

Aplikace metody ukotvujících vinět v pedagogickém výzkumu: přehled literatury a metodologická doporučení^{1, 2}

Hana Voňková^{a, b}, Ondřej Papajoanu^b, Stanislav Bendl^b

^a Univerzita Karlova v Praze, Pedagogická fakulta, Ústav výzkumu a rozvoje vzdělávání

^b Univerzita Karlova v Praze, Pedagogická fakulta, Katedra pedagogiky

Redakci zasláno 15. 2. 2016 / upravená verze obdržena 5. 8. 2016 /
k uveřejnění přijato 17. 8. 2016

Abstrakt: Dotazníková šetření se (sebe)hodnotícími položkami jsou v pedagogickém výzkumu využívána často. Respondenti však mohou na tyto položky odpovídat rozdílně, tj. volit jiné škálové kategorie, přestože jejich objektivní situace je stejná. Metoda ukotvujících vinět nabízí řešení tohoto problému. Odpovědi na tyto položky koriguje o rozdílné užívání škály a činí je tak porovnatelné. Tento příspěvek si klade dva cíle. Zaprvé poskytnout přehled do této chvíle publikovaných prací z oblastí pedagogického výzkumu, ve kterých byla tato metoda aplikována. Zadruhé shrnout metodologická doporučení týkající se správné formulace a užití ukotvujících vinět. Metoda byla zatím úspěšně aplikována při měření nečestného chování žáků, ICT znalostí a dovedností, spokojenosti rodičů se školou, řízení třídy učitelem a podpory žáků ze strany učitele. Ukazuje se, že respondenti s různými charakteristikami skutečně užívají škálu rozdílným způsobem a že korigování dat s využitím metody ukotvujících vinět může pomoci vysvětlit některá paradoxní zjištění o vztahu vybraných proměnných. Naše metodologická doporučení pro využívání vinět se týkají podpory základních předpokladů metody, diskriminační síly vinět, počtu vinět, obsahové relevance a jazykové složky vinět. Na základě shrnutí dosavadního rozšíření a přínosu této metody pedagogickému výzkumu diskutují autoři možnosti jejího dalšího využití pro měření pedagogických konceptů.

Klíčová slova: metoda ukotvujících vinět, konzistence odpovědí, ekvivalence vinět, diskriminační síla vinět, počet vinět, formulace vinět

¹ Vznik článku byl umožněn na základě podpory projektu *The relationships between skills, schooling and labor market outcomes: A longitudinal study* (P402/12/G130), který je financován Grantovou agenturou České republiky.

² Poznámka redakce: Po poradě s autory rukopis nebyl anonymizován, neboť anonymizace by způsobila citelné oslabení argumentační i výkladové linie textu. Anonymita recenzentů vůči autorům zůstala zachována.

K měření pedagogických jevů se velmi často využívají dotazníková šetření se (sebe)hodnotícími otázkami s posuzovacími škálami. Odpovědi respondentů na (sebe)hodnotící otázky se následně používají k porovnání různých zemí, regionů, škol, tříd, socio-ekonomických skupin či jednotlivců. V literatuře se však objevují doklady o tom, že porovnání založené na (sebe)hodnocení může vést k zavádějícím závěrům o objektivním stavu, který je cílem měření. Respondenti totiž mohou ve své odpovědi reflektovat nejen objektivní stav, ale také svůj styl užívání škály. Například dva žáci se stejným chováním ve škole mohou hodnotit své chování odlišně – jeden jako výborné, jiný jen jako dobré (více viz Voňková, Bendl, & Papajoanu, 2016). V literatuře se tento jev označuje např. jako heterogenita ve stylu odpovídání na otázky či tzv. *differential item functioning* (DIF). Termín *differential item functioning* pochází z teorie odpovědí na položku (*item response theory*). Otázka v testu vykazuje DIF, pokud testovaní z různých skupin se stejnou schopností odpoví na otázku správně s různou pravděpodobností (Voňková, 2012).

Řešení heterogenity ve stylu odpovídání na položky nabízí metoda ukotvujících vinět. Jejím cílem je přímo změřit rozdílné užívání škály u respondentů a tyto identifikované rozdíly dále využít ke korekci (sebe)hodnocení. Je zřejmé, že mimo (sebe)hodnocení musí být ještě využita další informace, neboť samotné sebehodnocení je modelově složeno ze dvou částí, objektivní úrovně a stylu užívání škály. K získání této informace jsou využity právě ukotvující viněty, což jsou krátké příběhy popisující hypotetickou osobu v dané oblasti. Příklad ukotvující viněty vztahující se k celkové neukázněnosti žáka je následující: *Pavel je při hodině roztěkaný, nedává pozor, nikdy neví, co se při hodině děje, ruší, a tím zdržuje výklad učitele. Často používá sprostá slova. Občas je vůči učitelům drzý. Když se nudí, čmárá po la- vici. Párkrát už byl za školou, omluvenku si napsal sám.* (Voňková et al., 2016). Všichni respondenti jsou požádáni, aby mimo (sebe)hodnocení provedli i hodnocení ukotvující viněty. Vzhledem k tomu, že všichni respondenti dostávají k hodnocení stejnou ukotvující vinětu, můžeme rozdíly v jejich hodnocení interpretovat jako rozdílné užívání škály a následně tuto informaci využít k očištění sebehodnocení.

Ke statistickému zpracování dat se využívají dva základní přístupy práce s metodou ukotvujících vinět. Prvním je neparametrický přístup. Ten je založen na relativním porovnání (sebe)hodnocení s hodnocením situace popsané v ukotvujících vinětách. Například v případě tří odstupňovaných

vinět popisujících nízkou, střední a vysokou úroveň měřeného konceptu ($v1 < v2 < v3$) existuje sedm možností relativní pozice sebehodnocení s vůči vinětám: (1) $s < v1$, (2) $s = v1$, (3) $v1 < s < v2$, (4) $s = v2$, (5) $v2 < s < v3$, (6) $s = v3$ a (7) $v3 < s$. Korigované sebehodnocení pak nabývá hodnot 1–7 dle uvedené relativní pozice sebehodnocení. Více o neparametrickém přístupu viz Wand, King a Lau (2011) či v česky psané literatuře Voňková (2012). Druhým přístupem je parametrický model metody ukotvujících vinět, tzv. *CHOPIT model* (*Compound Hierarchical Ordered Probit Model*). Tento model je rozšířením běžně využívaného *probitového modelu pro ordinální proměnnou* (*ordered probit model*). Základní rozdíl spočívá v tom, že meze pro rozlišení jednotlivých kategorií ordinální škály mohou být v případě CHOPIT modelu variabilní v závislosti na pozorovaných (a případně i nepozorovaných) proměnných (Voňková, 2012). Více k matematické specifikaci modelu lze najít například u Kinga a kol. (2004) či v česky psané literatuře u Voňkové (2012, 2013).

Využití metody ukotvujících vinět je v některých sociálních vědách časté. Lze jmenovat například aplikace v oblasti možnosti občanů ovlivňovat politická rozhodnutí (např. King et al., 2004), zdraví (např. Bago d'Uva, O'Donnell, & van Doorslaer, 2008; Kok et al., 2012; Voňková & Hullegie, 2011), pracovní neschopnosti (např. Angelini, Cavapozzi, & Paccagnella, 2012; Kapteyn, Smith, & van Soest, 2007), životní spokojenosti (např. Angelini, Cavapozzi, Corazzini, et al., 2012) a spokojenosti se sociálními kontakty (Bonsang & van Soest, 2012). I přes slibné výsledky v ostatních sociálních vědách se tato metoda v pedagogice prozatím nepoužívá často. V budoucnu lze očekávat rozšíření aplikace této metody díky mezinárodnímu srovnávacímu šetření PISA, které zařadilo ukotvující viněty do žákovského dotazníku v roce 2012 i 2015.

