Title: Psychiatric symptoms and response quality to self-rated personality tests: Evidence from the PsyCoLaus study.
Authors: Dupuis M, Meier E, Rudaz D, Strippoli MF, Castelao E, Preisig M, Capel R, Vandeleur CL
Journal: Psychiatry research
Year: 2017 Feb 20
Volume: 252
Pages: 118-125
DOI: 10.1016/j.psychres.2017.02.037

Psychiatric symptoms and response quality to self-rated personality tests: Evidence from the PsyCoLaus study

Marc Dupuis[1], Emanuele Meier[1], Dominique Rudaz[2], Marie-Pierre F. Strippoli[2], Enrique Castelao[2], Martin Preisig[2], Roland Capel[1] & Caroline L. Vandeleur[2]

[1] Institute of Psychology, University of Lausanne, Geopolis building, CH-1015 Lausanne, Switzerland

[2] Centre for Research in Psychiatric Epidemiology and Psychopathology, Lausanne University Hospital, Department of Psychiatry

Role of the funding source

The PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 3200B0–105993, 3200B0-118308, 33CSCO-122661, 33CS30-139468 and 33CS30-148401).

Corresponding author: Marc Dupuis; Institute of Psychology, University of Lausanne, Geopolis, CH-1015 Lausanne, Switzerland, Marc.Dupuis@unil.ch; Tel: +41 21 692 32 53; Fax: +41 21 692.32.65

Abstract

Despite the fact that research has demonstrated consistent associations between self-rated measures of personality dimensions and mental disorders, little has been undertaken to investigate the relation between psychiatric symptoms and response patterns to self-rated tests. The aim of this study was to investigate the association between psychiatric symptoms and response quality using indices from our functional method. A sample of 1,784 participants from a Swiss population-based cohort completed a personality inventory (NEO-FFI) and a symptom checklist of 90 items (SCL-90-R). Different indices of response quality were calculated based on the responses given to the NEO-FFI. Associations among the responses to indices of response quality, sociodemographic characteristics and the SCL-90-R dimensions were then established. Psychiatric symptoms were associated with several important differences in response quality, questioning subjects' ability to provide valid information using self-rated instruments. As suggested by authors, psychiatric symptoms seem associated with differences in personality scores. Nonetheless, our study shows that symptoms are also related to differences in terms of response patterns as sources of differences in personality scores. This could constitute a bias for clinical assessment. Future studies could still determine whether certain subpopulations of subjects are more unable to provide valid information to self-rated questionnaires than others.

Key words: functional method; psychiatric symptoms; response reliability; response validity; personality assessment inventory; self-rated questionnaires.

1. Introduction

Over the last several decades, a formidable corpus of research has been provided highlighting associations between personality using self-rated tests and mental disorders. The evidence that some personality traits are either risk factors or protective factors for mental health is now consistent, and covers different disorders so far. Regarding psychiatric disorders, personality has been particularly studied among patients with major depressive disorders (Bagby et al., 1996; Bagby et al., 1997; Huprich, 2000; Huprich et al., 2012; Quilty et al., 2013), bipolar disorders (Akiskal, 1983; Young et al., 1995; Engstrom et al., 2003; Almeida and Lafer, 2009; Quilty et al., 2009; Almeida et al., 2011; Kim et al., 2011; Jabben et al., 2012; Dupuis et al., 2016), and schizophrenia and other psychotic syndromes (Bagby et al., 1997; Lonnqvist et al., 2009; Boyette et al., 2013; Gurrera et al., 2014; Schirmbeck et al., 2015).

Nevertheless, little has been undertaken in order to check whether subjects affected by mental disorders are able to provide the same response patterns compared to individuals not affected by a mental disorder; likewise, nothing has been done to question this among individuals who suffer from subthreshold syndromes or even isolated symptoms. Despite the evidence of the relation between personality measures and mental disorders, mental disorders might be associated with something other than personality itself, namely, response patterns related to psychological status that interfere with personality measurement. Associations between personality and psychological issues also depend on factors such as age (Graham and Lachman, 2014; Lechner and Rammstedt, 2015); such associations are thus less generalizable than usually stated, and might still be attributable to confounders.

Two studies, Gurrera, O'Donnell, Rosenberg and McCarley (2005) and Gurrera, McCarley and Salisbury (2014) concluded that consistent abnormalities in personality measured in patients suffering from schizophrenia appeared to be caused by the cognitive deficits and symptoms related to the disorder. In addition, Lysaker, Bell, Kaplan, Greig & Bryson (1999)

stated that these consistent differences were related to positive symptoms of schizophrenia and emotional discomfort, while Bell and colleagues (2007) concluded that impaired insight makes self-rated measures of some personality factors less valid. These conclusions therefore contradict the rarely questioned assumption that personality might contribute to the development of schizophrenic syndromes. These findings are of major importance for psychological assessment suggesting that the sources of observed differences in test scores belong to different levels that we have summarized in Table 1. Indeed, the classical *true score theory,* assuming that an observed score is the function of the true score and a random error, could be enriched by hypothesizing that observed trait score differences are a function of trait-level differences, responding-level differences and response-level differences. *Trait-level differences* consist of true differences in a given psychological trait while *responding-level differences* are due to responding conditions, i.e. expectations, cognitive abilities, honesty, faking adequate abilities, so that some differences in test scores between individuals might be attributable to these responding-level differences. Last, *response-level differences* consist of the differences in scores among individuals that are neither attributable to difference in traits nor attributable to differences in responding, corresponding to the errors in measurement, including transient errors (Schmidt et al., 2003). This distinction between response-level differences and responding-level differences is also consistent with recent research that highlighted that differences in self-reported personality tests can be induced by the weather (Rammstedt et al., 2015), or by experimentally induced emotional states (Querengässer and Schindler, 2014) which are linked to responding-level differences. Yet, such an approach is still unique and should be applied to different populations in order to generalize the results, which is the purpose of the current study.

