, H. R. (1998). A comparison of Likert scale and traditional measures of self-efficacy. *Journal of Applied Psychology*, 83, 324–329.
- Mazalová, L. & Marečková, J. (2012). Types of Validity in the Research of NANDA International Components. *Profese on-line*, 5(1), 11–15. Dostupné z <http://profeseonline.upol.cz/en/pdfs/pol/2012/02/03.pdf>
- McDonald, R. P. (1978). Generalizability in factorable domains: “domain validity and generalizability”. *Educational and Psychological Measurement*, 38(1), 75–79.
- McLeod, S. A. (2008). *Likert scale*. Dostupné z <http://www.simplypsychology.org/likert-scale.html>
- Meloun, M. & Militký, J. (2002). *Kompendium statistického zpracování dat: metody a řešené úlohy*. Praha: Academia.

- Nisbet, E. K. (2011). *A nature relatedness intervention to promote happiness and environmental concern* [Disertační práce]. Ottawa: Carleton University.
- Nisbet, E. K., Zelenski, J. M. & Murphy, S. A. (2009). The Nature relatedness scale. Linking individuals' connection with nature to environmental concern and behavior. *Environment and Behavior*, 41(5), 715–740.
- Nisbet, E. K., Zelenski, J. M. & Murphy, S. A. (2011). Happiness is in our nature: Exploring nature relatedness as a contributor to subjective well-being. *Journall of Happiness Studies*, 12(2), 303–322.
- Nisbet, E. K. & Zelenski, J. M. (2013). The NR-6: a new brief measure of nature relatedness. *Frontiers in Psychology*, 4, 1–11.
- Novick, M. R. & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32(1), 1–13.
- Nun-Nally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Parker, L. & Lunney, M. (1998). Moving beyond content validation of nursing diagnoses. *Nursing Diagnosis*, 9(4), 144–150.
- Pavlasová, L. (2015). Disertační práce se zaměřením na didaktiku biologie, geologie a ekologie v České republice v letech 2004–2013. *Scientia in educatione*, 6(2), 4–15. Dostupné z <http://www.scied.cz/index.php/scied/article/view/234>
- Pelikán, J. (2011). *Základy empirického výzkumu pedagogických jevů*. Praha: Karolinum.
- Pell, G. (2005). Uses and misuses of Likert scales. *Medical Education*, 39(9), 970.
- Polit, D. F. & Beck, T. C. (2008). *Nursing Research*. Philadelphia: J. B. Lippincott Co.
- Půlpán, Z. (2004). *K problematice zpracování empirických šetření v humanitních vědách*. Praha: Academia.
- Revelle, W. & Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145–154.
- Richardson, M. W. & Kuder, G. F. (1939). The calculation of test reliability coefficients based upon the method of rational equivalence. *Journal of Educational Psychology*, 30(9), 681–687.
- Rod, A. (2012). Likertovo škálování. *E-Logos Electronic Journal for Philosophy*, 13, 2–14. Dostupné z <http://nb.vse.cz/kfil/elogos/science/rod12.pdf>
- Rusek, M. (2015). Analýza disertačních prací z didaktiky chemie obájených v České republice v letech 2003–2014. *Scientia in educatione*, 6(2), 16–34. Dostupné z <http://www.scied.cz/index.php/scied/article/view/233>
- Říčan, J. & Pešout, O. (2013, 16.–18. 9.). Výzkum kalibrace: nové perspektivy vědního objevování v pedagogické psychologii. In Sborník kol. autorů (Eds.), *Efektivita vzdělávání v proměnách společnosti* (311–317). Ústí nad Labem: PF UJEP.
- Sachs, L. (1984). *Applied Statistics, A Handbook of Techniques*. New York: Springer.
- Sarıçam, H., Şahin, S. H. & Soyucok, E. (2015). Doğayla ilişki olma ile dpresyon, anksiyete ve stres arasındaki ilişkinin incelenmesi. *Internation Journal of Psychiatry and Psychological Researches*, 4, 38–57.
- Sekaran, U. (1992). *Research methodsfor business:A skill building approach*. New York: John Wiley & Sons, Inc.

- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4), 591–601.
- Shoukri, M. M. & Edge, V. L. (1996). *Statistical methods for health sciences*. Michigan University: CRC Press.
- Stevens, S. S. (1959). Measurement, Psychophysics and Utility. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and Theories* (18–63). New York: John Wiley.
- Subedi, B. P. (2016). Using Likert type data in social science research: Confusion, issues and challenges. *International Journal of Contemporary Applied Sciences*, 3(2), 36–49.
- Šmejkal, P., Skoršepa, M., Stratilová Urválková, E. & Teplý, P. (2016). Chemické úlohy se školními měřicími systémy: motivační orientace žáků v badatelsky orientovaných úlohách. *Scientia in educatione*, 7(1), 29–48.  
Dostupné z <http://www.scied.cz/index.php/scied/article/view/280>
- Vácha, Z. & Ditrich, T. (2016). Efektivita badatelsky orientovaného vyučování na primárním stupni základních škol v přírodovědném vzdělávání v České republice s využitím prostředí školních zahrad. *Scientia in educatione*, 7(1), 65–79.  
Dostupné z <http://www.scied.cz/index.php/scied/article/view/293>
- Vickers, A. (1999). Comparison of an ordinal and a continuous outcome measure of muscle soreness. *International Journal of Technology Assessment in Health Care*, 15, 709–716.
- Weng, L. & Cheng, C. P. (2000). Effects of response order on Likert-type scales. *Educational and Psychological Measurement*, 60, 908–924.
- Wilson, E. O. (1984). *Biophilia: The human bond with other species*. Cambridge: Harvard University Press.
- Zvára, K. (2002) Měření reliability aneb bacha na Cronbacha, *Informační bulletin České statistické společnosti*, 13(2), 13–20.
- Žák, V. (2015). Disertační práce z didaktiky fyziky obhájené v České republice v letech 2004 až 2013 – přehled a analýza. *Scientia in educatione*, 6(2), 35–50.  
Dostupné z <http://www.scied.cz/index.php/scied/article/view/235>
- Žák, V. (2016). Metody sběru dat využívané didaktikou fyziky v mezinárodním prostředí. *Scientia in educatione*, 7(2), 18–33.  
Dostupné z <http://www.scied.cz/index.php/scied/article/view/349>

---

VLASTIMIL CHYTRÝ, [vlastimil.chytry@ujep.cz](mailto:vlastimil.chytry@ujep.cz)  
ROMAN KROUFEK, [kroufek@gmail.com](mailto:kroufek@gmail.com)  
Univerzita Jana Evangelisty Purkyně v Ústí nad Labem, Pedagogická fakulta  
Katedra preprimárního a primárního vzdělávání  
Pasteurova 1, 400 96, Ústí nad Labem, Česká republika