Tento příspěvek si klade dva cíle. Prvním cílem je seznámit čtenáře s dosavadní aplikací metody ukotvujících vinět v pedagogickém výzkumu. Článek tím navazuje na předchozí publikovanou studii Voňkové (2012), která se zabývala možnostmi využití metody ukotvujících vinět v pedagogickém výzkumu.

Druhým cílem příspěvku je shrnout praktická doporučení pro formulaci ukotvujících vinět v pedagogice i jiných oborech. Dostupnost metodických instrukcí pro uživatele jakékoliv metody je nutnou podmínkou jejího širokého využití v praxi. Ani přes zmíněnou rozšířenost metody ukotvujících vinět v sociálních vědách taková doporučení v české ani světové literatuře zatím nebyla formulována.

1 Příklady aplikace metody ukotvujících vinět v pedagogice

Od jejího představení v roce 2004 Gary Kingem z Harvardské univerzity ve výzkumu možností obyvatel ovlivňovat politická rozhodnutí (King et al., 2004) se metoda ukotvujících vinět rozšířila do mnoha sociálních věd. V pedagogice je však stále spíše okrajově využívanou metodou. Teprve nedávno vznikla v České republice organizovaná skupina výzkumníků, která se soustředí na možnosti a limity při využití metody ukotvujících vinět v pedagogickém výzkumu.

V této části se pokusíme přehledně shrnout dosavadní výzkumné práce v pedagogice, které tuto metodu využívají. Nejprve čtenáři nabídneme krátkou deskripci každé z těchto publikací a základní informace o jejich autorech. Dále následuje tabulka 1 shrnující klíčová zjištění, která tyto práce přinesly. Pro studie využívající metodu ukotvujících vinět je charakteristická určitá společná struktura:

- 1) Výsledky před korekcí – výsledky měření sledovaného jevu (např. hodnocení vlastní neukázněnosti) u různých skupin respondentů (dívky versus chlapci, žáci s nízkým versus vysokým socio-ekonomickým statutem atd.) před očištěním dat o možné rozdílné užívání škály. Jedná se o výsledky, které jsou v případě rozdílného užívání škály respondentů zkrusleny a nevypovídají o objektivní situaci. Jsou typicky prezentovány pomocí deskriptivních statistik (tabulky frekvencí a sloupcové grafy) a statistických modelů, často se jedná o probitový model s ordinální proměnnou.
- 2) Heterogenita užívání škály – porovnání užívání škálových kategorií měřeného jevu různými skupinami respondentů. Cílem je identifikovat skupiny respondentů s relativně vysokým či nízkým standardem pro hodnocení. Respondenti s relativně vysokým standardem pro hodnocení se ve srovnání s ostatními respondenty při hodnocení určité úrovně měřeného jevu přiklánějí spíše k negativním kategoriím škály. V některých studiích jsou tyto respondenti označováni jako negativní při hodnocení. Naopak respondenti s nízkým standardem pro hodnocení jsou někdy označováni jako pozitivní při hodnocení. Identifikace heterogenity užívání škály je založena na hodnocení ukotvujících vinět a je typicky prezentována pomocí deskriptivních statistik (tabulky četností a sloupcové grafy hodnocení ukotvujících vinět) a statistických modelů, často se jedná o odhady mezí v parametrickém modelu metody ukotvujících vinět (tzv. *CHOPIT model*).

- 3) Výsledky po korekci – výsledky měření sledovaného jevu u různých skupin respondentů po očištění dat o rozdílné užívání škály. Zkreslení výsledků při porovnávání skupin je odstraněno. Je-li použit ke korekci parametrický model metody ukotvujících vinět, prezentují se výsledky pomocí odhadnutých parametrů CHOPIT modelu. Tento model je dále využit pro simulování rozložení korigovaného hodnocení různých skupin respondentů – užívání škály jedné ze skupin je zafixováno a na této fixní škále se vyjádří rozložení výsledků ostatních skupin. Výsledky všech skupin jsou tedy prezentovány tak, jakoby všechny skupiny využívaly škálu stejným způsobem jako vybraná skupina. Je-li použit ke korekci neparametrický přístup metody ukotvujících vinět, jsou korigované výsledky typicky prezentovány pomocí tabulek četností či sloupcových grafů, přičemž jsou porovnávány závěry korigovaného hodnocení v závislosti na metodě zpracování přesmyček a nekonzistencí hodnocení vinět (více např. Wand et al., 2011).
- 4) Validizace výsledků – výsledky před korekcí a po korekci jsou porovnávány s hodnotami externí proměnné. Externí proměnná buď měří objektivní úroveň daného konceptu (např. výsledky očního testu při porovnávání hodnocení vlastní ostrosti vidění různými respondenty), nebo jiný koncept, o kterém se předpokládá, že by měl úzce korelovat s měřeným konceptem ve studii (např. hodnocení učitelova řízení třídy před korekcí a po korekci je porovnáváno s výsledky žáků v předmětu, který je učitel učí).

Tabulka 1 tedy reflektuje tuto strukturu a přináší v přehledné formě hlavní výzkumná zjištění. Poznamenejme, že některé studie se nezabývají validizací výsledků. Prezentují se výsledky před korekcí, heterogenita užívání škály a výsledky po korekci. Navíc některé studie nemají výše uvedenou strukturu. Zabývají se například možnostmi využití metody ukotvujících vinět v pedagogice, popřípadě metodologickými aspekty užívání této metody. Tyto studie jsou uvedeny v tabulce 2.

Jednou z nejnověji publikovaných prací je odborný článek autorů H. Voňková, S. Bendl a O. Papajoanu z Pedagogické fakulty Univerzity Karlovy v Praze s názvem *How students report dishonest behavior in school: Self-assessment and anchoring vignettes*, který byl publikován v *The Journal of Experimental Education* v roce 2016. V této studii se autoři zaměřují na fenomén nečestného chování žáků druhého stupně základních škol. Autoři ukazují, jak

rozdílné užívání škály respondenty zkresluje výsledky dotazníkových šetření využívajících žákovské sebehodnocení nečestného chování ve škole. Článek též dokládá, že ukotvující viněty popisující jiné druhy kázeňských přestupků nemohou být použity pro očištění žákovského sebehodnocení nečestného chování, což s sebou nese mnohé implikace pro využívání ukotvujících vinět mimo jiné i v mezinárodním šetření PISA.

Další z nedávno publikovaných studií, která popisuje využití metody ukotvujících vinět v pedagogice, je příspěvek autorů J. He a F. J. R. Van de Vijvera z Tilburské univerzity v Nizozemsku nacházející se v souborném díle *The psychology of Asian learners: A festschrift in honor of David Watkins* z roku 2016. V tomto příspěvku s názvem *The motivation-achievement paradox in international educational achievement tests: Toward a better understanding* se autoři pokoušejí vysvětlit paradoxní zjištění týkající se čínských žáků, kteří sice vykazují dobré studijní výsledky, nicméně nízkou motivaci k učení. Autoři na datech z mezinárodního šetření PISA 2012 ukazují, že metoda ukotvujících vinět může pomoci reflektovat rozdílné užívání škály žáky z různých zemí. Způsob, jakým čínští žáci užívají škálu, je pak interpretován zejména s ohledem na konfuciánský hodnotový systém, který zdůrazňuje hodnotu vzdělání, skromnosti či sebekritičnosti.