[*Insert Table 1 about here*]

Concerning personality tests, differences in response consistency across individuals have been largely discussed. In particular, authors like Tellegen (1988) or Reise and Waller (1993) have introduced the concept of *traitedness*, referring to the extent to which a respondent's answers to a given test fit the trait construct, and they have provided techniques in order to assess variation in traitedness based on item response theory. They concluded that traitedness is a phenomenon of high importance for psychological assessment that might lead to major methodological issues and that measuring such a phenomenon is difficult.

From a different theoretical perspective, Gendre has developed the *functional method as* a new scoring method that consists of modeling individual response patterns in multidimensional questionnaires (Dupuis et al., 2015). Based on response modeling, Gendre and colleagues proposed various indices (two of them specific to the functional method) to estimate the overall quality of a set of responses to a given self-rated test, and to determine whether the responses are valid enough for interpretation. These indices (*detailed below*) are useful to answer four main questions about response patterns: *how coherent, predictable and informative they are? How stable and reliable they are? How normative or even banal the responses are? How many positive and negative aspects are assumed in self-description?* Such indices are thus interesting in order to highlight responding-level-differences.

The only application of the functional method to psychiatric patients published so far has highlighted that subjects suffering from schizophrenia provided less coherent and less stable self-descriptions that were thus less reliable (Boulanger et al., 2013). Moreover, an application of the method to subjects from the general population resulted in important differences in responses that were associated with bipolar disorder, although they were not directly attributable to the disorder itself but to its correlates (Dupuis et al., 2016). Yet, the functional method was introduced to the English-speaking scientific community only very

recently (Dupuis et al., 2015), and systematic investigation of the role of psychopathological issues in responses to self-rated questionnaires remains largely unstudied.

The aim of this study was thus to use this new method to measure the associations between psychiatric symptoms and the response quality of a self-rated personality test in a community-based sample in Switzerland.

2. Methods

2.1. *Study design*

Cross-sectional data from the CoLaus|PsyCoLaus cohort study (Firmann et al., 2008; Preisig et al., 2009), a population-based study conducted in the city of Lausanne in the French-speaking region of Switzerland, was used. Briefly, the CoLaus study assessed cardiovascular risk factors and diseases and collected various genetic variants and biomarkers. The baseline recruitment and medical assessment of the CoLaus sample, which was completed between 2003 and 2006, has already been described in detail (Firmann et al., 2008). CoLaus was completed with a psychiatric assessment (PsyCoLaus) conducted after an interval of approximately one year (Preisig et al., 2009). Participants were recontacted five years after the initial somatic and psychiatric assessment s, respectively, in order to complete the follow-up investigation s, and the follow-up of PsyCoLaus also included self-reported measures similarly to the PsyCoLaus baseline assessment. Indeed, during the psychiatric parts, subjects were asked to complete a set of psychological self-questionnaires at home and to send them back to the research unit by post. Some additional self-report measures were introduced at the PsyCoLaus follow-up assessment. This study focuses on data from the follow-up of PsyCoLaus. Participation was voluntary and only transportation costs to the sites where the investigations took place were reimbursed.

2.2. *Participants*

A total of 2,848 subjects participated in the psychiatric arm of the study at baseline and also at follow-up. They received a first set of questionnaires at baseline that 2,162 (75.9%) participants agreed to complete and send back. All participants of the psychiatric follow-up received the second set of questionnaires which included both questionnaires of interest for the present study. The final sample consists of 1,784 (62.6%) participants who completed both of the self-questionnaires administered at follow-up used in this study.

2.3. *Measures*

Regarding sociodemographic variables, four characteristics were taken into account: age, sex (female = 1, male = 0), first language (French = 1, others = 0) and socioeconomic status (SES) using the Hollinghead's index (1975). Psychiatric symptoms were assessed using items from the revised Symptom Checklist (SCL-90-R), a 90-item self-reported screening inventory which covers the symptoms of the Axis-I disorders of the DSM manuals (Tatu et al., 1994; Paap et al., 2011). The original SCL-90-R is coded according to a Likert scale; instead for this study we recoded items as dichotomous in order to report whether the listed symptoms were present or absent over the past week, including during the day of filling out the questionnaire. Symptoms were grouped into nine dimensions: somatization (Cronbach's $\alpha = .86$), obsessive-compulsive patterns ($\alpha = .82$), interpersonal sensitivity ($\alpha = .78$), depression ($\alpha = .86$), anxiety ($\alpha = .78$), hostility ($\alpha = .68$), phobic anxiety ($\alpha = .68$), paranoid ideations ($\alpha = .72$) and psychoticism ($\alpha = .71$); and a group of seven additional items not related to these dimensions was also investigated. Concerning personality, the participants completed the French version of the NEO Five-Factor Inventory (NEO-FFI). The NEO-FFI is a 60-item questionnaire that assesses the Big Five dimensions of personality, also using a 5-point Likert-type scale (Rolland et al., 1998; Aluja et al., 2005) ranging from 0 (*strongly agree*) to 4 (*strongly disagree*).

Concerning the NEO-FFI, different indices of response quality were calculated, two of them using Gendre's functional method; the method of calculation of these indices is detailed below. Such indices are of great interest to detect data that are unusable or even misleading for both clinical and research purposes.

