Employee Recruitment: Identifying Response Distortion on the Personality Measure

Wahyu Widhiarso
Fathul Himam

Abstract

The aim of this study was to identify individuals who fake their response on personality assessments in the context of employee recruitment. In the study experiment, participants were randomly divided into two groups. The first group was instructed to complete a measure (BFI-44) honestly while the second group was instructed to complete the inventory as if they were job applicants participating in an employee recruitment. It was hoped that the second group would be induced to give fake responses. Cluster analysis and latent class analysis for a two-class constrained model was applied to and fitted to the data. The correlation between actual group (honest vs. faking) and predicted group obtained from analysis result (normal vs. unique) was moderate: values for the phi correlation using a 2x2 crosstab table yielded values from 0.02 to 0.37. The percentage of participants detected as giving fake responses ranged from 37 to 69 %.

Keywords: Faking Responses, Cluster/Class Membership, Applicants-Respondents

Introduction

In addition to technical skills, psychological attributes such as personality traits play important roles in predicting individuals' work performance. Barrick and Mount (1991) reviewed several studies that examine the relationship between personality and work performance. They found that personality traits have a strong relationship with work performance. In the context of the big five personality factors (*openness, conscientiousness, extraversion, agreeableness, and neuroticism*), conscientiousness is most associated with performance. Given the important impact of personality factors on job performance, Ryan and colleagues (1999) conducted a survey comprising 959 organizations located in 20 countries. They found that most organizations surveyed employed personality measures for personnel selection. Although self-report measures have been extensively used in research settings, such assessments are used less in employee selection or human resources mapping due to this format's susceptibility to faking (Converse, Peterson, & Griffith, 2009).

Self-report measure requires persons to understand themselves and be able to make subjective judgments about their experiences. On personality measures of this type, the "correct" answer (i.e., the one that will make applicants appear the best in the eyes of the potential employer) is often clear. As a result, test-takers often deliberately provide inaccurate responses (Eid & Zickar, 2007); (Levashina, Morgeson, & Campion, 2012). The susceptibility of self-report measures to faking is largely due to respondents' full autonomy and freedom in selecting responses; items are usually easy to understand and the assessed attribute is usually easily predicted, as the item statement provides a clear clue about the assessed attribute.

Fake responses represents errors that can affect decision-making in the personnel selection process (Mueller-Hanson, Heggestad, & Thornton, 2003). Griffith, Chmielowski, and Yoshita (2007) found that faking interferes with decision-making regarding personnel ranking based on obtained scores. Using an experimental study, these authors found that about 30–50 % of individuals who gave fake responses passed the selection procedure and were offered jobs. These results were consistent with previous research conducted by Rosse et al. (1998), who found that in rigorous selection processes, many applicants who fake their response make it the top of the list of potential candidates. As fake responses represent a variable that is unrelated to either work performance or other variables that predict success in work-related performance, efforts should be made to reduce the presence of fake responses in measurement in this context. Weiner and Gibson (2000) found that the intensity of response deception has no relationship with either job performance or performance during training, and even has a negative relationship with cognitive test scores. Other studies have also found similar results regarding the negative correlation between giving more deceptive responses and intelligence (e.g. Moutafi, Furnham, & Crump, 2003). These studies offer even more motivation to development assessments that are resistant to faking, and to give researchers and employers more tools with which to detect individuals who are faking their responses

Faking on Personality Test

There has been a great deal of research on faking on personality scales (e.g. Converse et al., 2009; Dilchert, Ones, Viswesvaran, & Deller, 2006; Hartman & Grubb, 2011). Most researchers found that personality measures were susceptible to faking. In a meta-analysis of 51 studies about faking, Viswesvaran and Ones (1999) found that the mean scores for samples of applicant samples tend to be 0.48–0.65 standard deviations above the mean scores of incumbent samples. Birkeland and colleagues (2006) also conducted a meta-analysis of 33 studies, and discovered that applicants obtained significantly higher scores than non-applicants on four of the five personality factors: conscientiousness (effect size/d = .45), emotional stability (d = .44), openness (d = .13), and extraversion (d = .11). These reported effect size were, however, smaller than those reported by Viswesvaran and Ones (1999) .

Various terms can be used to describe

faking one's responses on an assessment, such as *deception, impression management, response distortion,* and *response set*. A faking proxy, *social desirability,* refers to the tendency of individuals to produce a deceptive answer. Another behavior associated with faking is *acquiescence,* the tendency to endorse all given statement on the scale in a similar manner. For example, *yea-saying* refers to the tendency to agree with all questionnaire or personality test items, regardless of their content (Bachman & O'Malley, 1984). However, the current study restricts the term *faking* to conscious efforts to manipulate one's responses to create a positive impression (McFarland & Ryan, 2000).