H. Voňková a J. Hrabák (2015b), autoři z Pedagogické fakulty Univerzity Karlovy v Praze, publikovali v časopise *Computers & Education* odbornou studii *The (in)comparability of ICT knowledge and skill self-assessments among upper secondary school students: The use of anchoring vignette method*. Na vzorku českých žáků středních škol autoři ukazují, jak může rozdílné užívání škály zkreslit závěry o úrovni znalosti ICT, je-li využíváno pouze žákovské sebehodnocení. Na základě hodnocení vlastních znalostí ICT mohou výzkumníci dojít k paradoxním výsledkům, kdy by se zdálo, že žáci studující jiné než ICT obory (SOŠ s humanitním zaměřením) mají větší znalosti ICT než žáci ICT oborů (SOŠ se zaměřením na ICT). Autoři tyto rozdíly ve způsobu sebehodnocení a následně i prezentaci vlastních dovedností diskutují například v kontextu pracovního trhu, kde mohou hrát roli při hledání zaměstnání a výběru zaměstnanců, a navrhují užití metody ukotvujících vinět k dalšímu zkoumání v této oblasti.

Tabulka 1

Přehled hlavních zjištění výzkumných šetření aplikujících metodu ukotvujících vinět v pedagogice

| Autoři | Výsledky před korekcí | Heterogenita užívání škály | Výsledky po korekci | Validizace výsledků/ poznámky |
|------------------------------------|---|--|--|----------------------------------|
| Voňková, Bendl, a Papajoanu (2016) | Slabým prediktorem nečestného chování ve škole u žáků sedmých ročníků z Prahy 6 je vzdělání rodičů žáků. Pohlaví a známky jsou nevýznamné. | Vyšší standardy pro hodnocení nečestného chování mají žáci s charakteristikami: dívky, vyšší vzdělání rodičů, lepší studijní výsledky. | Nejsilnějším prediktorem nečestného chování jsou studijní výsledky žáků, následované vzděláním rodičů žáků. Slabším prediktorem je i pohlaví žáků. | |
| He a Van de Vijver (2016) | Korelace výsledků z matematiky s vnitřní/vnější motivací k učení se matematice ve výzkumu PISA 2012: Výsledky pro jednotlivé země Vnitřní: 0,04 až 0,38 Vnější: 0,02 až 0,38 Výsledky na úrovni zemí Vnitřní: -0,11 Vnější: -0,04 | | Korelace výsledků z matematiky s vnitřní/vnější motivací k učení se matematice ve výzkumu PISA 2012: Výsledky pro jednotlivé země Vnitřní: -0,13 až 0,42 Vnější: -0,11 až 0,43 Výsledky na úrovni zemí Vnitřní: -0,53 Vnější -0,52 | |
| Voňková a Hrabák (2015b) | Průměrné sebehodnocení znalostí ICT je u českých žáků ve srovnání s neICT žáky výrazně vyšší (rozdíl je statisticky signifikantní). | Žáci studující ICT mají vyšší standardy pro hodnocení vlastních znalostí informačních technologií než neICT žáci. | Průměrné sebehodnocení znalostí ICT je u českých žáků studujících ICT (ICT žáci) ve srovnání s žáky humanitně zaměřenými (neICT žáci) mírně nižší (nebyl nalezen statisticky signifikantní rozdíl). | |

Tabulka 1 (pokračování)

| Autoři | Výsledky před korekcí | Heterogenita užívání škály | Výsledky po korekci | Validizace výsledků/ poznámky |
|---|--|---|---|---|
| Voňková a Hrabák (2015a) | Průměrné sebehodnocení ICT znalostí je u dívek a žáků studujících ICT mírně nižší než u chlapců a neICT žáků (rozdíl je statisticky nesignifikantní). | Standardy hodnocení podle pohlaví a studijního oboru od nejvyšších po nejnižší: Chlapci ICT Dívky ICT Dívky neICT Chlapci neICT | Průměrné sebehodnocení ICT znalostí je u chlapců a žáků studujících ICT signifikantně vyšší než u dívek a neICT žáků. | Analýza dat ICILS 2013: Čeští žáci dopadli nejlépe v testu CIL, ale sebehodnocení ICT dovedností mají jedno z nejnižších. Dívky ve všech zemích dopadly v CIL, testu lépe než chlapci, avšak hodnotí své pokročilé znalosti hůře než chlapci. |
| Voňková, Zamarro, DeBerg, a Hitt (2015) | Pořadí vybraných PISA 2012 zemí dle žákovského hodnocení učitelova řízení třídy (od nejlepší úrovně po nejhorší úroveň učitelova řízení): Jordánsko 6. Rumunsko 11. Indonésie 18. Norsko 57. Dánsko 60. Nizozemsko 68. | Evropské severské země mají největší standardy pro hodnocení. Nejnižší standardy mají naopak země z jihozápadní Asie. | Pořadí vybraných PISA 2012 zemí dle žákovského hodnocení učitelova řízení třídy (od nejlepší úrovně po nejhorší úroveň učitelova řízení): Norsko 22. Dánsko 33. Nizozemsko 39. Rumunsko 60. Jordánsko 64. Indonésie 68. | (Ne)korigované hodnocení korelovalo s řadou externích proměnných jako je výsledek žáků v PISA 2012 testu matematické gramotnosti a HDP na obyvatele: Nekorigované hodnocení koreluje s výsledkem v testu matematické gramotnosti i s HDP záporně. Naopak uvedené korelace s korigovaným hodnocením jsou kladné. |

Tabulka 1 (pokračování)

| Autoři | Výsledky před korekcí | Heterogenita užívání škály | Výsledky po korekci | Validizace výsledků/ poznámky |
|----------------------------|--|---|---|----------------------------------|
| Kyllonen a Bertling (2013) | Korelace výsledků z matematiky s a) učitelovou podporou žáků, b) vztahem mezi učitelem a žákem, c) zájmem o matematiku ve výzkumu PISA 2012: <i>Průměrná korelace pro jednotlivé země:</i> a) 0,03 b) 0,05 c) 0,17 <i>Korelace na úrovni země</i> a) -0,45 b) -0,41 c) -0,50 | Způsoby užívání škály se mezi respondenty z různých zemí výrazně liší. | Korelace výsledků z matematiky s a) učitelovou podporou žáků, b) vztahem mezi učitelem a žákem, c) zájmem o matematiku ve výzkumu PISA 2012: <i>Průměrná korelace pro jednotlivé země:</i> a) 0,13 b) 0,14 c) 0,22 <i>Korelace na úrovni země</i> a) 0,29 b) 0,40 c) 0,14 | |
| Buckley & Schneider (2007) | Rodiče žáků chartrových škol hodnotí spolupráci rodič–učitel, rodič–rodič, rodič–ředitel lépe (jsou spokojenější) než rodiče žáků tradičních veřejných škol. | Volba chartrové školy je po korekci o ostatní proměnné (rasa, vzdělání apod.) nesignifikantní. Ostatní proměnné však signifikantní jsou. Rozdílné charakteristiky obou vzorků (rasa, vzdělání apod.) činí z rodičů dětí chartrových škol kritičtější skupinu. | Rodiče žáků chartrových škol jsou díky svým odlišným demografickým charakteristikám v průměru skutečně (po očištění o heterogenitu užívání škály) spokojenější než rodiče žáků tradičních veřejných škol. | |

H. Voňková a J. Hrabák (2015a) dále publikovali v časopise *Pedagogika* svůj článek s názvem *Řešení nesrovnatelnosti hodnocení ICT znalostí a dovedností žáků skrze ukotvující viněty*, ve kterém se zabývají srovnáváním sebehodnocení znalostí ICT. Autoři poukazují na rozdíly mezi žákovským sebehodnocením ICT znalostí a jejich skutečnými znalostmi zachycenými v mezinárodním výzkumném šetření ICILS 2013 a diskutují výsledky v kontextu vlastního výzkumného šetření. Autoři též nastiňují několik cest pro další výzkum, jako je zaměření na jednotlivé dimenze ICT znalostí zvláště, zkoumání žákovských znalostí netikety či zařazení ukotvujících vinět do dalších mezinárodních výzkumů ICT znalostí, jako je právě ICILS.