2.4. *Statistical analyses*

Several statistical analyses were performed in a procedure summarized as follows : first, eight indices of response quality (detailed below) were calculated using the NEO-FFI. Second, hierarchical regression models were conducted to determine the relations between sociodemographic variables (step 1), symptoms of the SCL-90R (step 2) and, each index of response quality, defined as the dependent variable of the models. Third, each model was linearized using simple power transformations calculated from the former corresponding non-linearized model.

Prior to these statistical analyses, the data needed to be transformed in order to calculate indices of response quality. Such indices result from the application of the functional method recently detailed by our group (Dupuis et al., 2015). In brief, the corner stone of Gendre's functional method is the creation of a Cartesian orthonormal and hyperspheric measurement space in which classical theorems and axioms of vectorial geometry are applicable. In this metric space, responses to items, factor scores, individual response strategies to an entire questionnaire, and other measures can be compared. This specific measurement space is obtained by repeated iterations of principal component analysis (PCA). The first PCA is performed on the answers to the items, in the usual way; however, factor analysis can be used instead at this step of the procedure. The second step consists of conducting PCA on the resulting loading matrix; the number of extracted components must be constrained to be the same as those resulting from the first PCA. The last step consists of reiterating PCAs on extracted factor scores. Then, each matrix line is divided by its norm. This procedure results

in the creation of a loading matrix called “matrix of item characteristics” where lines are the vectorial expression of the items in a same orthonormal and hyperspheric space. The aim of the successive PCA iterations is to warrant that factors are strictly orthonormal, that is to say independent, and that item coordinates are expressed in the very same standard metric. For this study, the third-step PCAs were iterated 10 times to ensure that both conditions were met. Then, functional factor scores can be calculated by establishing the “correlations” between responses to items on the original Likert-type scale and each column of item characteristic using their scalar product. The scalar product of responses and item characteristics are thus considered as functional factor scores and as the coordinates of a vector of response strategy. Nonetheless, given that the mathematical procedure constrains factor independence, the resulting dimensions do not necessarily have psychological meaning; the most interesting indices of response quality are calculable even when PCA or factor analyses result in a structure with no clear meaning. Thus, the functional factor scores were not used within the present study, and only the eight indices of response quality were calculated to determine whether classical factor scores were valid, reliable and interpretable.

Four specific functional indices of response quality can be calculated based on different specific vectors of response strategy. Because of the mathematical procedure applied, these indices can be considered as correlations, which makes them easily interpretable and comparable between individuals or between questionnaires. Two of them are only available using the functional method (i.e. response coherence and response reliability) and are the most important indices, while the two others (response positivity and response negativity) represent functional applications of already existing indices.

- *Response coherence* indicates how coherent a whole set of responses is. It consists of the norm of the vector of response strategy; in other words, this index is equal to the square root of the sum of squared correlations. Given that factor independence is

constrained by the mathematical procedure, response coherence is a multiple correlation index that indicates how predictable the response strategy is. Response coherence can be considered as a measure of traitedness applied on responses to a whole multidimensional questionnaire.

- *Response reliability* results from the application of the bisection method on the vector of response strategy. It consists of the scalar product of two vectors of strategy to the two halves of a questionnaire, which indicates whether one's response strategy was stable throughout the whole test. For this study, we identified pairs of closest items using minimal Euclidean distance in order to split the questionnaire into similar halves.

Response coherence and reliability have the main advantage of using the functional method to assess response quality in questionnaire data. Moreover, compared to classical indices measuring the internal consistency of a test (e.g. Cronbach's α) and test reliability (e.g. split-half reliability), these two indices assess personal consistency and personal reliability. In fact, both indices measure, in individual responses, that which classical indices measure in the test itself. The next two indices are also expressed in a correlational metric, and are based on the calculation of vectors of strategy; nonetheless, as aforementioned, simpler and unidimensional versions of these indices have been developed to be calculated without the functional method:

- *Response positivity* and *response negativity* measure the extent to which socially desirable and undesirable aspects of personality features have been chosen. Both response positivity and negativity result from the sum of the product of the general vector of strategy and a vector of strategy specific to desirable and undesirable items, respectively. For this study, items from the NEO-FFI were considered as positive when the mean score was higher than 2.5 on a Likert-type scale, and as negative when the mean score was below 1.5. Items with scores between 1.5 and 2.5 were considered

as neutral and were not used in these calculations. The two resulting indices are of interest in order to detect individuals giving "unbalanced" self-descriptions with only desirable (or undesirable) qualities endorsed by one respondent.

The main qualities of these indices are that they measure response strategy to the whole test, and that they rely on the multidimensionality of the questionnaire. In addition, four other indices that can be calculated independently from the functional method have been proposed; the two following indices can also be considered as correlations:

- *Response modality* is an index that measures how often the most frequently given answers were chosen by the respondent. Mathematically speaking, response modality is a weighted form of Cohen's kappa coefficient that is used to measure how many answers given are the item modal answers.
- *Response normativity* indicates the extent to which a whole set of responses fits general response tendencies to the questionnaire. Response normativity consists of the correlation between responses and the mean scores on each item.

The two remaining indices differ from the others regarding their metric qualities; indeed, they are expressed in the same Likert-type scale as the responses to items.

- *Response level* corresponds to the mean score of items before taking reverse coding into account. This index is useful to assess bias, particularly acquiescence and opposition.
- *Response variability* consists of the standard deviation of the answers. This index can be used to detect retention of information and extreme responses.