The present study attempts to identify individuals who have faked their responses. A widely used procedure to accomplish this goal is to administer an instrument that serves as a good proxy for faking; common assessments for this purpose include the Social Desirability Scale (Edwards, 1957), the Marlowe-Crowne Social Desirability Scale (Marlowe & Crowne, 1961), and the Jackson Social Desirability Scale (Jackson, 1984). These instruments assess the tendency of individuals to adapt what is considered good or ideal to social norms. Although in a number of studies these instruments have highly correlated with faking (Andrews & Meyer, 2003), some researchers have objected to making this connection. They argue that individuals who try to respond to item stem such as to make a good social impression are different from individuals who are deliberately faking responses (e.g., Holden, 2007).

Numerous studies have confirmed that self-report is fakable (see Morgeson et al., 2007); the next logical step is therefore to figure out how to detect whether individual have faked their responses (in the present study, among job applicants in industrial and organizational setting). A number of tests for detecting fake responses have been conducted, mostly in clinical psychological research. Behaviors associated with faking are often termed *dissimulation* or *malingering,* i.e., symptom faking or exaggeration. Common methods are using specific tests to detect malingering (Greve, Ord, Curtis, Bianchini, & Brennan, 2008), embedding sets of items or subtests that are able to detect faking into the larger instrument (Sellbom, Toomey, Wygant, Kucharski, & Duncan, 2010), examining consistency within individuals' responses (Heinrich & Borkenau, 1998), and using statistical analysis (Elhai & Frueh, 2001).

Research on faking in the context of self-report assessments generally uses statistical analyses of responses on assessments that use social desirability as a faking proxy. Using social desirability scales as a method for detecting fake responses has been widely examined (e.g. Paulhus, 1984). However, using such scales to detect fake responses is complex. Even in normal situations, social desirability scales can incorrectly identify honest respondents as faking, since social desirability is a different attribute than the true intention to fake responses (Burns, Christiansen, Griffith, & Peterson, 2006). The statistical techniques that have been employed to detect faking are mixed-Rasch modeling (Zickar, Gibby, & Robie, 2004), hybrid Rasch-latent modeling (Holden & Book, 2009), and structural model equation modeling (SEM; Ferrando & Anguiano-Carrasco, 2013). Identifying fake responses using statistical techniques is a promising method with several advantages: it is easy to apply and inexpensive, it can be applied for various measurement purposes, and it does not require any change to the instruments used. The present study follows up this approach by using two statistical analyses to simultaneously process the same data. The results from these two techniques are then compared to examine which technique is appropriate for detecting fake responses.

Purpose of Current Study

Research on faking has reached a consensus: personality self-reports are susceptible to faking. To address this problem there are two approaches proposed, increase the resistance of instrument against faking and develop methods that can detect individuals who faked their responses. Little studies detecting faking on personality self-reports in the context of employee selection have been conducted as compared to studies examined the effects of faking on the validity of the instrument (Zickar et al., 2004). Hence, studies that explore the method for faking detection are important because it can help employer in selecting qualified job applicant. If the faking detection method has a high level of accuracy, then all fakers are identified and warned, whereas no honest respondents are unnecessarily warned (Lukoff, 2012).

The aim of this study is to compare the accuracy of two statistical mapping techniques in detecting individuals who fake their responses. We hypothesize that there will be a high correlation between participant category as created by research design (performed as research respondents and applicant) and category resulted from statistical analyses. We expect our statistical analyses to have both low rates of false positives (i.e. detecting honest applicants as providing fake responses) and low rates of false negatives (i.e. detecting respondents who provided false responses as being honest). The statistical techniques used will be cluster analysis (CA) and latent class analysis (LCA). These techniques have not been widely applied in the professional/industrial setting, especially as studies have revealed response hoaxes in the context of applicant selection. However, these techniques have been applied to other research topics. For example, cluster analysis was used by Fouladi and Lambert (2005) to identify error response on Likert Scales, and Eid and Zickar (2007) used Mixed Rasch Models to identify yea-saying in respondents. Kankaraš & Moors (in press) used LCA to identify extreme responses in a measure of attitudes.

The research questions posed in the present study are: (1) to what extent can the selected statistical techniques identify individuals who are providing fake responses? (2) Between the two selected statistical clustering techniques (CA and LCA), which is the most effective in detecting individuals who provide fake responses?

Methods

Participants and Procedure

Participants were psychology students (N = 412) at Universitas Gadjah Mada in Yogyakarta, Indonesia. Participants were selected using purposive sampling techniques by the researchers. To create two groups of more standard sizes, the 412-person sample was reduced to 400 using random selection. Participants were then randomly divided into two groups that would receive different instructions regarding how they were to complete the scale: 200 participants were designated as *respondents* and 200 designated as *job applicants*.

