



Brief Report

The effect of face-to-face interviewing on personality measurement



Luisa Hilgert*, Martin Kroh, David Richter

German Institute for Economic Research – DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany

ARTICLE INFO

Article history:

Received 5 November 2015

Revised 6 March 2016

Accepted 11 May 2016

Available online 20 May 2016

Keywords:

Personality

Big Five

ESEM

Face-to-face

Interviewer bias

ABSTRACT

In recent years, an increasing number of nationally representative surveys in the social sciences and economics have implemented the Big Five model of personality. While many personality inventories were originally developed in the context of self-administered questionnaires, they are often used by large surveys in face-to-face interview settings instead. Drawing on an experimental research design, we studied the effect of this switch in the method of data collection on measurement invariance as well as measurement error and interviewer effects in the Innovation Sample of the German Socio-Economic Panel (SOEP-IS). Although in some cases we found slightly stronger associations between interviewer and respondent personality in face-to-face settings, the results generally suggested strict measurement invariance—and therefore full comparability—across methods of data collection.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

In recent years, personality measurements have enjoyed increasing popularity outside of psychological research, particularly in the social sciences and economics (e.g., Heineck, 2011; Uysal & Pohlmeier, 2011). Studies pioneering the use of personality measures include the Socio-Economic Panel Study (SOEP), the Household, Income and Labour Dynamics in Australia (HILDA) Survey, and the British Household Panel Survey (BHPS). While psychological research traditionally uses self-administered questionnaires to measure personality, many surveys in the social sciences and economics use face-to-face interviews instead. However, little is known about the effect of this mode-switch on the measurement of personality.

A large body of survey research documents that the method of data collection can indeed have an impact on the quality of survey responses (for a meta-analysis, see Leeuw, 1992). Whereas face-to-face interviews enable the use of complex instruments and help to prevent misunderstandings, self-administered interviews often ensure greater privacy and fewer reactivity problems such as socially desirable responding (i.e., the tendency of respondents to give positive self-descriptions and to seek approval from the interviewer in their survey responses). Previous investigations of method effects in personality measures either were restricted to sub-populations (Sawhney & Cigularov, 2014) or used non-experimental data (Lang, John, Lüdtke, Schupp, & Wagner, 2011) to compare these two methods. In the present experimental study based on a nationally representative sample

of the German population, personality measures were collected through the use of (1) face-to-face interviews and (2) self-administered questionnaires (with an interviewer present in the household). In the basic setup, respondents were randomly assigned to one of the two groups, which differed only in the methods used for the collection of data on personality. First, we investigated (strict) measurement equivalence, which is required to compare personality scores and reliability in the two modes. Specifically, we compared measurement and structural invariance across both interviewer-administered and self-administered interviews. Second, we studied interviewer bias as previous research has shown that the presence of interviewers and specific interviewer characteristics, such as gender, race, and political attitudes, can influence response behavior (for an overview, see Schaeffer, Dykema, & Maynard, 2010). We expect the same to hold for interviewer personality: Since respondents tend to seek approval from interviewers as a form of impression management, they may adjust their responses in personality tests to fit their perceptions of the interviewer's personality. We considered both the variance in responses attributable to interviewers as well as the extent to which interviewer personality correlates with respondents' answers on the personality questionnaire.

2. Materials and methods

2.1. Sample

The SOEP is a large, random household sample that has been conducting annual interviews since 1984. The SOEP currently has around 28,000 adult respondents across Germany. The SOEP Innovation Sample, SOEP-IS (Richter & Schupp, 2012), provides

* Corresponding author.

E-mail address: lhilgert@diw.de (L. Hilgert).

an additional longitudinal sample for particularly innovative research projects. Both SOEP and SOEP-IS implement the Big Five constructs using a short version known as the BFI-S (Big Five Inventory-SOEP, Gerlitz & Schupp, 2005). In the 2013 version of the SOEP-IS study, half of a total of 5101 participants answered the BFI-S in the standard computer-based face-to-face method (Computer Assisted Personal Interviewing, CAPI), while the other half answered the questions through a self-administered questionnaire during the personal interview (Computer Assisted Self Interviewing, CASI). In the CAPI group, the interviewer read the questions and answer options given on the computer screen aloud to the respondents. The respondents then stated their answers and the interviewer recorded the responses on the computer. In the CASI group, interviewers were asked to interrupt the standard face-to-face interviewing at a given point during the interview for the section on personality measurement. Interviewers then turned the screen toward the respondents, who were asked to read the corresponding instructions and complete their answers on their own. In the CASI, the interviewers did not leave the households. However, they were asked not to intervene in the self-administered parts of the interview. Due to the longitudinal character of SOEP-IS, it is important to note that 1606 of the participants in 2013 had answered the BFI-questionnaire in previous years, typically in the face-to-face mode.