A lack of response reliability is the most detrimental issue that can be highlighted by analyzing response strategy (Dupuis et al., 2015). Indeed, null values in response reliability indicate random responding, which totally invalidates test results. In contrast, despite the fact that low values in response coherence, modality, normativity, positivity and negativity may

reflect different types of unusual response patterns, test results might still be interpretable and reflect aspects of respondents' personality (e.g. incoherence) that could exist in the population. In contrast, a low response reliability automatically implies a low coherence warranting that results are totally unusable. Results lacking response reliability reflect nothing but measurement error.

Once the functional method was applied to the NEO-FFI, two-step hierarchical multiple regression analyses assessing the associations among the previously defined indices of response quality to the NEO-FFI, sociodemographic variables and SCL-90-R scores were performed. Hierarchical regression analysis is a technique that consists of conducting sequential regressions in order to isolate the part of variance in a dependent variable that is explained by a given group of predictors, and to measure which part of the remaining unexplained variance is explained by others groups of predictors in subsequent steps. The aim of these analyses was to quantify the amounts of variance in indices of response quality that were explained by symptoms after taking the variance explained by sociodemographic characteristics into account. Thus, in a first step, age, sex, first language and SES were entered into the models; SCL-90-R dimensions were entered into the regression models in a second step. From this perspective, the analyses were conducted with a particular focus on the increase in determination coefficients between step 1 and step 2. In addition, since indices of response quality are highly correlated, separate regression models were necessary to circumvent problems related to multicollinearity. Given that strict linear relationships between predicted values and regressors were unlikely to be met, power transformations were performed to linearize each model; because non-strictly positive values might have been problematic, we added 1 to each variable with values distributed near 0 (i.e. response reliability, positivity, normativity and modality), and 1.1 with response negativity which values are mostly varying between -1 and 0. The aim of this study was to determine *to which*

degree symptoms explain response quality in general; thus the models were analyzed with a specific focus on the coefficients of determination.

Finally, each model was then tested for heteroscedasticity (i.e. differences in the variance of a variable associated with the value of another variable) using the Breusch-Pagan test. The Breusch-Pagan test is expected to be non-significant to consider that there is no heteroscedasticity of the model. The analyses were two-tailed, with $\alpha = 5\%$, and p-values were adjusted to multiple testing using the Bonferroni correction. The statistical analyses were performed using R (version 3.2.1) and the package *car*.

3. Results

3.1. *Sample description*

The sample consisted of 771 (43.2%) men and 1,013 women aged 57.23 years ± 8.74. Regarding the first language, French was the mother tongue for 1,233 (69.1%) participants, and a second language for 551 participants. Concerning SES, 1,479 participants (84.9%) were from middle-to-upper classes, while only 302 subjects (16.9%) were either from lower or lower middle classes. Finally, 3 participants (0.2%) did not provide information on SES. The scores of the SCL-90-R and NEO-FFI dimensions as well as the indices of response quality are reported in Table 2.

[*Insert Table 2 about here*]

3.2. *Associations between psychiatric symptoms and response quality to the NEO-FFI*

The Breusch-Pagan test (see Tables 3 and 4) was not significant for five models, highlighting that dependent variables could be considered as normally distributed; on the other hand, the test was significant for the three remaining models, suggesting that linearizing was insufficient to make the dependent variables normally distributed.

Table 3 reveals that significant associations were measured between three sociodemographic variables, namely, sex, first language and SES, and response coherence, explaining 4.1% of

the variance of the model. In addition, significant correlations between two SCL-90-R dimensions and response coherence were measured. Indeed, the second step of the model resulted in negative associations with obsessive-compulsive symptoms, and hostility, explaining a substantial proportion of the variance (i.e. 13.0%).

Regarding the second model (still reported in Table 2), only sex was significantly associated with response reliability. However, the full model explained 4.8% of the variance in this index.

Regarding response positivity and negativity, the hierarchical multiple regression models highlighted that both indices were significantly associated with sex in step 1, and with obsessive-compulsive symptoms and psychoticism in step 2. SES was specifically associated with response positivity, while interpersonal sensitivity, depression and hostility were associated with response negativity. Though both indices of response quality are strongly correlated, the models did not explain the same amount of the variance; the second step of the third model explained about 12.3% of variance in response positivity, while the second step of the fourth model was able to explain about 31.3% of variance in response negativity.

[*Insert Table 3 about here*]

As reported in Table 4, the fifth model highlighted several associations between sociodemographic variables, psychiatric symptoms and response normativity. In step 1, response normativity was associated with sex, first language and SES. In step 2, response normativity was more strongly associated with obsessive-compulsive symptoms and depression, and lower but still significant associations were found for interpersonal sensitivity, hostility and psychoticism. Despite these associations being small in effect size, they explained 31.4% of the variance in step 2.

Next, no significant associations were measured in step 1, and response modality was only significantly associated with somatization and depression. The model explained 5.5% of the

variance in the index, nonetheless, the Breusch-Pagan test was significant, indicating problems of heteroscedasticity.

Response level was only significantly associated with depression and paranoid ideations which explained about 7.1% of its variance. However, the Breusch-Pagan test was also significant.

Finally, response variability was associated with sex in step 1, and with obsessive-compulsive patterns, interpersonal sensitivity, anxiety and paranoid ideation in step 2, but the Breusch-Pagan test was significant revealing problems of heteroscedasticity.

