Computers & Education 85 (2015) 191–202

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

The (in) comparability of ICT knowledge and skill self-assessments among upper secondary school students: The use of the anchoring vignette method

Hana Vonkova*, Jan Hrabak

Charles University in Prague, Faculty of Education, Magdalény Rettigové 4, Prague 1 116 39, Czech Republic

ARTICLE INFO

Article history:
Received 16 December 2014
Received in revised form
2 March 2015
Accepted 3 March 2015
Available online 12 March 2015

Keywords:
Heterogeneity in reporting behavior
Self-assessment
Information and communication technology
The nonparametric approach of the anchoring vignette method
Entropy

ABSTRACT

Self-reporting of ICT knowledge and skills is commonly used in questionnaire surveys among students and in job search situations to provide information about actual ICT knowledge and skills. The main advantage of this approach is its low cost. The responses can, however, reflect not only the actual level of knowledge and skills but also the self-assessment style. Two students with the same actual knowledge and skills level may give different self-assessments since one might be optimistic and overestimate his/her skills and the other could be pessimistic and underestimate his/her skills. The anchoring vignette method helps to adjust self-reports for the differences in scale usage. We compare the ICT knowledge and skills of two distinct groups of Czech upper secondary school students, ICT students (N = 228) and non-ICT (business and pedagogy) students (N = 147), based on their self-reported general ICT knowledge and skills, before and after the anchoring vignette adjustment for the different usage of scale. We show that the anchoring vignette method helps to distinguish between the two groups of students' differences in scale usage and that adjusted self-reports correspond to the actual level of students' ICT knowledge and skills (in contrary to the unadjusted self-reports). The anchoring vignette method is a very promising tool for increasing the validity of surveys that use students' ICT knowledge and skill self-reports.

1. Introduction

Students around the world today are becoming increasingly familiar with computers and how to use them. This reflects the fast development of information and communication technologies (ICT) and their use in broad areas of human activities. An important research issue is finding valid ways to measure ICT knowledge and skills so that the data can be used, for example, to compare students, schools, regions or countries. There are several commonly used methods for measuring the ICT knowledge and the skills of students.

One important source are achievement tests. An example of these might be the International Computer and Information Literacy Study 2013 (ICILS 2013) which focuses on how young people have developed their computer and information literacy in order to support their capacity to participate in the digital age. It analyzes differences in computer and information literacy within and between twenty-one participating countries and investigates the factors that influence computer and information literacy among students in the eighth grade (Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014).

Another method of ICT knowledge and skills measurement is based on (self-)assessments used in questionnaire surveys. The big advantage of this method is its low cost compared to achievement tests, both in terms of time and finance. Self-assessments of ICT knowledge and skills are used, for example, to compare different groups of students or teachers (Danner & Pessu, 1995; Hakkarainen et al., 2000; Ilomäki & Rantanen, 2007; Lau & Yuen, 2014; Watkins & Cheung, 1995; de Wit, Heerwegh, & Verhoeven, 2014). They are also commonly used in curriculum vitaes (European Communities, 2003) or in job search situations with the aim of providing information about the real level of competency of job applicants. However, self-reported ICT knowledge and skills may reflect not only the objective situation but also the scale usage. For example, there may be two students with the same skills, but one of them might evaluate his skills as excellent

* Corresponding author. Tel.: +420 221 900 531.
E-mail addresses: h.vonkova@gmail.com (H. Vonkova), kabarh.naj@seznam.cz (J. Hrabak).

http://dx.doi.org/10.1016/j.compedu.2015.03.003
0360-1315/

while the other evaluates his as average. The reason behind this may be that the students use the scale differently. The first student is optimistic and overestimates his skills and/or the other student is pessimistic and underestimates his skills. A comparison of such self-reported competencies is then misleading. Therefore, it can threaten the validity of research and in a job-search situation it may disadvantage applicants with high self-evaluation standards.

Self-assessment questions are not only used in ICT knowledge and skills assessments, but also in the measurement of other educational concepts like non-cognitive skills, attitudes, values, classroom management, teacher support behavior, and teacher–student relationships etc. (see, for example, OECD, 2013). The heterogeneity of scale usage among different groups of students has been documented in many educational domains. For example, Buckley (2009) analyzes response styles for students from 57 countries in PISA 2006, dividing them into the acquiescence response style (also called positivity bias, a tendency to agree with items regardless of actual attitude), the disacquiescence response style (a tendency to disagree with items regardless of their content), the extreme response style (a tendency to choose the endpoints of an item's scale) and noncontingent responding (a term used to describe the random or careless response to items). He finds big differences between the response styles and among the countries. Chen, Lee, and Stevenson (1995) focus on cross-cultural comparisons of rating scales among East Asian and North American students. They conclude that Japanese and Chinese students use the midpoints on the 7-point Likert-type scales more often than North American students and that within each cultural group respondents who are individualistic tend to use the scale's extreme values while those who are more collectivist-oriented tend to favor the midpoints.