The present study had an experimental design. Participants were divided into two groups, *honest* and *faking*. In the first group (henceforth *respondents*), participants were instructed to fill out the scale honestly, with responses exactly representing their actual feelings and experiences. In the second group (henceforth *applicants*), participants were asked to pretend that they were applying for a job, and that their score on the assessment would be used to determine whether they matched the requirements of the prospective job. This modified instruction was hoped to

serve as an inducement for participants to fake their response in order to make a positive impression and thereby pass the selection process. Our hypothesis was that the applicants would be more motivated to make a positive impression, and would therefore be more likely than respondents to give fake responses.

Instrument

This study used a scale measuring the big five personality factors, which was adapted from the Big Five Inventory (BFI-44; John, Donahue, & Kentle, 1991). Responses to the BFI-44 are given on a Likert scale ranging from 1= "strongly agree" to 5= "strongly disagree." This self-report measure asks participants to rate themselves in terms of the five personality factors of extraversion (e.g., actively participate in conversations), agreeableness (e.g., are helpful and do not envy others), conscientiousness (e.g., take their work seriously), neuroticism (e.g., are easy to distress; reverse scored to assess emotional stability), and openness (e.g. willingness to share with others). This procedure is recommended by John, Donahue, & Kentle (1991)

A sample of Indonesian students (N = 185) was used to conduct a validation study, which produced satisfactory reliability: extraversion (alpha coefficient = 0.839), agreeableness (0.789), conscientiousness (0.924), emotional stability (0.848), and openness (0.807). These results are similar to those originally reported by John and Srivastava (1999). They found reliability (α) to range between 0.75 and 0.80, with test–retest reliabilities of 0.80 to 0.90. They also found a high concurrent validity of the BFI with two instruments assessing five factor personality, such as NEO Five Factor Inventory (NEO-FFI, Costa & Mc-Crae, 1989) and Trait Descriptive Adjectives (TDA, Goldberg, 1992), producing an average correlation of 0.83 to 0.91.

Data Analyses

Two statistical techniques— cluster analysis (CA) and latent class analysis (LCA)—were used to explore categories across individuals, based on their pattern responses to the study assessment. Cluster Analysis is an exploratory data analysis tool that sorts sets of objects into new categories based on certain similarities. The usual way to run CA is to first calculate the similarity of all object under analysis (here, individuals' responses), and then to group those objects into clusters based on their similarities (Ruscio & Ruscio, 2008). Most of the work on CA has been done for continuous variables that were assumed to be normally distributed within produced classes (McLachlan, Bean, & Peel, 2002).

Latent class analysis has the same goal as CA, i.e., to find the categories for a set of objects. However, LCA emphasizes the probabilistic approach, identifying unobserved latent categorical variables using the covariance between two or more empirical indicators. This technique can be used to determine the number of types of individuals regarding a certain variable, and can also be used to determine the probability of a given individual belonging to a certain class. This is done by estimating the likelihood of class membership for each response pattern (Thomas, Lanyon, & Millsap, 2009). Latent class analysis is a multivariate technique that attempts to identify distinct classes of individuals for a psychometric scale (see Lazarsfeld & Henry, 1968). Individuals within a single class are assumed to behave similarly for the relevant behavior, while members of different classes are assumed to behave differently.

This study used the confirmatory analysis approach, whereby both CA and LCA were employed to produce two clusters or classes: honest condition and faking condition. Results were therefore given in a 2x2 cross-tabulate form, with two categories representing the type of instruction (honest vs. faking condition) and two categories obtained from the statistical analysis (also honest vs. faking condition). Perfect results would be one cluster exclusively representing participants from the honest condition (respondents), and another cluster exclusively representing participants from the faking condition (applicants).

Results

Descriptive Statistics

Descriptive statistics of the data are presented in Table 1. Personality factor scores were computed from summed scores divided by the number of item for that factor. Since possible item scores ranged from 1 to 5, the hypothetical mean score of each factor is 3. Means scores for the personality measure exceeded 3 for both the honest and the faking condition, meaning that participants' scores for both conditions were above average. Participants' mean scores for those in the faking condition were higher than for those in the honest condition for all five factors of the assessment.

Table 1 also demonstrates that both the minimum and maximum scores for participants in the faking condition were higher than for those in the honest condition. Participants obtained the highest overall scores for extraversion, agreeableness, and conscientiousness: 7 (3.5%) of participants obtained the maximum score for extraversion, 12 participants (6.0%) obtained the maximum score for agreeableness, and 13 (6.5%) obtained the maximum score for conscientious.