[*Insert Table 4 about here*]

4. Discussion

The main aim of this research was to quantify the amount of variance in response quality that could be accounted for by symptoms. Despite the fact that only small partial regression coefficients were found (ranging from -.206 to .220) the present study highlights determination coefficients ranging from 4.8% to 34.3%. Our findings support the hypothesis that symptoms of mental disorders are associated with differences in terms of response patterns to self-rated personality questionnaires based on the five-factor model. This is consistent with the literature, but this single finding adds very little to the current state of knowledge on this topic. In fact, the core findings of this study are that psychiatric symptoms are associated with differences in the *quality* of the response patterns.

Up to 13% of the variance in response coherence was explained by the SCL-90-R dimensions, which suggests that psychiatric symptoms are associated with less consistent and less predictable response patterns. Comparatively, approximately 4% of the variance in response reliability was explained by sociodemographic characteristics and psychiatric symptoms. This result is reassuring, suggesting that subjects with psychiatric symptoms are able to provide stable reliable responses throughout their entire self-description, and this even if they suffer

from psychotic syndromes. However, such a result does not alleviate the question of mental disorders' correlates that are also associated with a lower response reliability, as pointed out concerning subjects suffering or having suffered from bipolar disorders (Dupuis et al., 2016).

Regarding response normativity, positivity and negativity, larger associations were found. Indeed, the SCL-90-R dimensions explained 31% of the variance in response normativity, 12% in positivity, and 31% in negativity, which can be considered as very large effect sizes. This suggests that subjects suffering from psychiatric symptoms provide less normative and more distinctive response patterns. These results support the hypothesis of an inverse relation between mental disorders and adequate abilities to improve, consciously or not, one's self-image in the eyes of others (Verschuere et al., 2014). On the other hand, such differences might also reflect lower self-esteem among subjects suffering from symptoms of psychiatric disorders. The fact that the models explained larger parts of variance in negativity was predictable; nonetheless such a result highlights that positivity and negativity are close but divergent dimensions, especially regarding their relations with symptoms.

Finally, homoscedasticity was not assumable in the three last models; however, only small parts of variance in response modality, level and variability were explained by the models.

The present study proposed to investigate response quality using eight indices. Given their theoretical and mathematical properties and according to the current results, a distinction between indices could be made. On one hand, there are indices specifically covering the validity and the usability of questionnaire results: these indices consist of response coherence and of response reliability (calculable only with the functional method), and of response level and variability. On the other hand, the four remaining indices are rather investigating how well individuals fit social expectations. The fact that response normativity and response negativity are in particular largely explained by SCL-90-R dimensions was the most predictable finding, and is consistent with former studies (DeVylder and Hilimire, 2015;

Dupuis et al., 2016). Another major finding is that symptoms seem to explain a large part of the variance in response coherence, and a non negligeable part of the variance in response reliability, in terms of bias-attributable parts of variance in test results.

Despite the fact that this study focused on response patterns instead of factor scores, the current study is consistent with other studies that focused on specific populations like patients suffering from schizophrenia (Lysaker et al., 1999; Gurrera et al., 2000; Bell et al., 2007; Boulanger et al., 2013; Gurrera et al., 2014) or psychopathy (Verschuere et al., 2014) that stated that symptoms may have an effect on response patterns in self-rated personality tests. Nevertheless, the current results rely on data from a population-based survey, which contributes to the generalizability of the former findings.

As pointed out throughout the paper, the corpus of evidence of differences in personality scores in patients is ever-growing; yet, the possible explicative theories on such differences deserve to be discussed. On one hand, a first hypothesis suggests that psychopathology has an effect on responding abilities, in line with the scar model of neurocognitive impairments (Hesse et al., 2015; Dupuis et al., 2016). On the other hand, a second, alternative hypothesis assumes that some psychiatric syndromes have an effect on identity (Boulanger et al., 2013) instead of responding. The current study suggests that an index such as response coherence may reflect differences in the responding level (i.e. cognitive ability to respond to a questionnaire) and in identity as well. However, response reliability is more likely to reflect essentially cognitive issues. Given the difference in parts of the explained variance in response reliability and response coherence, one could argue that both identity and responding might differ in subjects suffering from mental disorders. In any case, further research on this question is recommended.

Admitting that self-rated personality measures reflect the presence of psychiatric symptoms, three more main questions can still be raised. First, do self-rated personality tests still measure

personality in subjects with symptoms? In other words, is self-rated personality assessment valid among symptomatic subjects? Second, if not, what are they measuring instead? Third, are these associations generalizable to other kinds of self-description? Those questions appear far from being answered yet. Nonetheless, these new results do not disqualify all the previous research on personality and mental disorders; on one hand, they do question the validity of personality measures, but on the other hand, they keep confirming the consistent significant associations stated in the literature.

The current results suggest that important differences exist in terms of normative responding. These differences concern not only *what* symptomatic subjects describe (i.e. response-level differences), but also *how* they respond (i.e. responding-level differences). For example, two individuals with similar levels of neuroticism might obtain radically different scores on this factor depending on their social awareness and their ability to enhance their self-image.

This study has some limitations that must be acknowledged. First, cross-sectional data was used, which implies that causal relations between variables cannot be assumed. In addition, it is the participants themselves, who had to report on the presence of symptoms; though reporting only on the presence or absence of symptoms reduces the impact of bias in responding compared to reporting frequencies on a Likert-type scale, possible bias attributable to self-report remains. Further research is suitable to ascertain whether response quality is explained either by symptoms only, or by both symptoms and specific response patterns to the SCL-90-R. Using external criteria for future studies is thus of high importance to improve knowledge on the relations between symptoms and response quality. Nevertheless, the current results clearly highlighted strong relations between response quality to the NEO-FFI and symptoms; these results suggest that those relations are probably meaningful even if every possible confounding factor has not been excluded so far. Finally, since the study is based on a cohort of 40-to-80 year-old citizens from a large Swiss city, its results could be

representative of the populations from Western European cities, but not necessarily of younger populations nor rural ones.