Participants who did not answer all of the BFI-S questions in 2013 (1.4%) were excluded from the analyses. In the face-to-face group, 2490 respondents provided complete data for the BFI-S, while in the self-administered condition complete data was available for 2511 participants. There were no differences with respect to nonresponse (χ^2 -test: $p = 0.723$) between the two study groups. There were no statistically significant differences in terms of the gender composition (χ^2 -test: $p = 0.295$) or family status distribution (χ^2 -test: $p = 0.396$). Mean age, however, was slightly higher in the face-to-face ($M = 52.0$) than in the self-administered ($M = 50.6$) group (t -test: $p < .01$).

The mode of data collection was randomized at the household level, so that the same mode of data collection was applied to all household members. The interviews were conducted by a total of 324 interviewers. The median number of interviews carried by interviewers was 13. Almost all interviewers conducted both face-to-face interviews as well as self-administered interviews. Additionally, the information on the personality of 217 interviewers was available from interviewer questionnaires that were conducted in 2006 and 2012.

2.2. Material: The BFI-S

Personality was measured with a 16-item German short version of the Big Five Inventory (BFI-S; Gerlitz & Schupp, 2005; validity: Hahn, Gottschling, & Spinath, 2012). The five-factor model of personality identifies the traits of openness (O), conscientiousness (C), extraversion (E), agreeableness (A), and neuroticism (N) (McCrae & Costa, 1997). All personality factors, with the exception of openness (for which a four-item scale was used) were measured with three items on a 7-point Likert scale that ranged from '1' (does not apply at all) to '7' (applies completely).

As for the consistency of the respective items, McDonalds ω (McDonald, 1999) was 0.65 for openness, 0.59 for conscientiousness, 0.67 for extraversion, 0.45 for agreeableness, and 0.63 for neuroticism. The interpretation of the ω values is the same as of the commonly used Cronbachs α .¹ The moderate consistency

values reflect the few items per dimension. As to the stability of BFI mean scores, we drew on the longitudinal nature of personality measurement in the SOEP-IS. More specifically, we inspected test-retest correlations of the factor mean scores across four years. For the participants that had answered the BFI-S questionnaire before, the retest correlation was $r = 0.57$ for openness, $r = 0.46$ for conscientiousness, $r = 0.58$ for extraversion, $r = 0.41$ for agreeableness, and $r = 0.57$ for neuroticism.² The moderate stability found for these values can be attributed to the long test-retest time interval of four years or more.

2.3. Procedures for inspecting measurement invariance

Since the assumption of the absence of cross-loadings in classical confirmatory factor analysis was too restrictive for our data, we instead applied Exploratory Structural Equation Modeling (ESEM; Asparouhov & Muthén, 2009) using quartimin rotation. Exploratory structural equation modeling is an integration of exploratory and confirmatory factor analysis. It provides a tool for multigroup comparisons of means and covariances of the latent factors (see below: Cross-Group Structural Equivalence) as well as for multigroup comparisons of factor loadings, intercepts and residual variances (see below: Cross-Group Measurement Equivalence). As Horn's Parallel Analysis (Horn, 1965) indicated a five-factor structure for both methods of data collection, we extracted five factors in our ESEM-Analysis.

2.4. Procedures for inspecting interviewer bias

We employed hierarchical linear modeling of the factor scores in order to decompose variance between interviewers and respondents nested within interviewers. Moreover, we used hierarchical modeling to study the effects of interviewers' personalities on respondents' personality measures.

3. Results and discussion

3.1. Invariance across modes of assessment

3.1.1. Cross-group measurement equivalence

To assess whether personality traits were measured equally in both groups, we tested for measurement equivalence. We followed a stepwise procedure, adding restrictions and comparing models.³ The model with strict invariance (equal factor loadings as well as intercepts and variances of residuals in both groups) was not rejected on a 1% α -level (χ^2 diff. test against strong invariance: $p = 0.034$). Hence, we can safely assume strict measurement equivalence across both groups, i.e., scores calculated from the items belonging to the same factor are comparable across groups.

3.1.2. Cross-group structural equivalence

To assess if the latent factor means and covariances between factors were the same across groups, we compared a model that set these restrictions to a model that did not. The model with strict invariance plus equal factor means and equal factor covariances was not rejected (χ^2 diff. test against strict invariance: $p = .336$). Hence, we can also assume structural equivalence across both

² For openness, the test-retest correlation was significantly lower in the face-to-face ($r = 0.578$) than in the self-administered group ($r = 0.595$). This was surprising insofar as the stability was slightly lower when the method of data collection did not change. A tentative interpretation may be that there was a higher reliability of measurements obtained through the self-administered personality measures.

³ Details in the online appendix.

¹ RMSEAs and model comparisons based on the Akaike information criterion (AIC) and likelihood ratio tests for goodness of fit suggested the use of congeneric models, which imply that all items meant to belong to a common factor measure the same phenomenon. Cronbachs α would have relied on essential tau-equivalence.

Download English Version:

<https://daneshyari.com/en/article/951229>

Download Persian Version:

<https://daneshyari.com/article/951229>

[Daneshyari.com](https://daneshyari.com)