The anchoring vignette method has been introduced to adjust self-assessments so they take account of the differences in scale usage and allow them to be used to make comparisons between respondents (King, Murray, Salomon, & Tandon, 2004). The basic idea is to let respondents evaluate not only themselves but also hypothetical situations for other people (anchoring vignettes). Since the hypothetical situation is the same for all respondents, the differences in their evaluations of the situations may be interpreted as the differences in scale usage. This information is then used to correct (adjust, anchor) their self-reports. The method has been used widely in many research domains like political efficacy (King et al., 2004), health (Bago d'Uva, Lindeboom, O'Donnell, & van Doorslaer, 2008; Peracchi & Rossetti, 2012; Vonkova, 2013; Vonkova & Hullegie, 2011), life satisfaction (Angelini, Cavapozzi, Corazzini, & Paccagnella, 2012; Kapteyn, Smith, & van Soest, 2010), job satisfaction (Kristensen & Johansson, 2008), work disability (Angelini, Cavapozzi, & Paccagnella, 2011; Kapteyn, Smith, & van Soest, 2007).

Though the method has been used in many research domains, its use in educational research is rare (Buckley, 2008; Buckley & Schneider, 2007; Kyllonen & Bertling, 2013; Vonkova, Zamarro, & DeBerg, 2014) despite its high potential for enhancing the validity of research results. In our paper we compare the ICT knowledge and skills of two distinct groups of upper secondary school students by looking at their self-reported general ICT knowledge and skills levels, both before and after adjusting for the different usage of scale. Specifically, our sample consists of ICT and non-ICT (business and pedagogy) upper-secondary school students. The two groups differ in many of the observed characteristics related to ICT knowledge and skills, e.g. the ICT curriculum at their school, the ICT knowledge and skills required at the beginning of their studies and their motivation to actually study ICT. The implication of this is that the ICT knowledge and skills of the ICT students can be expected to be much higher compared to the non-ICT students. We first compare the ICT knowledge and skills of the two groups based on the raw data, i.e. their self-assessments. Then we identify the differences in scale usage based on their vignette evaluations and compare self-assessments adjusted for the two groups' differences in scale usage. Additionally, we also analyze the discriminatory power of our anchoring vignettes based on entropy and discuss the implications of this on the optimal choice of vignettes for our sample. As far as we are aware, this is the first study of the anchoring vignette method to self-reported ICT knowledge and skills assessments and one of the first applications of it in educational science. We show the methods' high potential for enhancing the validity of research that uses self-reported ICT knowledge and skills assessments and the possibilities for its use in the international context.

2. Methods

2.1. The anchoring vignette method

Responses to self-assessment questions in surveys may reflect not only the objective situations we seek to measure but also differences in scale usage. The anchoring vignette method has been proposed to avoid misleading results caused by differences in respondents' scale usage. The method uses additional information to reveal the potential differences – these being evaluations of hypothetical persons described in short stories, called anchoring vignettes. The differences between the respondents' vignette evaluations may be interpreted as the differences in their scale usage since the respondents evaluate the same stories. Note that the domain described in the anchoring vignettes should correspond to (or is in most applications of the method at least highly related to) the domain being measured by the self-assessments, since the scale usage may differ across different domains. For example, one should be careful about using vignettes concerning language skills to adjust a self-assessment about ICT skills.

The basic idea of the non-parametric approach of the anchoring vignette method is to relate the self-assessment s of a respondent to his/her vignette evaluations $v_1, v_2, \ldots, v_J$, assuming that the vignettes are naturally ordered (for example, from the lowest skills level of a hypothetical vignette person to the highest). For respondents with consistently ordered rankings on all vignettes ($v_j < v_{j+1}, j = 1, \ldots, J-1$), the vignette-corrected self-assessment C is then equal to 1 if $s < v_1$, equal to 2 if $s = v_1$, equal to 3 if $v_1 < s < v_2$, equal to 4 if $s = v_2$, …, and finally equal to $2J + 1$ if $v_J < s$. The meaning of each category is given by the hypothetical situations described in the vignettes used. For respondents with ties (for example, $v_1 = v_2$) and inconsistencies (for example, $v_2 < v_1$), C will have an interval value if multiple conditions are true. Consider, for example, a respondent with answers $s = v_2 < v_1 < v_3 < \ldots < v_J$. The equality $s = v_2$ means that $C = 4$, while $s < v_1$ means that $C = 1$. In such a situation we say that C has the interval value 1–4 and write $C = [1,4]$. The list of all possible rankings of self-assessments and vignette evaluations are given for two vignettes (Wand, King, & Lau, 2011) and three vignettes (van Soest & Vonkova, 2014).

We usually aim to summarize the vignette-corrected self-assessments of groups of respondents. In the case of only scalar values of C, we can draw a standard histogram. If there are interval values of C, the drawing of a histogram is not so straightforward. Four approaches have been proposed for dealing with the interval values: a) omitting interval values, b) uniform allocation within intervals, c) the censored ordered probit model – distributing each vector-valued response according to the proportion of "similar" respondents (gender, age, school type etc.) whose C values are spanned by the vector (King & Wand, 2007), and d) minimum entropy – the information in C that we know