

Part 6

Research Education & Research Practice

Ondrej Papajoanu, Katerina Kralova, Hana Vonkova & Jiri Stipek

Examining the Assessment of Anchoring Vignettes in Different Information and Communication Technology Competence Domains: The Results of a Pilot Study Among Upper-Secondary Students

Abstract

The anchoring vignette method has been proposed as an innovative approach to solve the problem with self-assessment data incomparability caused by the differences in scale usage between respondents. In this study, we use a set of 15 anchoring vignettes describing different domains of ICT skills (Information, Communication, Content creation, Safety, and Problem solving) to examine the differences in scale usage between Czech upper-secondary students (N = 166) studying at different types of schools in these domains of ICT competence. Our preliminary results suggest that students studying at different types of schools seem to have different standards for vignette evaluation (i.e. they use the scale differently), however, we also identified certain similarities in the way students from a particular school use the scale across the five ICT competence domains. Such findings might be of high relevance for the further use of the anchoring vignette method in ICT competence research, however, further investigation in this research area is necessary.

Keywords: differences in scale usage, anchoring vignettes, self-assessment, ICT skills, DigComp framework

Introduction

In educational research, student self-assessment questionnaires containing items with rating scales are a frequently used approach of data collection. Such data are then commonly used to compare different groups of students based on their characteristics such as gender, ethnicity, or socioeconomic status. However, there has been a long-term concern about the comparability of such data, which might be hindered by the differences in the way respondents use scales to assess various concepts (Vonkova, Papajoanu & Bendl, 2016).

For example, it might be the case that two students with the same level of particular concept (e.g. ICT skills) assess their skills on a scale: (1) *excellent*, (2) *very good*, (3) *good*, (4) *poor*, (5) *very poor* differently – one student as *excellent*, the other student only as *very good* (Vonkova & Hrabak, 2015). These differences

might lead researchers to erroneous conclusions about the actual level of the measured concept, making the comparison between students inaccurate.

In the literature, some paradoxical findings have been documented both at the international level (He & van de Vijver, 2016; Kyllonen & Bertling, 2013; Vonkova, Zamorro & Hitt, 2018) and intra-country level (Vonkova & Hrabak, 2015; West et al., 2016). For example, using PISA 2012 country-aggregated data, He and van de Vijver (2016) found a negative relationship between student mathematic achievement and their motivation to learn mathematics. These striking results might be attributable to the differences in which respondents use scales.

The anchoring vignette method

Several methodological approaches have been proposed to solve the problem with scale usage differences among respondents, such as the anchoring vignette method (King et al., 2004), the identification of response styles irrespective of the item content (Baumgartner & Steenkamp, 2001; He & van de Vijver, 2016), or the overclaiming technique (Paulhus et al., 2003; Vonkova, Papajoanu & Stipek, 2018). The anchoring vignette method, the approach that we use in this paper, was first introduced by King et al. (2004) in research on political efficacy.

The basic idea of the method is that we let respondents not only assess themselves but also a set of anchoring vignettes, which are short stories describing certain levels of the measured concept (King et al., 2004). Since all the respondents assess the same anchoring vignettes, the differences in their assessments might be interpreted as the differences in which they use the scale. This information can be then used to adjust their self-assessments, leading to better data comparability.

Since its introduction, the method has been successfully applied in many research areas such as health (Bago d'Uva, O'Donnell & van Doorslaer, 2008), personality (Weiss & Roberts, 2018), life satisfaction (Angelini et al., 2012), or satisfaction with social contacts (Bonsang & van Soest, 2012). In educational research, there is a growing number of studies using the method in domains such as school discipline (Vonkova, Bendl & Papajoanu, 2017) or ICT knowledge and skills (Vonkova & Hrabak, 2015). The method was also included in Programme for International Student Assessment (PISA) studies in 2012 and 2015 and several secondary analyses of PISA vignette data have been published (e.g. He & van de Vijver, 2016; Vonkova, Zamorro & Hitt, 2018).

Application of the anchoring vignette method in the area of ICT knowledge and skills measurement

Technologies affect people's lives in many areas being it at school, in the workplace, or in the community and the development of ICT skills have become a major part of people's education (Fraillon et al., 2014). These developments warrant that we accurately measure these skills across different groups of students according to their characteristics such as gender, cultural, or socioeconomic background. Self-assessment measures of ICT skills have been widely reported and interpreted as a proxy measure of actual ICT literacy (Siddiq et al., 2016). Their main advantage in comparison to performance tests might be low monetary costs and time requirements of their administration, making them an appropriate tool for large-scale data collection.

The previous study by Vonkova and Hrabak (2015), however, documented that the comparison of different groups of students based solely on their self-assessment of ICT skills might lead to contra-intuitive results, suggesting that scale usage differences between respondents are an important issue in this area of research. They found that students studying ICT reported a lower level (though non-significantly) of ICT skills than students studying pedagogy and business (non-ICT students). After the adjustment of students' self-assessment of ICT skills using the anchoring vignette method, the results were the opposite – ICT students reporting a higher level of ICT skills than non-ICT students. These intuitively more compelling results suggest that the anchoring vignette method is a promising approach in this research area that needs to be further examined.