Autoři H. Voňková (PedF UK), G. Zamarro, V. DeBerg z University of Arkansas v USA a C. Hitt z Southern Illinois University School of Medicine v USA zveřejnili v roce 2015 svůj working paper s názvem *Comparisons of student perceptions of teacher's performance in the classroom: Using parametric anchoring vignette methods for improving comparability*, ve kterém provádějí sekundární analýzu dat z PISA 2012. Autoři pomocí vinět, které byly obsaženy v žákovském dotazníku PISA 2012, analyzují styly užívání škály jednotlivých zúčastněných zemí. Článek též ukazuje, jak může rozdílné užívání škály respondenty z různých zemí zkreslit mezinárodní komparaci a ovlivnit umístění jednotlivých států v mezinárodním žebříčku podle úrovně učitelova řízení třídy v daných zemích. Nakonec jsou analyzovány vztahy mezi úrovní učitelova řízení třídy v dané zemi a některými dalšími proměnnými, jako je například HDP na obyvatele či výsledky žáků dané země v matematice. Autoři na základě svých zjištění důrazně doporučují užívání metody ukotvujících vinět v jakémkoliv dotazníkovém šetření zaměřeném na zkoumání kvality učitelů. Ve svém dalším výzkumu se chtějí zaměřit na rozdíly v hodnocení kvality učitelů mezi školami s různým složením žáků v rámci jednotlivých států a vhodnost užití metody ukotvujících vinět při tomto zkoumání.

P. C. Kyllonen a J. B. Bertling (2013), vědečtí pracovníci pod Educational Testing Service s hlavním sídlem v Princetonu v New Jersey, přispěli do publikace *Handbook of international large-scale assessment* svojí statí s názvem *Innovative questionnaire assessment methods to increase cross-country comparability*. V části příspěvku věnované metodě ukotvujících vinět autoři ukazují na datech z výzkumného šetření PISA 2012 (viněty týkající se míry učitelovy podpory), jak mohou ukotvující viněty pomoci osvětlit zdánlivé anomálie ve výsledcích získaných v rámci tohoto šetření, v němž dochází až

ke změně znaménka korelačních koeficientů mezi sledovanými proměnnými. Autoři též diskutují možnost užití ukotvujících vinět zkonstruovaných pro určitý koncept pro očištění dat týkajících se jiných konceptů a tuto možnost empiricky dokládají očištěním jiných škál, než je míra učitelovy podpory. To by umožňovalo využití jedné sady vinět pro očištění široké škály sebehodnotících dat, což by vedlo k výraznému zkrácení komplexních dotazníkových šetření. Tato zjištění jsou však v rozporu s výsledky publikovanými Voňkovou a kol. (2016).

J. Buckley, v té době pracující pod Boston College, a M. Schneider ze Stony Brook University publikovali v roce 2007 svoji knihu *Charter schools: Hope or hype?*, která je prvním dílem, v němž byla metoda ukotvujících vinět aplikována v pedagogickém výzkumu. Autoři srovnávají hodnocení různých typů amerických škol (charterové a tradiční veřejné) rodiči, jejichž děti jeden z daných typů škol navštěvují. Otázka, na kterou se výzkumníci snaží odpovědět pomocí užití metody ukotvujících vinět, je, nakolik může rozdílné užívání škály (v tomto případě kritičnost rodičů, jejichž děti navštěvují charterové školy) ovlivnit srovnání těchto dvou typů škol. Autoři tak vrhají světlo i na problematiku tzv. *efektu růžových brýlí* (tj. psychologický jev, který spočívá v tom, že pokud si lidé vyberou nějaký program nebo službu, mají tendenci jej hodnotit pozitivně).

Tabulka 2

Heslovitý přehled stěžejních bodů teoretických a metodologických článků o metodě ukotvujících vinět v pedagogice

| Autoři | Stěžejní informace |
|-----------------------|---|
| OECD (2013) – Annex 6 | Představení vinět použitých v žákovském dotazníku PISA 2012; představení 12 indexů, které byly očištěny pomocí ukotvujících vinět; vysvětlení principu očištění dat a možných limitací při práci s nimi. |
| Voňková (2012) | Popis základního principu metody ukotvujících vinět; představení parametrického a neparametrického přístupu; diskuze nad možnostmi aplikace metody v pedagogice. |
| Buckley (2008) | Umístění sebehodnotící otázky vůči vinětám, řazení vinět v sestupném pořadí a fixní pořadí jednotlivých skupin vinět má vliv na odpovědi respondentů a může vést ke zkreslení dat. Randomizace pořadí jednotlivých vinět, pořadí skupin vinět a pozice vinět vůči sebehodnotící otázce by měla být užívána při generování dotazníků, aby se tomuto zkreslení dalo předejít. |

Mezi studie, které nemají typickou strukturu článku využívajícího ukotvující viněty, patří teoretické a metodologické publikace popsané v tabulce 2. Tyto publikace se zabývají převážně popisem základního principu této metody, popisem různých statistických přístupů či konkrétního zpracování dat, doporučeními pro správné užití ukotvujících vinět v dotaznících či možnostmi aplikace metody v pedagogickém výzkumu.

2 Doporučení pro formulaci ukotvujících vinět

Přestože metoda ukotvujících vinět má již mimo pedagogiku četné využití, přehled doporučení pro formulaci ukotvujících vinět ani v zahraniční literatuře neexistuje. Níže je uveden seznam klíčových doporučení, která buď odkazují na literaturu, nebo vyplývají z empirického šetření autorů. První dvě doporučení se týkají dodržení klíčových předpokladů metody ukotvujících vinět – tzv. *konzistence odpovědí (response consistency)* a *ekvivalence vinět (vignette equivalence)*. Další doporučení se vztahují k maximalizaci diskriminační síly vinět a s ní souvisejícímu vhodnému počtu vinět. Poté následují doporučení k obsahové relevanci a jazykové složce vinět.

2.1 Konzistence odpovědí (*response consistency*)

Jedná se o první z klíčových předpokladů metody ukotvujících vinět. Tento předpoklad odkazuje na to, že užívání škály v (sebe)hodnotících otázkách a ukotvujících vinětách je stejné. K porušení tohoto předpokladu může dojít, když se respondenti ve srovnání s hypotetickými osobami popsanými ve vinětách nadhodnocují či naopak podhodnocují. Při konstrukci vinět a jejich následném zadávání respondentům je nutné tento předpoklad podpořit.