Condition	Personality Factors	Min	Max	Mean	SD
Honest	Extraversion	1.38	4.75	3.48	0.53
	Agreeableness	1.44	4.78	3.49	0.53
	Conscientiousness	1.44	4.89	3.34	0.50
	Emotional Stability	1.63	4.13	3.12	0.37
	Openness	1.10	0.90	3.59	0.48
Faking	Extraversion	2.50	5.00	3.80	0.54
	Agreeableness	2.67	5.00	3.73	0.70
	Conscientiousness	2.22	5.00	3.75	0.69
	Emotional Stability	2.38	4.50	3.34	0.47
	Openness	2.80	4.90	3.88	0.44

Note. *Min* = the lowest obtained score; *Max* = The highest obtained score; *SD* = Standard deviation.

Table. 1 Descriptive statistics for the five-factor personality assessment: Comparing participants under honest and faking conditions

The *t*-test for mean comparison is presented in Table 2 (p. 16), and indicates significant differences between the honest and the faking conditions for all personality sub-scores. The mean largest difference was found for conscientiousness (.41), and the smallest difference was found for emotional stability (.22).

Descriptive statistics of a comparison of composite scores found that only 13 people (6.5%) of applicants received the maximum possible score for conscientiousness, which is related to job performance. The low number of applicants who obtained the maximum score for this factor indicates that participants gave fake responses—while getting the maximum score may indicate that the applicant is an ideal prospective employee, test-takers know that maximum scores can raise red flags regarding dishonesty. Such a red flag can result in employer suspicion that the applicant's score is fake, and the employer

Table 2. Comparison of sub-test scores for the five factors of personality: Comparing participants under honest and faking conditions

Personality Factors	Mean Diff.	t-value	Effect Size
Extraversion	0.32**	-6.076	.53
Agreeableness	0.24**	-3.871	.40
Conscientiousness	0.41**	-6.921	.65
Emotional Stability	0.22**	-5.226	.43
Openness	0.29**	-6.376	.39
Note. ** = $p < 0.001$			

may then seek to verify the applicant's responses more carefully, (e.g., via interviewing), thus diminishing the applicant's chances of getting the job. These results are consistent with a previous study conducted by Zickar and Robie (1999), which found that applicants were careful to not obtain the maximum score due to the threat of these negative consequences.

Categorization Analysis

The techniques of CA and LCA provided information about differences in the response patterns in participants from the honest and faking conditions. Figure 1 shows an example response patterns for eight items measuring extraversion, denoted by eight points on the X-axis. The Y-axis shows the probability of endorsement, ranging from 0 to 1. Almost all participants in Category 1 (honest) had a similar probability (.2–.4) of endorsing all items on the scale. In contrast, participants in Category 2 (faking) had an inconsistent probability (.2–.6) of endorsing all scale items .
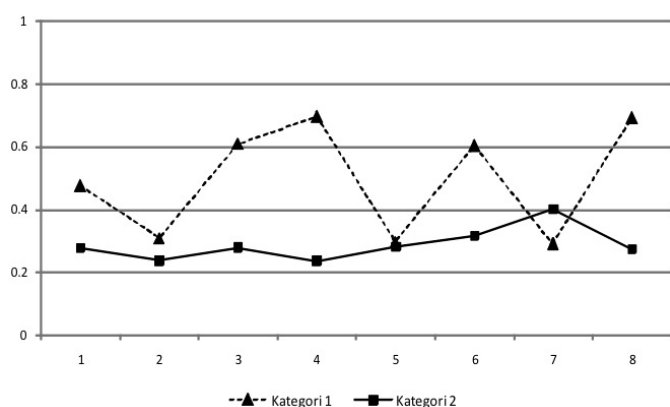

Figure 1. An example of the different response patterns for Category 1 (honest) and Category 2 (faking) participants: Eight items on the sub-scale measuring extraversion.

Based on the response pattern, Category 2 can be called a unique group, since members of this group only endorsed specific items. This unique response pattern could represent either a true trait or a strategy to make a good impression. Because the data being analyzed consisted of both honest- and faking-condition participants, we assume that this unique pattern represents a strategy on the part of participants to give a positive impression. These results will support our initial hypotheses if all members of Category 1 are from the honest condition, while all members of Category 2 are from the faking condition.

Tables 3 and 4 depict the results of statistical analyses of participant classification. Most participants from the honest condition fell into the *normal* category while most participants from the faking condition fell into the *unique* category. However, this is not always true because most of the faking participants for agreeableness in CA and most of the faking participants for nearly all personality factors in LTA fell into the normal

Personality Factors	Type of Instruction	Cluster Membership	
		Normal Pattern	Unique Pattern
Extraversion	Honest	119 (60)	81 (41)
	Faking	95 (48)	105 (53)
Agreeableness	Honest	108 (54)	92 (46)
	Faking	104 (52)	96 (48)
Conscientiousness	Honest	138 (69)	62 (31)
	Faking	99 (49)	101 (51)
Emotional Stability	Honest	113 (57)	87 (43)
	Faking	88 (44)	112 (56)
Openness	Honest	121 (61)	79 (40)
	Faking	62 (31)	138 (69)
Note: Numbers in parentheses indicate percent. Italicized numbers indicate the number of participants who performed as an applicant detected provides a unique response.			