Our study builds on the previous literature in several ways. In their study, Vonkova and Hrabak (2015) used only one set of anchoring vignettes describing general ICT skills that were based primarily on national curriculum. In this study, we use five sets of anchoring vignettes based on the five areas of ICT competencies as defined in the DigComp framework – a framework for developing and understanding digital competence in Europe: Information, Communication, Content creation, Safety, and Problem solving (Ferrari, 2013). This more detailed approach allows to identify the differences in scale usage for the particular areas of ICT skills.

In PISA 2012, anchoring vignettes for teacher's classroom management were used to compute twelve adjusted indices ranging from attitude towards school to instrumental motivation for mathematics (OECD, 2013). The assumption was that the rating standards (i.e. use of scale) of respondents do not vary across domains and vignettes describing one domain can be used to adjust self-reports in other domains. However, as documented by Vonkova et al. (2017), in their study students used the scale differently when assessing vignettes describing different subdomains of school behavior (e.g. dishonest behavior in school, bullying).

It still remains to be investigated whether the use of scale by the respondents differs across different areas of ICT skills. The use of a single set of vignettes to adjust different self-assessment questions might decrease the administration costs of the questionnaire and decrease the demands placed on the respondents, especially when younger students are surveyed. However, if there are differences in the way the respondents use scales in different areas of ICT skills, using a single set of vignettes might make the adjustment inaccurate.

In this study, we use the anchoring vignette method to examine the differences in scale usage between different groups of Czech upper-secondary students in various domains of ICT skills as defined by the DigComp framework. Our main research question is as follows: What are the differences in the way students from different schools use the scale in different domains of ICT skills as defined by DigComp?

Methodology

In this section, we (a) describe the anchoring vignettes used in this study that portray different areas of ICT skills based on the DigComp framework and (b) provide an example of a vignette from the domain Information. Then we describe the sample consisting of Czech upper-secondary schools of different types. The data

were collected as a part of a pilot study preceding a large-scale data collection on a representative sample of Czech upper-secondary schools of different types.

Anchoring vignettes describing different areas of ICT competence

For the creation of the anchoring vignettes, we used the DigComp framework (Ferrari, 2013) which defines five areas of ICT competence: Information, Communication, Content creation, Safety, and Problem solving. In each of these areas, three competence levels are distinguished: foundation, intermediate, and advanced. For each area, we developed three vignettes corresponding to the three competence levels, making the total of fifteen vignettes.

Note that our vignettes were shortened in comparison to the original DigComp descriptions with the aim of keeping the fundamental information but not overburdening the respondents with too long vignettes. We also added concrete examples into the vignettes in order to make them more understandable for the respondents. Also, the requirements described in the advanced level vignettes (corresponding to the advanced DigComp level) were slightly lowered to be more relevant to upper-secondary students. An example of an intermediate level vignette in the Information domain is (Vonkova et al., 2019):

Elisabeth uses internet search engines to search for information. She compares and validates information obtained from various sources on several web pages. She uses search engine filters (e.g. only pictures, only pages in the Czech language) to get the best results. She downloads the information and files she needs and sorts them into folders.

The respondents assessed the vignettes in a randomized order for each domain on a 7-point scale ranging from 1 = *the lowest level* to 7 = *the highest level*.

Sample

Data collection took place at four Czech upper-secondary schools (ISCED 3, age 15 – 19) in the first (initial) year and the fourth (final) year of study (total N = 166). We included students from: (a) a humanities oriented school offering a study programme Pedagogical Lyceum (56 students), (b) a technical non-ICT school focused on transport and mechanization (25 students), (c) a grammar school providing general education (47 students), and (d) a technical school offering a study programme focused on ICT (38 students). All four study programmes are four years long and are completed by School Leaving Examination. Since each chosen school was of a different type, the students were expected to have a different ICT background.

Results and discussion

Our results suggest that there are certain regularities in the assessments of students from different schools across the different ICT skills domains. For each school, we computed the mean assessment of each of the vignettes. Then, for each of the vignettes, we ranked the schools according to their mean assessment. The analysis showed that schools tended to rank similarly across the different ICT skills domains.

These results indicate that there might be a certain regular pattern in the way the respondents from different schools use scale across different domains of ICT skills. Such a conclusion might be of great importance to researchers in the domain of ICT skills measurement since they could administer only one set of anchoring vignettes describing a single domain and, using these vignettes, adjust a variety of self-assessment questions aiming at different domains of ICT skills. However, our results are only preliminary and further investigation of this issue seems advisable before making any definitive conclusions.

The issue of using vignettes describing a certain domain of a concept to adjust self-assessments of other unrelated concepts (or different subdomains of the same concept) seems still inconclusive and requires further exploration. Despite the approach used in PISA (OECD, 2013), Vonkova et al. (2017) showed that using vignettes describing other subdomains of student school behavior such as bullying to adjust student self-assessment of dishonest behavior is not recommendable. It might be possible that the different nature of the measured concepts plays role in the way respondents use scales across its subdomains.