Příklad takovéto podpory je popsán ve studii Voňkové a kol. (2016), která se zaměřuje na měření kázně žáků (více k celkovým výsledkům studie viz předchozí část o aplikacích metody v pedagogice). Žáci účastníci se výzkumu byli požádáni, aby hodnotili sami sebe stejně přísně jako žáky popsané ve vinětách, což je v podstatě samotný předpoklad. Dále dotazník použitý v tomto výzkumu byl anonymní a žákům bylo při sběru dat zdůrazněno, že data sebraná v tomto výzkumu budou využita pouze pro výzkumné účely a nebudou tudíž ukázána ani jejich učitelům ani řediteli. Třídní učitel byl při sběru dat přítomen (vyplňoval učitelský dotazník, v němž hodnotil chování jednotlivých žáků i ukotvující viněty), ale seděl v první lavici a byl požádán, aby během sběru dat nechodil třídou a nenahlížel do odpovědí žáků. Toto mělo

redukovat následující dva efekty, které by vedly k porušení předpokladu. Zaprvé, aby žáci nehodnotili sami sebe o mnoho lépe, než jak tomu ve skutečnosti je, a nechtěli učitelé ukázat, jak je jejich chování vzorné. Zadruhé, aby žáci nehodnotili hypotetické žáky ve vinětách hůře, než jak si myslí, že se skutečně chovají, tj. aby neudávali sociálně žádoucí odpověď (Fisher, 1993; Wei & Huang, 2005), že popsané chování ve vinětách je pro ně nepřijatelné a nemělo by být akceptováno. Jinými slovy, mělo být zabráněno tomu, aby žáci vlastní chování v tomto výzkumu nadhodnocovali a naopak chování žáků ve vinětách podhodnocovali.

Obecně může docházet i k opačnému efektu, tj. že respondenti sami sebe podhodnocují a naopak viněty nadhodnocují. King a kol. (2004) například uvádí, že k této situaci může dojít, pokud se respondenti nepovažují za až tak dobré či důležité jako osoby popsané ve vinětách. Měření různých konceptů může přinášet odlišné tendence respondentů, a je tudíž nutné zvážit specifika daného konceptu.

V literatuře existují statistické testy tohoto předpokladu (např. Au & Lorgelly, 2014; Bago d'Uva et al., 2011; Kapteyn et al., 2011; Peracchi & Rossetti, 2012). Je tedy možné otestovat, zdali je tento předpoklad u zformulovaných vinět dodržen, a pokud tomu tak není, viněty přeformulovat.

2.2 *Ekvivalence vinět (vignette equivalence)*

Jedná se o druhý z klíčových předpokladů metody ukotvujících vinět. Tento předpoklad spočívá v tom, že všichni respondenti interpretují viněty stejným způsobem (King et al., 2004).

Bonsang a van Soest (2012) uvádí, že k porušení tohoto předpokladu může dojít v případě, kdy respondenti ve vinětě postrádají nějakou informaci. Respondenti si pak mohou chybějící informaci doplnit dle vlastního uvážení či zkušeností a to může následně vést k systematickým rozdílům ve způsobu interpretace viněty. K porušení tohoto předpokladu může dle Bonsanga a van Soesta (2012) dojít i v případě, kdy je měřený koncept multidimenzionální a formulace viněty nepokryje komplexnost konceptu, což pak může vést k nežádoucímu doplnění informace jednotlivými respondenty dle jejich vlastního uvážení. Ve vinětě bychom sice mohli multidimenzionální koncept popsat podrobněji, nicméně je nutné mít na paměti, že délka viněty musí být pro respondenty únosná. Existuje zde tedy určitý kompromis mezi důrazem

na úplnost poskytnutých informací a délkou viněty. V pedagogickém výzkumu je zachování rozumné délky viněty zásadní a musí být přizpůsobeno věkové kategorii dětských respondentů.

Příklad podpory tohoto předpokladu je popsán ve Voňkové a kol. (2016). Autoři uvádějí, jakým způsobem se snaží podpořit splnění tohoto předpokladu při měření nekázně založeném na hodnocení vlastního chování a chování hypotetických žáků popsaném v ukotvujících vinětách. Žáci účastníci se výzkumu byli požádáni, aby si pozorně přečetli charakteristiku chování žáka popsanou ve vinětě a hodnotili toto chování pouze a jen na základě toho, co přečetli, a nedoplňovali do viněty žádnou další informaci.

Co se týká multidimenzionálních konceptů, je jedním z možných řešení dle webových stránek autora této metody Garyho Kinga (King, 2009) rozdělit koncept do jednotlivých dimenzí a formulovat viněty ke každé dimenzi zvlášť. Nicméně viněty se v praxi často formulují i pro multidimenzionální koncepty, jako je celková spokojenost se životem (Bonsang & van Soest, 2012). Pokud formulujeme vinětu k multidimenzionálnímu konceptu a v popisu viněty vynecháme jednu či více dimenzí, pak pro korektní aplikaci metody ukotvujících vinět musíme předpokládat, že užívání škály je ve vynechaných dimenzích srovnatelné s užíváním škály v dimenzích popsaných ve vinětě. K tomuto je ovšem zapotřebí důkladného odůvodnění. To může vyplývat například z předchozího empirického výzkumu, v němž se ukázalo, že užívání škály je u jednotlivých dimenzí měřeného konceptu srovnatelné u všech respondentů.

V literatuře existují statistické testy *vignette equivalence* (např. Bago d'Uva et al., 2011; Peracchi & Rossetti, 2012). Stejně jako u předpokladu konzistence odpovědí můžeme tedy předpoklad ekvivalence vinět testovat. Pokud je předpoklad ekvivalence vinět zamítnut, viněty přeformulujeme.

2.3 Diskriminační síla vinět

Zeptáme-li se náhodného výběru žáků na základní škole, zdali umí integrovat a derivovat, nepřinese to mnoho informace o jejich znalosti matematiky. Podobně je tomu u formulace vinět. Úroveň popsaného konceptu ve vinětě by měla odpovídat úrovni tohoto konceptu v populaci. Představme si, že budeme měřit nějaký koncept na desetibodové škále, většina sebehodnocení se bude pohybovat v rozmezí jedna až pět a zároveň většina hodnocení viněty

se bude pohybovat v rozmezí šest až deset. V takovémto případě nejsme schopni pomocí hodnocení vinět očistit užívání škály v sebehodnocení, většina sebehodnocení je pod úrovní hodnocení vinět.

V praxi je tedy vhodné provést předvýzkum, ve kterém zjistíme, jestli viněta není pro danou populaci irelevantní. Prvním indikátorem je, zdali distribuce hodnocení vinět odpovídá distribuci sebehodnocení. Pro získání prvního vhledu lze tedy sestavit jeden histogram s hodnocením vinět a druhý histogram se sebehodnocením a porovnat, zda nejsou histogramy příliš různé. Pokud na první pohled vidíme, že se například histogramy nepřekrývají, lze z toho vyvodit, že daná viněta má mizivou diskriminační sílu. Pro určení diskriminační síly vinět existují i přesné ukazatele, není nutné se opírat o první vhled získaný porovnáním dvou histogramů. Postup výpočtu se opírá o teorii entropie (King & Wand, 2007). Příklad výpočtu diskriminační síly vinět užitých v pedagogickém výzkumu lze najít ve Voňkové a Hrabákovi (2015b).

2.4 Počet vinět

Volba počtu vinět souvisí především s heterogenitou populace, na kterou se ve výzkumu zaměřujeme. Čím je populace heterogennější, tím více vinět bychom měli formulovat. Avšak nelze formulovat velké množství vinět, neboť tím by se dotazník a potažmo respondent zcela zahltil. Existuje zde tedy jistý kompromis mezi množstvím informací získaných pomocí vinět a počtem vinět.