To conclude, as pointed out by Bell et al. (2007), despite the formidable corpus of studies that explored the relation between psychopathology and personality, work that considers symptoms as potential sources of bias in personality assessment is still scarce. Yet, much is at stake for both clinical practice and research. We recommend that future studies still determine whether certain specific subpopulations are more unable to provide valid information on self-rated questionnaires than others.

Conflicts of interest

None.

Role of the funding source

The PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 3200B0–105993, 3200B0-118308, 33CSCO-122661, 33CS30-139468 and 33CS30-148401).

Acknowledgments

The authors would like to express their sincere gratitude to the participants of the study and to the colleagues who contributed to data collection and management of the study. The authors also wish to thank Dr Mehdi Gholam-Rezaee and Dr Jean-Philippe Antonietti for their advice regarding the statistics.

References

Akiskal, H.S., 1983. The Relationship of Personality to Affective Disorders. Archives of General Psychiatry 40, 801-810.

Almeida, K.M., Lafer, B., 2009. A sib-pair study of personality traits in bipolar disorder. Bipolar Disorders 11, 15-15.

Almeida, K.M., Nery, F.G., Moreno, R.A., Gorenstein, C., Lafer, B., 2011. Personality traits in bipolar disorder type I: a sib-pair analysis. Bipolar Disorders 13, 662-669.

Aluja, A., Garcia, O., Rossier, J., Garcia, L.F., 2005. Comparison of the NEO-FFI, the NEO-FFI-R and an alternative short version of the NEO-PI-R (NEO-60) in Swiss and Spanish samples. Personality and Individual Differences 38, 591-604.

Bagby, R.M., Bindseil, K.D., Schuller, D.R., Rector, N.A., Young, L.T., Cooke, R.G., Seeman, M.V., McCay, E.A., Joffe, R.T., 1997. Relationship between the five-factor model of personality and unipolar, bipolar and schizophrenic patients. Psychiatry Research 70, 83-94.

Bagby, R.M., Young, L.T., Schuller, D.R., Bindseil, K.D., Cooke, R.G., Dickens, S.E., Levitt, A.J., Joffe, R.T., 1996. Bipolar disorder, unipolar depression and the Five-Factor Model of Personality. Journal of Affective Disorders 41, 25-32.

Bell, M., Fiszdon, J., Richardson, R., Lysaker, P., Bryson, G., 2007. Are self-reports valid for schizophrenia patients with poor insight? Relationship of unawareness of illness to psychological self-report instruments. Psychiatry Research 151, 37-46.

Boulanger, M., Dethier, M., Gendre, F., Blairy, S., 2013. Identity in schizophrenia: a study of trait self-knowledge. Psychiatry Research 209, 367-374.

Boyette, L.L., Korver-Nieberg, N., Verweij, K., Meijer, C., Dingemans, P., Cahn, W., de Haan, L., Group, 2013. Associations between the Five-Factor Model personality traits and psychotic experiences in patients with psychotic disorders, their siblings and controls. Psychiatry Research 210, 491-497.

DeVylder, J.E., Hilimire, M.R., 2015. Screening for psychotic experiences: social desirability biases in a non-clinical sample. Early Interv Psychiatry 9, 331-334.

Dupuis, M., Capel, R., Meier, E., Rudaz, D., Strippoli, M.-P.F., Castelao, E., Preisig, M., Vandeleur, C.L., 2016. Do bipolar subjects' responses to personality questionnaires lack reliability? Evidence from the PsyCoLaus study. Psychiatry Research 238, 299-303.

Dupuis, M., Meier, E., Capel, R., Gendre, F., 2015. Measuring individuals' response quality in self-administered psychological tests: an introduction to Gendre's functional method. Frontiers in Psychology 6, 1-12.

Engstrom, C., Brandstrom, S., Sigvardsson, S., Cloninger, R., Nylander, P.O., 2003. Bipolar disorder. II: Personality and age of onset. Bipolar Disorders 5, 340-348.

Firmann, M., Mayor, V., Vidal, P.M., Bochud, M., Pecoud, A., Hayoz, D., Paccaud, F., Preisig, M., Song, K.S., Yuan, X., Danoff, T.M., Stirnadel, H.A., Waterworth, D., Mooser, V., Waeber, G., Vollenweider, P., 2008. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. BMC Cardiovascular Disorders 8, 6.

Graham, E.K., Lachman, M.E., 2014. Personality Traits, Facets and Cognitive Performance: Age Differences in Their Relations. Personality and Individual Differences 59, 89-95.

Gurrera, R.J., McCarley, R.W., Salisbury, D., 2014. Cognitive task performance and symptoms contribute to personality abnormalities in first hospitalized schizophrenia. Journal Psychiatric Research 55, 68-76.

Gurrera, R.J., Nestor, P.G., O'Donnell, B.F., 2000. Personality Traits in Schizophrenia. The Journal of Nervous & Mental Disease 188, 31-35.

Gurrera, R.J., Nestor, P.G., O'Donnell, B.F., Rosenberg, V., McCarley, R.W., 2005. Personality Differences in Schizophrenia Are Related to Performance on Neuropsychological Tasks. The Journal of Nervous and Mental Disease 193, 714-721.

Hesse, K., Kriston, L., Wittorf, A., Herrlich, J., Wolwer, W., Klingberg, S., 2015. Longitudinal relations between symptoms, neurocognition, and self-concept in schizophrenia. Front Psychol 6, 1-9.