In school behavior research, students might have very different evaluation standards when a student bullies another student and when a student cheats on his or her homework. These two subdomains of school behavior might be perceived as very different in nature and their severity. In the research on ICT knowledge and skills, on the other hand, students might have very similar evaluation standards across its subdomains, since the particular subdomains (e.g. programming, graphics) might be perceived in a similar fashion.

Conclusion

In this study, we investigated the differences in scale usage between different groups of Czech upper-secondary students in different domains of ICT skills as defined by DigComp. Our results indicate that students studying at different types of upper-secondary schools (and, presumably, with different ICT background), indeed, seem to use scale differently when assessing anchoring vignettes describing different domains of ICT skills. However, a closer analysis revealed that there might be certain regularities in the way these students assess anchoring vignettes across these different domains. These similarities across the different ICT competence areas might suggest that students tend to use the scale across these areas consistently.

Future research could investigate the differences in scale use across different ICT skills domains on a larger, perhaps a representative sample of students. Our results are only preliminary since the data were collected as a part of a pilot study on a limited number of schools. It would also seem appropriate to further investigate the possibilities and limitations of using vignettes describing one concept to adjust self-assessments of other concepts (or other subdomains of the same concept) with respect to the nature of the measured concepts.

Acknowledgements

This study is a result of the research funded by the Czech Science Foundation as the project GA ČR 17-02993S “Factors influencing the ICT skill self-assessments of upper-secondary school students”.

References

- Angelini, V., Cavapozzi, D., Corazzini, L. & Paccagnella, O. (2012): Age, health and life satisfaction among older Europeans. *Social Indicators Research*, 105(2), 293-308.
- 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.
- Baumgartner, H. & Steenkamp, J. B. E. (2001): Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38(2), 143-156.
- Bonsang, E. & van Soest, A. (2012): Satisfaction with social contacts of older Europeans. *Social Indicators Research*, 105(2), 273-292.
- Ferrari, A. (2013): *DIGCOMP: A framework for developing and understanding digital competence in Europe*. <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC83167/lb-na-26035-enn.pdf> (Accessed 6 January 2019).
- Frailon, J., Ainley, J., Schulz, W., Friedman, T. & Gebhardt, E. (2014): *Preparing for life in a digital age: The IEA International Computer and Information Literacy Study international report*. <https://link.springer.com/content/pdf/10.1007%2F978-3-319-14222-7.pdf> (Accessed 6 January 2019).
- 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* (pp. 253-268). Singapore: Springer Science.
- King, G., Murray, C. J., Salomon, J. A. & Tandon, A. (2004): Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191-207.
- Kyllonen, P. C. & Bertling, J. (2013): Innovative questionnaire assessment methods to increase crosscountry 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* (pp. 277-286). London, England: Chapman Hall/CRC Press.
- OECD (2013): *PISA 2012 results: What makes a school successful? Resources, policies and practices. Vol. 4, Annex A6: Anchoring vignettes in PISA 2012 student questionnaire*. Paris: OECD.
- Paulhus, D. L., Harms, P. D., Bruce, M. N. & Lysy, D. C. (2003): The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, 84(4), 890-904.
- Siddiq, F., Hatlevik, O. E., Olsen, R. V., Thronsen, I. & Scherer, R. (2016): Taking a future perspective by learning from the past – A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educational Research Review*, 19, 58-84.
- Vonkova, H., Bendl, S. & Papajoanu, O. (2017): How students report dishonest behavior in school: Self-assessment and anchoring vignettes. *The Journal of Experimental Education*, 85(1), 36-53.
- Vonkova, H. & Hrabak, J. (2015): 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.
- Vonkova, H., Hrabak, J., Kralova, K. & Papajoanu, O. (2019): *Developing a framework for the examination of anchoring vignette assumptions using cognitive interviews for (not only) the domain of ICT knowledge and skills*. Manuscript submitted for publication.

- Vonkova, H., Papajoanu, O. & Bendl, S. (2016): Aplikace metody ukotvujících vinět v pedagogickém výzkumu: přehled literatury a metodologická doporučení. *Pedagogická orientace*, 26(3), 537-559.
- Vonkova, H., Papajoanu, O. & Stipek, J. (2018): Enhancing the cross-cultural comparability of self-reports using the overclaiming technique: An analysis of accuracy and exaggeration in 64 cultures. *Journal of Cross-Cultural Psychology*, 49(8), 1247-1268.
- Vonkova, H., Zamorro, G. & Hitt, C. (2018): Cross-country heterogeneity in students' reporting behavior: The use of the anchoring vignette method. *Journal of Educational Measurement*, 55(1), 3-31.
- Weiss, S. & Roberts, R. D. (2018): Using Anchoring Vignettes to Adjust Self-Reported Personality: A Comparison Between Countries. *Frontiers in Psychology*, 9, 1-17.
- West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F. & Gabrieli, J. D. (2016): Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, 38(1), 148-170.

Mgr. Ondrej Papajoanu, Charles University, Czech Republic

Mgr. Katerina Kralova, Charles University, Czech Republic

Assoc. Prof. Dr. Hana Vonkova, Charles University, Czech Republic

Dr. Jiri Stipek, Charles University, Czech Republic