V praxi lze postupovat tak, že v předvýzkumu sestavíme větší množství vinět, ověříme jejich diskriminační sílu a následně z nich vybereme ty, které nám přinášejí nejvíce informací (King & Wand, 2007). Pokud se při zpracování dat využívá neparametrický postup, je typické, že se pracuje s více než jednou vinětou – běžně se používají tři viněty. Neparametrický postup lze aplikovat i v případech, kdy pracujeme pouze s jednou vinětou, ale korigované sebehodnocení může mít pak teoreticky maximálně tři kategorie (sebehodnocení je nižší než / je na úrovni / je vyšší než hodnocení viněty). Parametrický přístup je často aplikován i v případech, kdy máme jednu vinětu. To je obecně ve všech statistických přístupech založeno na tom, že neparametrické přístupy vyžadují větší množství dat než parametrické, neboť se u nich na rozdíl od parametrických přístupů nedělají (či dělají v menší míře) dodatečné statistické předpoklady.

2.5 *Obsahová relevance a jazyková složka vinět*

Nyní následují některá další obecná doporučení, která jsou nezbytná pro širokou využitelnost zkonstruovaných vinět. Tato doporučení doprovázíme naším příkladem viněty týkající se žákovy pevné vůle. Viněta je určena žákům devátých ročníků základních škol a příslušných ročníků víceletých gymnázií:

Jana se na začátku října rozhodla, že se bude až do konce školního roku každý den půl hodiny učit matematiku, aby tak zlepšila své známky. Poté, co v listopadu dostala pětku z písemky, se na učení celý víkend ani nepodívala. V pondělí se však přeci jen začala opět pravidelně učit. V lednu dostala třída novou učitelku angličtiny, která žáky svou výukou zaujala. Jana se proto rozhodla, že se bude každý den věnovat poslechu písniček, čtení blogů nebo sledování videí v angličtině. Od té doby se Jana do konce školního roku učila matematiku dvakrát týdně, dobu učení zkrátila na čtvrt hodiny.

První se týká interkulturní relevance vinět. Pokud plánujeme viněty využívat k mezinárodnímu porovnávání, je nutné kromě kvalitního překladu zajistit i kulturně/národně univerzální obsah, kterému budou všichni respondenti rozumět a bude pro ně relevantní.

Jako příklad můžeme uvést výběr školních předmětů ve výše uvedené vinětě. Uvádíme zde dva školní předměty, přičemž u jednoho bylo potřeba více specifikovat, co žáka pro daný předmět nadchlo. Bylo tedy nezbytné vybrat takové předměty a jejich obsah, které by mohly být univerzálně použity při mezinárodním srovnávání. Pokud předem víme, na jaké země chceme svými vinětami cílit, můžeme pomocí analýzy vzdělávacích dokumentů identifikovat společné studijní obsahy a jim obsah vinět přizpůsobit. Pokud srovnávané země nemáme předem určené a chceme viněty formulovat co nejuniverzálněji, jako tomu bylo v našem případě, je nutné volit obezřetně. My jsme pro naše viněty zvolili matematiku a anglický jazyk, jejichž výuka je napříč různými školami a zeměmi poměrně rozšířená a existuje zde předpoklad, že v mnoha zemích tyto předměty budou pro žáky relevantní.

Viněty musí být relevantní pro danou cílovou skupinu (hraje zde roli věk cílové skupiny, obvyklé zájmy jejích členů apod.). Ve výše zmiňované vinětě týkající se žákovy pevné vůle byl v jedné z jejích vývojových fází zakomponován přírodopis jakožto druhý předmět k matematice. Popisovaného žáka tolik zaujala výuka geologie, že se rozhodl chodit do přírody a sbírat různé typy hornin. Ačkoliv i dnes mezi mladou populací jistě existují nadšenci do geologie

a daný popis by dobře zapadal do charakteristiky měřeného konceptu, relevance tohoto příběhu je pro většinu cílové populace nízká. Respondenti by měli problém uvěřit pravdivosti popisovaného příběhu a ztotožnit se s jejím protagonistou. Proto byla tato verze viněty přepracována, aby lépe odpovídala své cílové skupině.

Problém s relevancí může nastat především v případě, kdy se výzkumník (věkově vzdálený dětské populaci) nedokáže dostatečně vžít do dané cílové skupiny a formulované viněty se budou žákům zdát cizí a nereálné. Podobná situace se může objevit, má-li výzkumník, který vinětu tvoří, sám velmi vysokou či naopak velmi nízkou úroveň měřeného konceptu a z toho plynoucí zkreslené vnímání. Může se poté stát, že viněta, o které výzkumník věří, že reprezentuje nízkou úroveň měřeného konceptu (podle jeho standardů), pro žakovskou populaci reprezentuje úroveň vysokou. Další dvě (vyšší) viněty jsou pak pro většinu žáků zcela neuvěřitelné a nereálné.

Je však nutné dát pozor, abychom se kvůli zvyšování relevance vinět pro cílovou skupinu neodchýlili od původního konceptu. Namísto chození do přírody a sbírání hornin v pracovní verzi viněty bylo zvažováno, že by se žák přestal učit matematiku kvůli tomu, že v televizi začne běžet nový seriál, který žáka velmi zaujme. Toto je jistě velmi relevantní pro naši cílovou skupinu, avšak již se nejedná o aktivitu stejné jakosti (jde spíše o volnočasovou aktivitu). Takto formulovaná viněta by tedy již neodpovídala ani tak žakově pevně vůli (kde se snažíme vystihnout tendenci k nahrazení jednoho cíle cílem druhým, a to při zachování návaznosti cílů na školní prostředí), jako spíše sebekontroly (u které jde o odolávání pokušení, což jistě odpovídá rozhodování, zdali se učit, či se dívat na seriál) – to je však již rozdílný koncept.

Viněty by měly být *konkrétní*. Pokud je to možné, je cílem redukovat počet slov, která lze nahradit specifickými slovy vyjadřujícími objektivní míru (např. místo *dělá danou činnost často* užít *dělá danou činnost hodinu denně*). Srovnajte například pracovní formulaci *od té doby se Jana sice učí matematiku, ale nepravidelně a u učení nevydrží dlouho* s formulací finální *od té doby se Jana do konce školního roku učila matematiku dvakrát týdně, dobu učení zkrátila na čtvrt hodiny*. Vidíme, že úroveň konkrétnosti je zcela rozdílná a prostor pro rozdílnou interpretaci se výrazně zmenšil.

Ve vinětách by se též neměly objevovat hodnoticí adjektiva, která mohou respondenty ovlivňovat při jejich hodnocení. Například není vhodné psát *od té doby se Jana do konce školního roku poctivě učila matematiku dvakrát týdně*, ale spíše neutrální variantu *od té doby se Jana do konce školního roku učila matematiku dvakrát týdně*.

Poznamenejme, že na viněty jako na jedny z dotazníkových položek se vztahují i další běžná doporučení pro formulaci položek v dotazníku, jako jsou vyhnutí se sugestivnímu obsahu a důraz na jasnost a srozumitelnost.