Table 3. Cluster Analysis results: Comparing the number of participants assigned normal and unique patterns according to instruction type

Personality Factors	Role of Participants	Class Membership		
		Normal Pattern	Unique pattern	N
Extraversion	Honest	150 (75)	50 (25)	200
	Faking	113 (57)	87 (44)	200
Agreeableness	Honest	135 (68)	65 (33)	200
	Faking	115 (58)	85 (43)	200
Conscientiousness	Honest	175 (88)	25 (13)	200
	Faking	104 (58)	96 (48)	200
Emotional Stability	Honest	158 (79)	42 (21)	200
	Faking	93 (47)	107 (54)	200
Openness	Honest	165 (83)	35 (18)	200
	Faking	127 (64)	73 (37)	200
Note: Numbers in parentheses indicate percent. Italicized numbers indicate the number of participants who performed as an applicant detected provides a unique response.				

Table 4. Latent Class Analysis results: Comparing the number of participants assigned normal and unique patterns according to instruction type

category. For example, CA of the factor *extraversion* gave false positives for 81 participants (41%), since the answers of these participants were in a unique pattern. For the same factor, 95 participants (48%) were given false negatives, and their faked responses were not identified.

Some possible causes of these errors in prediction will be explained in the discussion. However, comparing the results from CA and LCA, CA had a higher false-positive rate than LCA, with the number of participants who were detected as giving fake responses being higher. On average, 55% of participants from the honest condition were detected as presenting a unique response via CA, while only 45 % of participants from the honest condition were detected as presenting a unique response via LCA.

Both CA and LCA were found to have a moderate capacity (a maximum of 70 %) to detect fake responses in applicants. The remaining 30 % fell into the normal response category, representing applicants who generated response pattern similar

to those of respondents. Based on their response pattern, this appears to represent a group of participants who gave honest responses despite the strong motivation to fake their answers. According to Zickar and colleagues (2004), these individuals could have a wide range of reasons for their behavior—they could be extremely virtuous, naïve, poor at impression management, or unmotivated to pursue the job. Using mixed-Rasch Modeling, Zickar and colleagues found that there was considerable overlap between these groups of participants: 7.2–22.9% of participants in the honest group were detected as being participants who faked their responses, while, a large number of participants in the faking group were detected as giving honest responses. Their results suggest that several types of responses exist even when employing a controlled experimental design.

Correlation between Condition and Analyses Results

A subsequent analysis was performed to examine the correlation between condition categories (honest vs. faking) and analysis category (normal vs. unique pattern). The data being analyzed are represented in a 2x2 crosstab, since each category consist of two nominal codes. Table 5 showed the phi-correlation between condition categories (honest vs. faking) and response pattern categories (normal vs. unique). The correlation values indicate that participant categorization using LCA was higher than using CA. Using chi-square statistics to test whether categorization based on condition categories and statistical categorization support these findings. Significant correlations were found on the all factors using LCA, suggesting that LCA outperform CA for identifying individuals who faked their responses.

Personality Factors	Cluster Analysis		Latent Class Analysis	
	Phi.	Chi-square	Phi	Chi-square
Extraversion	0.12*	15.20***	0.20**	5.788*
Agreeableness	0.02	4.27*	0.10*	0.16
Conscientiousness	0.20**	59.72***	0.37**	15.749***
Emotional Stability	0.13*	45.19***	0.34**	6.250*
Openness	0.30**	18.32***	0.21**	35.06***
Note. * = p < 0.05 ; ** = p < 0.01; ** = p < 0.001				

Table 5. Phi and Chi-square coefficient that Indicate Relationship between Type of Instruction Categorization and Statistical Analysis Results

Correlation values ranged over low–moderate levels, indicating that neither CA nor LCA performed well at detecting fake responses. Overall, our analyses demonstrated that we were successful in inducing participants to make fake responses through the prompt of asking them to pretend to be job applicants, with mean *applicant* scores higher than mean *respondent* scores. Our statistical analyses had a moderate ability to identify participants who were giving fake responses, although their performance varied among the five personalities factors. Emotional stability and openness proved to be the best sub-tests for detecting fake responses, because the lowest false-negatives were obtained for these two factors. Finally, we found that CA had higher response accuracy in identifying fake responses than LCA.

Discussion

This study aimed to identify individuals who faked their responses on a self-report personality measure commonly used for personnel selection. Participants who were instructed to behave as job applicants were assumed to be more likely to fake responses due to their intention to pass the employee selection

process. Study findings supported this hypothesis, as mean scores from *applicants* were higher than those of *respondents* for all five personality factors. These results indicate that our effort to induce participants to fake their response was quite successful. These results are consistent with previous studies with similar designs.