Hollingshead, A.B., 1975. Four factor index of social status. Yale University Press, New Haven.

Huprich, S.K., 2000. Describing depressive personality analogues and dysthymics on the NEO-Personality Inventory-Revised. Journal of Clinical Psychology 56, 1521-1534.

Huprich, S.K., Pouliot, G.S., Bruner, R., 2012. Self-other representations mediate the relationship between Five-Factor Model depression and depressive states. Psychiatry 75, 176-189.

Jabben, N., Arts, B., Jongen, E.M.M., Smulders, F.T.Y., van Os, J., Krabbendam, L., 2012. Cognitive processes and attitudes in bipolar disorder: A study into personality, dysfunctional attitudes and attention bias in patients with bipolar disorder and their relatives. Journal of Affective Disorders 143, 265-268.

Kim, B., Joo, Y.H., Kim, S.Y., Lim, J.H., Kim, E.O., 2011. Personality traits and affective morbidity in patients with bipolar I disorder: The five-factor model perspective. Psychiatry Research 185, 135-140.

Lechner, C.M., Rammstedt, B., 2015. Cognitive Ability, Acquiescence, and the Structure of Personality in a Sample of Older Adults. Psychological Assessment.

Lonnqvist, J.E., Verkasalo, M., Haukka, J., Nyman, K., Tiihonen, J., Laaksonen, I., Leskinen, J., Lonnqvist, J., Henriksson, M., 2009. Premorbid Personality Factors in Schizophrenia and Bipolar Disorder: Results From a Large Cohort Study of Male Conscripts. Journal of Abnormal Psychology 118, 418-423.

Lysaker, P.H., Bell, M.D., Kaplan, E., Greig, T.C., Bryson, G.J., 1999. Personality and psychopathology in schizophrenia: the association between personality traits and symptoms. Psychiatry 62, 36-48.

Paap, M.C., Meijer, R.R., Van Bebber, J., Pedersen, G., Karterud, S., Hellem, F.M., Haraldsen, I.R., 2011. A study of the dimensionality and measurement precision of the SCL-90-R using item response theory. International Journal of Methods in Psychiatric Research 20, e39-55.

Preisig, M., Waeber, G., Vollenweider, P., Bovet, P., Rothen, S., Vandeleur, C., Guex, P., Middleton, L., Waterworth, D., Mooser, V., Tozzi, F., Muglia, P., 2009. The PsyCoLaus study: methodology and characteristics of the sample of a population-based survey on psychiatric disorders and their association with genetic and cardiovascular risk factors. BMC Psychiatry 9.

Querengässer, J., Schindler, S., 2014. Sad but true? - How induced emotional states differentially bias self-rated Big Five personality traits. BMC Psychology 2, 14.

Quilty, L.C., Pelletier, M., DeYoung, C.G., Bagby, R.M., 2013. Hierarchical personality traits and the distinction between unipolar and bipolar disorders. Journal of Affective Disorders 147, 247-254.

Quilty, L.C., Sellbom, M., Tackett, J.L., Bagby, R.M., 2009. Personality trait predictors of bipolar disorder symptoms. Psychiatry Research 169, 159-163.

Rammstedt, B., Mutz, M., Farmer, R.F., 2015. The Answer Is Blowing in the Wind. European Journal of Psychological Assessment 31, 287-293.

Reise, S.P., Waller, N.G., 1993. Traitedness and the Assessment of Response Pattern Scalability. Journal of Personality and Social Psychology 65, 143-151.

Rolland, J.P., Parker, W.D., Stumpf, H., 1998. A psychometric examination of the French translations of the NEO-PI-R and NEO-FFI. Journal of Personality Assessment 71, 269-291.

Schirmbeck, F., Boyette, L.L., van der Valk, R., Meijer, C., Dingemans, P., Van, R., de Haan, L., for GROUP, Kahn, R.S., de Haan, L., van Os, J., Wiersma, D., Bruggeman, R., Cahn, W., Meijer, C., Myin-Germeys, I., 2015. Relevance of Five-Factor Model personality traits for obsessive-compulsive symptoms in patients with psychotic disorders and their un-affected siblings. Psychiatry Res 225, 464-470.

Schmidt, F.L., Le, H., Ilies, R., 2003. Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-differences constructs. Psychological Methods 8, 206-224.

Tatu, P., Pellet, J., Lang, F., Pichon, M., Coerchon, C., 1994. [SCL 90 R in a sample of French-speaking subjects free of psychiatric disorders]. Annales Médico-Psychologiques 152, 187-190.

Tellegen, A., 1988. The Analysis of Consistency in Personality-Assessment. Journal of Personality 56, 621-661.

Verschuere, B., Uzieblo, K., De Schryver, M., Douma, H., Onraedt, T., Crombez, G., 2014. The inverse relation between psychopathy and faking good: not response bias, but true variance in psychopathic personality. Journal of Forensic Psychiatry & Psychology 25, 705-713.

Young, L.T., Bagby, R.M., Cooke, R.G., Parker, J.D.A., Levitt, A.J., Joffe, R.T., 1995. A Comparison of Tridimensional Personality Questionnaire Dimensions in Bipolar Disorder and Unipolar Depression. Psychiatry Research 58, 139-143.