Při formulaci vinět je nutné mít na paměti jejich primární cíl – identifikaci heterogenity užívání škály v (sebe)hodnoticích otázkách. Ta je pak následně využita pro korekci sebehodnocení, ale to už je technická záležitost (aplikujeme neparametrický či parametrický přístup a sebehodnocení korigujeme). Z obecného pohledu lze konstatovat, že viněta nemusí pokrývat celý koncept, nemusí být v tomto smyslu univerzální. Viněta může pokrývat jen určitou část konceptu, např. jednu dimenzi v případě multidimenzionálních konceptů. Pokud se díky krátké vinětě podaří identifikovat heterogenitu užívání škály při dodržení všech v předchozích částech diskutovaných předpokladů (konzistence odpovědí, ekvivalence vinět a diskriminační síla vinět), je to optimální stav. Dodržení všech uvedených předpokladů není však vždy snadné. Jedná se v podstatě o limity metody při její aplikaci na měření určitých jevů. Vše vyžaduje obezřetnost při formulování vinět a sběru dat. Statistické testování předpokladů uvedené v literatuře pomáhá zjistit, zdali jsme předpoklady splnili a dosáhli tak žádoucího stavu. Z pohledu univerzálnosti je vhodné se zamýšlet nad tím, zda respondent užívá škálu stejným způsobem při hodnocení různých konceptů. Lze si klást otázku, zda jsou respondenti stejně přísní při hodnocení různých konceptů, či zda existuje alespoň nějaká množina konceptů, u nichž by byli respondenti v tomto smyslu jednotní. Pokud by tomu tak bylo, stačilo by v podstatě zformulovat vinětu(y) k jednomu konceptu, a pak tuto vinětu / tyto viněty využít i pro očištění odlišného užívání škály u dalších konceptů. Literatura však ukazuje, že přenositelnost určitého užívání škály napříč různými jevy bude nutné ještě podrobně zkoumat. Stejně užívání škály totiž nemusí být zachováno nejen napříč různými jevy, ale také v rámci měření jednotlivých dimenzí jednoho jevu (Voňková et al., 2016).

Závěrem

Metoda ukotvujících vinět přináší do pedagogického výzkumu nové možnosti zkoumání způsobu, kterým respondenti hodnotí jevy a následně i míru, do jaké spolu různé jevy vzájemně souvisejí. Výzkumy dokazují, že žáci s různým pohlavím, studijními výsledky, či sociokulturním zázemím se mohou lišit ve způsobu hodnocení různých pedagogických konceptů. Voňková a kol. (2016) ukázali, že žáci s různým pohlavím, vzděláním rodičů a různými studijními výsledky hodnotí své nečestné chování různě přísně, stejně tak jako Voňková a Hrabák (2015a) zjistili, že žáci s různým pohlavím a studijním oborem mají různé standardy pro hodnocení svých ICT znalostí a dovedností. Voňková a kol. (2015) na základě sekundární analýzy dat z PISA 2012 rozdělili respondenty z různých zemí podle jejich tendence užívat různé škálové kategorie (země s tendencí užívat krajní body škály; prostřední body škály; nižší body škály; vyšší body škály; a země s nevyhraněným užíváním škály).

Pokud data očistíme o rozdílné užívání škály respondenty, můžeme získat přesnější informace například o tom, které charakteristiky respondentů jsou prediktory měřeného chování – jako např. ve Voňkové a kol. (2016). Autoři před očištěním dat o rozdílné užívání škály identifikovali pouze jediný signifikantní prediktor žákovského nečestného chování (vzdělání rodičů), po očištění dat se však ukázalo, že signifikantním prediktorem je též pohlaví žáka a vzdělání jeho rodičů. Po očištění dat se též může stát, že korelace mezi sledovanými proměnnými změní znaménko, a tím se mohou vysvětlit některá paradoxní zjištění o jejich vztahu před očištěním dat. Například Kyllonen a Bertling (2013) při sekundární analýze dat z PISA 2012 uvádějí, že před očištěním dat byla korelace (na úrovni zemí, $N = 63$) mezi škálou učitelovy podpory žáka a studijními výsledky žáků $-0,45$. Toto je zářející výsledek. Lze předpokládat, že tam, kde se žákům dostává větší podpory ze strany učitele, budou žáci dosahovat lepších výsledků. Po očištění dat o rozdílné užívání škály se tato korelace změní na $0,29$, což více odpovídá předpokládané podobě vztahu. Podobně Voňková a kol. (2015) měřili korelaci mezi úrovní řízení třídy učitelů matematiky a studijními výsledky žáků v matematice. Před korekcí o rozdílné užívání škály byl výsledek korelace $-0,33$, což je paradoxní výsledek. Po očištění dat se korelace změnila na $0,45$, což je výsledek, který více odpovídá předpokládanému vztahu těchto dvou proměnných.

Aby bylo možné metodu ukotvujících vinět správně aplikovat, je nezbytně nutné, aby byly dodrženy její dva základní předpoklady. K podpoře

konzistence odpovědí doporučujeme používat anonymní dotazníky. Dále je nutné se pokusit zajistit, aby respondenti neměli důvod se hodnotit lépe či naopak hůře než osobu ve vině. K podpoře předpokladu ekvivalence vině v případě multidimenzionálních konceptů doporučujeme formulovat viněty pro každou dimenzi zvlášť. Pokud například v předvýzkumu zjistíme, že heterogenita ve stylu odpovídání je totožná pro různé dimenze konceptu, můžeme dále použít viněty popisující jen jednu z dimenzí a nesbírat data k vinětám popisujícím další dimenze. Pro předpoklad ekvivalence vině je důležité, aby respondenti hodnotili viněty pouze na základě toho, co je v nich napsáno, a nedomýšleli si žádné informace. Diskriminační sílu viněty můžeme zvýšit tím, že formulujeme vinětu tak, aby úroveň konceptu ve vinětě více odpovídala úrovni v populaci. Počet viněty závisí především na heterogenitě populace, na kterou se ve výzkumu zaměřujeme. Obecně platí, že čím je populace heterogennější, tím více viněty bychom měli formulovat. Nicméně je nutno brát v potaz i celkovou délku dotazníku a únavu respondentů. Naše doporučení pro formulaci viněty se týkají i jejich obsahové relevance – je nutné brát v úvahu jednak interkulturní rozdíly při mezinárodním porovnávání, jednak specifika cílové skupiny, jako je věk respondentů či jejich zájmy. Dále doporučujeme vyhýbat se vágním termínům a raději užívat konkrétní výrazy, zvláště pokud jde o frekvenci různých jevů. Též není vhodné do viněty vkládat hodnotící adjektiva, která mohou ovlivnit hodnocení respondentů.

Výše uvedený přehled výzkumných prací v pedagogice, ve kterých byla metoda ukotvujících viněty použita, ukazuje nejen ty oblasti, kam již metoda pronikla, ale i velké množství oblastí pedagogického výzkumu, v nichž její aplikace zatím zcela chybí. I na základě rostoucího počtu aplikací této metody v jiných sociálních vědách je možno předpokládat, že její potenciál pro zvýšení validity měření je výrazný. Všude tam, kde jsou respondenti dotazováni otázkami s hodnotícími škálami a hodnotí přitom sami sebe či druhé, hrozí zkreslení subjektivním využíváním škály. Doporučujeme proto využívání této metody ve všech takových oblastech pedagogického výzkumu. Dosavadní aplikace poukazují na její velký potenciál.

Literatura

- Angelini, V., Cavapozzi, D., Corazzini, L., & Paccagnella, O. (2012). Age, health and life satisfaction among older Europeans. *Social Indicators Research*, 105(2), 293–308.
- Angelini, V., Cavapozzi, D., & Paccagnella, O. (2012). Cross-country differentials in work disability reporting among older Europeans. *Social Indicators Research*, 105(2), 211–226.