Our results demonstrate that participants who behaved as job applicants had various ways to make a positive impression and qualify for the job. Previous research supports this finding, as (McFarland & Ryan, 2000) found that individuals have different reasons for faking or responding honestly: some participants respond honestly for ethical reasons, or for fear of being found out, while other respondents feel no compunction about distorting their responses, and try to give the answers that they believe will be most likely to result in them being employed (Zickar & Robie, 1999). Although the mean score of *applicants* was significantly higher than *respondents* in the present study, the numbers of applicants who obtained a score below the mean score of respondents was notably high. For example, for *conscientiousness*, 67 applicants (34%) obtained a score below the mean score of respondents. This result also indicates there were differences in the magnitude and strategies of faking responses for the assessment used.

The differences in the magnitude and strategies of faking were also evidenced by the similarity of response patterns uncovered by the two statistical analyses. Two categories (i.e., normal vs. unique) were uncovered in the *respondents* using either CA or LCA, indicating a rather similar response pattern. On the other hand, the two categories revealed in *applicants* were different, as the result of different faking strategies. At a certain point members of the applicant group got higher scores for some sub-scales than for others.

Members of the applicant group not only tried to avoid getting the maximum possible score; there were also differences among participants as to how they perceived certain item statements: Overall, for items that they perceived to be most relevant to the job, they tried to obtain the maximum score; in contrast, for items that they perceived to be irrelevant to the job, they avoided getting a maximum score. These findings support the previous finding that applicants' response patterns are more heterogeneous than those of respondents (Zickar et al., 2004). In addition to differences in amounts or in faking strategies, there are several additional variables that might be valuable to control in future studies in this area. For example, previous research suggests that individuals with high *self-esteem* tend to see themselves more positively, and therefore tend to get higher scores on other personality measures (Tunnell, 1980). Another promising variable is *attitudes toward faking* (McFarland & Ryan, 2006), as well as situational factors, such as *stringency of process selection* (i.e., the chance that applicants can pass the selection process (Robie, 2006).

Regarding analysis techniques, CA emphasizes the similarity of responses, using both distance and a correlation yield group categorization that better matched the true categorization of participants (i.e., applicants vs. respondents) than LCA. Item analysis showed that the parameter of item threshold parameter was equivalent for most of the measure; because LCA uses item threshold as the basis of analysis, this technique was not sensitive to detecting unusual response patterns, resulting in fewer applicants being detected as having unique response patterns than when using CA. A second difference between CA and LCA was their different performance in detecting fake responses. Cluster Analysis aims to identify a manifest classification, with objects already assigned to exactly one cluster. In

contrast, LCA assumes a latent grouping variable, meaning that all participants belong to all produced latent classes, with varying degrees of probability (Rost, 2003). In general, both statistical techniques used in this study to differentiate response patterns were moderately successful in identifying fake responses. A promising technique to explore in future studies in this area would be Rasch-Mixed Modeling, which has already been implemented to detect a variety of response styles, and has been shown to be appropriate in analyzing data from personality assessments (Zickar & Robie, 1999).

Conclusion

Our findings highlight a crucial concern for researchers who use instruction manipulation to prompt fake responses: while participants who were asked to imagine themselves as job applicants did tend to give more fake responses, this was not univer-

sally true, and they may have answered honestly or dishonestly for a number of distinct motivations. Differences in the type of instruction, participant characteristics, and type of personality assessment should be taken into account. Another limitation of the present study was the use of a two-class constrained model for analysis, since, as exploratory analyses, both CA and LCA may yield more than two classes that fit the data. By leaving the number of classes open, individuals with more unique characteristics (e.g. moderate level of faking, safely faking) can be accommodated by the model. However, in the present study, both analyses were restricted to two classes so that their results would correlate with the existing categories of respondents and applicants; unfortunately, this restriction resulted in heterogeneous classes contaminated with individuals who had other characteristics than those defining their class. We therefore recommend that future research leave the number of classes to be generated by CA or LCA open.

References

Andrews, P., & Meyer, R. G. (2003). Marlowe–Crowne Social Desirability Scale and short Form C: Forensic norms. Journal of Clinical Psychology, 59(4), 483-492. doi: 10.1002/jclp.10136

Bachman, J. G., & O'Malley, P. M. (1984). Yea-Saying, Nay-Saying, and Going to Extremes: Black-White Differences in Response Styles. Public Opinion Quarterly, 48(2), 491-509. doi: 10.1086/268845

Barrick, M. R., & Mount, M. K. (1991). the Big Five Personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44(1), 1-26.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A Meta-Analytic Investigation of Job Applicant Faking on Personality Measures. International Journal of Selection and Assessment, 14(4), 317-335. doi: 10.1111/j.1468-2389.2006.00354.x

Burns, G. N., Christiansen, N. D., Griffith, R. L., & Peterson, M. H. (2006). Sensitive or senseless: On the use of social desirability measures in selection and assessment. In R. L. Griffith & M. H. Peterson (Eds.), A closer examination of applicant faking behavior. Greenwich, CT: Information Age.