Table 1

Summary of conceptual levels and differences in observed scores

Level	Definition	Meaning of individual differences
Trait	The latent dimension that an instrument aims to measure, its true score.	Differences in a trait are actual differences that a perfect measurement instrument is supposed to ensure once measurement error and biases are partialled out.
Responding	The voluntary and involuntary strategy adopted to respond to a questionnaire. Responding includes very different patterns; some of them correspond to response biases: situational social desirability bias (i.e. faking either good or bad), agreeing with every proposition of a questionnaire (i.e. acquiescence), completing a questionnaire with little care (i.e. insufficient effort responding), etc. Some patterns result from social (e.g. culture, gender, etc.), and some from individual characteristics (i.e. cognitive abilities, age, etc.).	Responding refers to a qualitative process, differences in responding are thus difficult to measure. Nevertheless, they imply that two true scores of a same trait are expressed in a different way which might not be comparable.
Response	The observed score on the latent construct, what the instrument actually measures: the response is an observed score or an answer which is provided to a given item.	Differences in responses are observed differences. They consist of both potential differences in traits and errors in measurement. Response-level differences correspond to differences that neither result from differences in traits nor from differences in responding, that-is-to-say to differences attributable to error in measurement.

Table 2

Sample characteristics

Variable	Mean	SD	Participants	
			N	Proportion
Age	57.23	8.74		
Sex				
Male			771	43.2%
Female			1,013	56.8%
First language				
French			1,233	69.1%
Other			551	30.9%
Socioeconomic status				
Lower class			123	6.9%
Lower middle class			179	10.0%
Middle class			471	26.4%
Upper middle class			544	30.5%
Upper class			464	26.0%
Unknown			3	0.2%
Symptom dimensions (SCL-90-R)				
Somatization (0-12)	2.78	2.71		
Obsessive-compulsive (0-10)	2.72	2.61		
Interpersonal sensitivity (0-9)	2.00	1.99		
Depression (0-13)	3.33	3.27		
Anxiety (0-10)	1.20	1.79		
Hostility (0-6)	1.05	1.28		
Phobic anxiety (0-7)	0.51	1.02		
Paranoid ideations (0-6)	1.35	1.56		
Psychoticism (0-10)	0.90	1.46		
Additional items (0-7)	2.19	1.56		
Personality (NEO-FFI)				
Neuroticism (0-48)	17.80	7.75		
Extraversion (0-48)	28.01	6.25		
Openness (0-48)	28.87	6.05		
Agreeableness (0-48)	33.47	5.20		
Conscientiousness (0-48)	34.96	5.78		
Response quality (NEO-FFI)				
Coherence	0.71	0.12		
Reliability	0.85	0.18		
Normativity	0.56	0.19		
Positivity	0.80	0.21		
Negativity	-0.80	0.23		
Modality	-0.16	0.26		
Level	1.99	0.23		
Variability	1.20	0.25		

Table 3

Associations[§] among response indices of coherence, reliability, positivity, negativity sociodemographic characteristics and SCL-90-R dimensions (N=1,784)

Variable	*Coherence*		*Reliability*		*Positivity*		*Negativity*	
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
Step 1 (age, sex, mother tongue & SES)	.041***		.011**		.036***		.012***	
Age		.004		.008		.011		.051
Sex		.142***		.081**		.161***		.098***
Mother tongue (French = 1, other = 0)		.089**		.020		-.020		.004
Socioeconomic status (SES)		.133***		.077		-.070*		.006
Step 2 (symptoms)	.130***		.037***		.123***		.313***	
Somatization		-.010		.045		.078		.081
Obsessive-compulsive		-.195***		-.075		-.123**		-.138*
Interpersonal sensitivity		-.045		-.002		-.091		-.129***
Depression		-.093		-.048		-.070		-.220***
Anxiety		.096		-.018		.026		.005
Hostility		-.103***		-.079		-.057		-.100***
Phobic anxiety		-.002		-.015		-.055		-.038
Paranoid ideations		-.032		.022		.002		-.002
Psychoticism		-.030		-.014		-.148***		-.123***
Additional items		-.005		-.035		.047		.020
Total R^2	.171***		.048***		.159***		.325***	
λ[a]	3.004		9.179		6.436		-.0559	
Breusch-Pagan test (df=14)	13.843		2.951		17.829		22.035	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

§ Based on hierarchical multiple regression analyses; [a] λ corresponds to the exponent used to linearize the models

Table 4

Associations[§] among response indices of normativity, modality, level, variability, sociodemographic characteristics and SCL-90-R dimensions (N=1,784)

Variable	*Normativity*		*Modality*		*Level*		*Variability*	
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
Step 1 (age, sex, mother tongue & SES)	.029***		.001		.010		.012	
Age		.021		.012		-.014		-.039
Sex		.147***		.009		-.057		.083**
Mother tongue (French = 1, other = 0)		.065*		.011		.030		.039
Socioeconomic status (SES)		.078**		.018		.062		.058
Step 2 (symptoms)	.314***		.055***		.061***		.050***	
Somatization		.055		.091*		.021		-.036
Obsessive-compulsive		-.183***		-.008		.077		-.145**
Interpersonal sensitivity		-.087*		.047		-.055		-.136**
Depression		-.206***		-.127*		-.161**		-.009
Anxiety		.022		-.091		-.002		.118*
Hostility		-.095**		.019		.059		-.037
Phobic anxiety		-.047		-.047		.027		.029
Paranoid ideations		.008		-.071		.213***		.126**
Psychoticism		-.131***		-.070		.073		-.056
Additional items		-.003		.041		.005		-.044
Total R^2	.343***		.055***		.071***		.062***	
λ[a]	4.613		1.125		1.886		.787	
Breusch-Pagan test (df=14)	15.083		46.257***		48.407***		61.610***	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

§ Based on hierarchical multiple regression analyses; [a] λ corresponds to the exponent used to linearize the models