- Au, N., & Lorgelly, P. K. (2014). Anchoring vignettes for health comparisons: An analysis of response consistency. *Quality of Life Research*, 23(6), 1721–1731.
- Bago d'Uva, T., Lindeboom, M., O'Donnell, O., & van Doorslaer, E. (2011). Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources*, 46(4), 872–903.
- Bago d'Uva, T., O'Donnell, O., & van Doorslaer, E. (2008). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*, 37(6), 1375–1383.
- Bonsang, E., & van Soest, A. (2012). Satisfaction with social contacts of older Europeans. *Social Indicators Research*, 105(2), 273–292.
- Buckley, J. (2008). *Survey context effects in anchoring vignettes*. Dostupné z <http://www.polmeth.wustl.edu/media/Paper/surveyartifacts.pdf>
- Buckley, J., & Schneider, M. (2007). *Charter schools: Hope or hype?* Princeton: Princeton University Press.
- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2), 303–315.
- He, J., & Van de Vijver, F. J. R. (2016). The motivation-achievement paradox in international educational achievement tests: Toward a better understanding. In R. B. King & A. B. I. Bernardo (Eds.), *The psychology of Asian learners: A festschrift in honor of David Watkins* (s. 253–268). Singapore: Springer Science.
- Kapteyn, A., Smith, J. P., & van Soest, A. (2007). Vignettes and self-reports of work disability in the US and the Netherlands. *American Economic Review*, 97(1), 461–473.
- Kapteyn, A., Smith, J. P., van Soest, A., & Voňková, H. (2011). *Anchoring vignettes and response consistency* (RAND Working Paper No. WR-840). Dostupné z http://www.rand.org/content/dam/rand/pubs/working_papers/2011/RAND_WR840.pdf
- King, G. (2009). *Anchoring vignettes: Frequently asked questions*. Dostupné z <http://gking.harvard.edu/files/gking/files/vfaq.pdf>
- King, G., Murray, C., Salomon, J., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 567–583.
- King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, 15, 46–66.
- Kok, R., Avendano, M., Bago d'Uva, T., & Mackenbach, J. (2012). Can reporting heterogeneity explain differences in depressive symptoms across Europe? *Social Indicators Research*, 105(2), 191–210.
- Kyllonen, P. C., & Bertling, J. (2013). Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *A handbook of international large-scale assessment data analysis: Background, technical issues, and methods of data analysis* (s. 277–286). London: Chapman Hall/CRC Press.
- Organisation for Economic Cooperation and Development (OECD). (2013). *PISA 2012 results: What makes a school successful? Resources, policies and practices (Volume IV). Annex A6: Anchoring vignettes in PISA 2012 student questionnaire*. Dostupné z <https://www.oecd.org/pisa/keyfindings/PISA-2012-results-Annex%20A6-VolIII-VolIV.pdf>

- Peracchi F., & Rossetti C. (2012). Heterogeneity in health responses and anchoring vignettes. *Empirical Economics*, 42(2), 513–538.
- Voňková, H. (2012). Metoda ukotvujících vinět a možnosti využití v pedagogice. *Orbis scholae*, 6(1), 27–40.
- Voňková, H. (2013). Subjektivní hodnocení problémů s pohybem: užití parametrického modelu metody ukotvujících vinět. *Orbis scholae*, 7(1), 49–66.
- Voňková, H., Bendl, S., & Papajoanu, O. (2016). How students report dishonest behavior in school: Self-assessment and anchoring vignettes. *The Journal of Experimental Education (advance online publication)*.
- Voňková, H., & Hrabák, J. (2015a). Řešení nesrovnatelnosti hodnocení ICT znalostí a dovedností žáků skrze ukotvující viněty. *Pedagogika*, 65(3), 274–291.
- Voňková, H., & Hrabák, J. (2015b). The (in) comparability of ICT knowledge and skill self-assessments among upper secondary school students: The use of the anchoring vignette method. *Computers & Education*, 85, 191–202.
- Voňková, H., & Hulleger, P. (2011). Is the anchoring vignettes method sensitive to the domain and choice of the vignette? *Journal of the Royal Statistical Society: Series A*, 174(3), 597–620.
- Voňková, H., Zamarro, G., DeBerg, V., & Hitt, C. (2015). *Comparisons of student perceptions of teacher's performance in the classroom: Using parametric anchoring vignette methods for improving comparability* (EDRE Working Paper No. 2015-01). Dostupné z <http://www.uaedreform.org/downloads/2015/05/comparisons-of-student-perceptions-of-teachers-performance-in-the-classroom-using-parametric-anchoring-vignette-methods-for-improving-comparability.pdf>
- Wand, J., King, G., & Lau, O. (2011). Anchors: Software for anchoring vignette data. *Journal of Statistical Software*, 42(3), 1–25.
- Wei, H. S., & Huang, C. K. (2005). Reviewing school bullying research: Empirical findings and methodical considerations. *NTTU Educational Research Journal*, 16(1), 69–112.

Autoři

PhDr. RNDr. Hana Voňková, PhD. et. PhD., Univerzita Karlova v Praze, Pedagogická fakulta, Ústav výzkumu a rozvoje vzdělávání / Katedra pedagogiky, Myslíkova 7, 110 00 Praha 1, e-mail: h.vonkova@gmail.com

Mgr. Ondřej Papajoanu, Univerzita Karlova v Praze, Pedagogická fakulta, Katedra pedagogiky, Magdalény Rettigové 4, 110 00 Praha 1, e-mail: ondrej.papajoanu@pedf.cuni.cz

Prof. PaedDr. Stanislav Bendl, PhD., Univerzita Karlova v Praze, Pedagogická fakulta, Katedra pedagogiky, Magdalény Rettigové 4, 110 00 Praha 1, e-mail: stanislav.bendl@pedf.cuni.cz

Applying the anchoring vignette method in educational research: Literature review and methodological recommendations

Abstract: Questionnaire surveys with (self-)assessment items have been frequently employed in educational research. Respondents can, however, assess the same situations differently, i.e. choose different scale categories. The anchoring vignette method provides a solution to this problem. It helps adjust (self-)assessment for heterogeneity in scale usage and make respondents' answers comparable. This paper has two aims. First, to provide a review of works in the area of educational research in which the method has been used. Second, to summarize methodological recommendations concerning the development and correct use of anchoring vignettes. The method has been successfully used for measuring students' dishonest behavior, ICT knowledge and skills, parents' satisfaction with a school, teacher's classroom management and teacher's support. The results suggest that respondents with different characteristics indeed use the scale differently and that the anchoring vignette method can help explain some paradoxical findings concerning relationships between variables. Our methodological recommendations for the use of the method focus on the support of its key assumptions, number and discriminatory power of vignettes, relevancy of their content, and some linguistic aspects of their formulation. Summarizing current applications and contributions of the method to educational research, the authors discuss the possibilities of its future utilization for the measurement of educational concepts.

Keywords: anchoring vignette method, response consistency, vignette equivalence, discriminatory power, number of vignettes, vignette formulation

Poláchová Vašátková J., & Rudnicki P. (2016). *Proměny české a polské pohraniční školy poskytující povinné vzdělávání po roce 1989*. Olomouc: Univerzita Palackého v Olomouci.

Kniha vznikla jako výstup z projektu řešeného v letech 2013–2014 na Univerzitě Palackého v Olomouci ve spolupráci s polským partnerem, jímž byla Dolnośląska Szkoła Wyższa ve Wroclavi. Záměrem projektu bylo analyzovat proměny českého a polského školství zajišťující povinné vzdělávání po roce 1989 se zaměřením na konkrétní regiony (Olomoucký a Moravskoslezský kraj a Dolnoslezské vojvodství). Zkoumány byly zejména funkce školy, s akcentem na kvalifikační, socializačně-integrační, individualizační, kulturní a ekologickou funkci. Smíšený výzkumný design koncipovaný komparativně se opíral o sběr dat s využitím různých metod: analýzy dokumentů, dotazníku a vícepřípadové studie. Specifikem projektu byla role výzkumníků v pozici kulturních cizinců: Čechů v Polsku a Poláků v České republice.