Converse, P. D., Peterson, M. H., & Griffith, R. L. (2009). Faking on Personality Measures: Implications for selection involving multiple predictors. International Journal of Selection and Assessment, 17(1), 47-60. doi: 10.1111/j.1468-2389.2009.00450.x

Costa, P. T., & McCrae, R. R. (1989). The NEO-PI/NEO-FFI Manual supplement. Odessa, FL: Psychological Assessment Resources.

Dilchert, S., Ones, D. S., Viswesvaran, C., & Deller, J. (2006). Response distortion in personality measurement: Born to deceive, yet capable of providing valid self-assessments? Psychology Science, 48(209-225).

Edwards, A. L. (1957). The social desirability variable in personality assessment and research. New York: Dryden.

Eid, M., & Zickar, M. J. (2007). Detecting response styles and faking in personality and organizational assessments by mixed Rasch models. In M. Von Davier & C. Carstensen (Eds.), Multivariate and Mixture Distribution Rasch Models. New York: Springer.

Elhai, J. D., & Frueh, B. C. (2001). Subtypes of clinical presentations in malingerers of posttraumatic stress disorder: an MMPI-2 cluster analysis. Assessment, 8(1), 75-84.

Ferrando, P. J., & Anguiano-Carrasco, C. (2013). A Structural Model–Based Optimal Person-Fit Procedure for Identifying Faking. Educational and Psychological Measurement, 73(2), 173-190. doi:

10.1177/0013164412460049

Fouladi, R. T., & Lambert, R. (2005). Two-step cluster analysis and detecting misresponding to Likert scales with given correlated measurement error. Paper presented at the American Educational Research Association Annual Meeting, Montreal, Canada.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. Psychological Assessment, 4(1), 26-42. doi: 10.1037/1040-3590.4.1.26

Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2008). Detecting Malingering in Traumatic Brain Injury and Chronic Pain: A Comparison of Three Forced-Choice Symptom Validity Tests. The Clinical Neuropsychologist, 22(5), 896-918. doi: 10.1080/13854040701565208

Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. Personnel Review, 36(3), 341-355.

Hartman, N. S., & Grubb, W. L. (2011). Deliberate Faking on Personality and Emotional Intelligence Measures. Psychological Reports, 108(1), 120-138.

Heinrich, C. U., & Borkenau, P. (1998). Deception and deception detection: The role of cross-modal inconsistency. Journal of Personality, 66(5), 687-712. doi: 10.1111/1467-6494.00029

Holden, R. R., & Book, A. S. (2009). Using hybrid Rasch-latent class modeling to improve the detection of fakers on a personality inventory. [doi: DOI: 10.1016/j.paid.2009.02.024]. Personality and Individual Differences, 47(3), 185-190.

Jackson, D. N. (1984). Personality research form manual. Port Huron: Research Psychologists Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory--Versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

Kankaraš, M., & Moors, G. (in press). Measurement equivalence and extreme response bias in the comparison of attitudes across europe. European Journal of Research Methods for the Behavioral and Social Sciences.

Lazarsfeld, P. F., & Henry, N. W. (1968). Latent structure analysis. Boston: Houghton Mifflin.

Levashina, J., Morgeson, F. P., & Campion, M. A. (2012). Tell me some more: Exploring how verbal ability and item verifiability influence responses to biodata questions in a high-stakes selection context. Personnel Psychology, 65(2), 359-383. doi: 10.1111/j.1744-6570.2012.01245.x

Lukoff, B. (2012). Is faking inevitable? In M. Ziegler, C. McCann &

R. D. Roberts (Eds.), New perspectives on faking in personality assessments. Oxford: Oxford University Press.

Marlowe, D., & Crowne, D. P. (1961). Social desirability and response to perceived situational demands. Journal of Consulting Psychology, 25, 109-115.

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. Journal of Applied Psychology, 85(5), 812-821. doi: 10.1037/0021-9010.85.5.812

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. Journal of Applied Social Psychology, 36(4), 979-1016.

McLachlan, G. J., Bean, R. W., & Peel, D. (2002). A mixture model-based approach to the clustering of microarray expression data. Bioinformatics, 18(3), 413-422. doi: 10.1093/bioinformatics/18.3.413

Morgeson, F. P., Campion, M. a., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the Use of Personality Tests in Personnel Selection Contexts. Personnel Psychology, 60(3), 683-729. doi: 10.1111/j.1744-6570.2007.00089.x

Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: a study using the Neo Personality Inventory and the Myers–Briggs Type Indicator. European Journal of Personality, 17(1), 79-94. doi: 10.1002/per.471

Mueller-Hanson, R., Heggestad, E. D., & Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. Journal of Applied Psychology, 88(2), 348-355. doi: 10.1037/0021-9010.88.2.348

Paulhus, D. L. (1984). Two component models of socially desirable responding. Journal of Personality and Social Psychology, 46, 598-609.

Robie, C. (2006). Effects of perceived selection ratio on personality test faking. Social Behavior and Personality, 34(10), 1233-1244.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. Journal of Applied Psychology, 83(4), 634-644. doi: 10.1037/0021-9010.83.4.634

Rost, J. (2003). Latent Class Analysis. In R. Fernández Ballesteros (Ed.), Encyclopedia of psychological assessment. Thousand Oaks, CA: SAGE Publications.

Ruscio, J., & Ruscio, A. M. (2008). Categories and Dimensions: Advancing Psychological Science Through the Study of Latent Structure. CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE, 17(3), 203-207. doi: 10.1111/j.1467-8721.2008.00575.x

Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. Personnel Psychology, 52(2), 359-392. doi: 10.1111/j.1744-6570.1999.tb00165.x

Sellbom, M., Toomey, J. A., Wygant, D. B., Kucharski, L. T., & Duncan, S. (2010). Utility of the MMPI-2-RF (Restructured Form) validity scales in detecting malingering in a criminal forensic setting: A known-groups design. Psychological Assessment, 22(1), 22-31. doi: 10.1037/a0018222

Thomas, M. L., Lanyon, R. I., & Millsap, R. E. (2009). Validation of diagnostic measures based on latent class analysis: A step forward in response bias research. Psychological Assessment, 21(2), 227-230. doi: 10.1037/a0015693

Tunnell, G. (1980). Intraindividual consistency in personality-assessment - the effect of self-monitoring. Journal of Personality, 48(2), 220-232.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. Educational and Psychological Measurement, 59(2), 197-210. doi: 10.1177/00131649921969802

Weiner, J. A., & Gibson, W. M. (2000). Practical Effects of Faking on Job Applicant Attitude Test Scores. Paper presented at the Annual meeting of the Society for Industrial and Organizational Psychology, New Orleans, April 2000, New Orleans.

Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. Organizational Research Methods, 7(2), 168-190.

Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. Journal of Applied Psychology, 84(4), 551-563.

Authors

Wahyu Widhiarso, MA
Faculty of Psychology
Universitas Gadjah Mada (UGM)
Jl. Humaniora No. 1 Bulaksumur
Yogyakarta, Indonesia

Email: wahyu_psy@ugm.ac.id

Fathul Himam, Ph.D
Faculty of Psychology
Universitas Gadjah Mada (UGM)
Jl. Humaniora No. 1 Bulaksumur
Yogyakarta, Indonesia

Email: fhimam@ugm.ac.id

Appendix

Items of BFI-44 (John & Srivastava, 1999).
1. Is talkative
2. Tends to find fault with others (rev)
3. Does a thorough job
4. Is depressed, blue
5. Is original, comes up with new ideas
6. Is reserved (rev)
7. Is helpful and unselfish with others
8. Can be somewhat careless (rev)
9. Is relaxed, handles stress well (rev)
10. Is curious about many different things
11. Is full of energy
12. Starts quarrels with others (rev)
13. Is a reliable worker
14. Can be tense
15. Is ingenious, a deep thinker
16. Generates a lot of enthusiasm
17. Has a forgiving nature
18. Tends to be disorganized (rev)
19. Worries a lot
20. Has an active imagination
21. Tends to be quiet (rev)
22. Is generally trusting
23. Tends to be lazy (rev)
24. Is emotionally stable, not easily upset (rev)
25. Is inventive
26. Has an assertive personality
27. Can be cold and aloof (rev)
28. Perseveres until the task is finished
29. Can be moody
30. Values artistic, aesthetic experiences
31. Is sometimes shy, inhibited (rev)
32. Is considerate and kind to almost everyone
33. Does things efficiently
34. Remains calm in tense situations (rev)
35. Prefers work that is routine (rev)
36. Is outgoing, sociable
37. Is sometimes rude to others (rev)
38. Makes plans and follows through with them
39. Gets nervous easily
40. Likes to reflect, play with ideas
41. Has few artistic interests (rev)
42. Likes to cooperate with others
43. Is easily distracted (rev)
44. Is sophisticated in art, music, or literatures

Note. (rev) denotes reverse scored items.
Extraversion: 1, 6, 11, 16, 21, 26, 31, 36
Agreeableness: 2, 7, 12, 17, 22, 27, 32, 37, 42
Conscientiousness: 3, 8, 13, 18, 23, 28, 33, 38, 43
Neuroticism: 4, 9, 14, 19, 24, 29, 34, 39
Openness: 5, 10, 15, 20, 25, 30, 35, 40, 41, 44